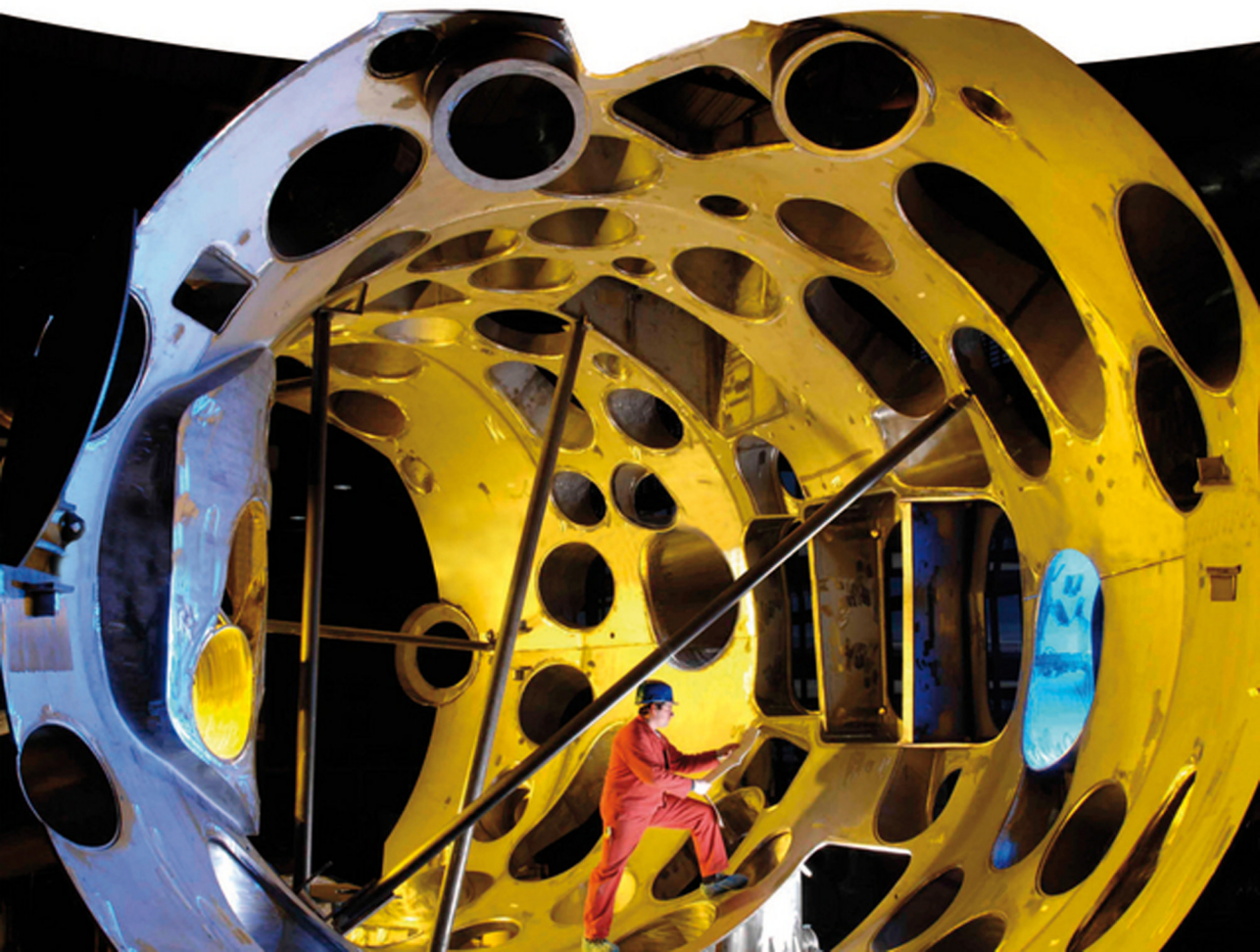
WILEY-VCH

Edited by Reinhard Stock

Encyclopedia of Nuclear Physics and its Applications



Encyclopedia of Nuclear Physics and its Applications

Edited by
Reinhard Stock

Related Titles

Martin, B.R.
Nuclear and Particle Physics - An Introduction

2006
ISBN: 978-0-470-03547-4

Lilley, J.
Nuclear Physics – Principles & Applications

2001
ISBN: 978-0-471-97935-7

Lieser, K.H.
Nuclear and Radiochemistry
Fundamentals and Applications

2nd Edition
2001
ISBN: 978-3-527-30317-5

Stacey, W.M.
Fusion Plasma Physics

2005
ISBN: 978-3-527-40586-2

Turner, J.E.
Atoms, Radiation, and Radiation Protection

3rd Edition
2007
ISBN: 978-3-527-40606-7

Gläser, M., Kochsiek, M. (eds.)
Handbook of Metrology

2010
ISBN: 978-3-527-40666-1

Stacey, W.M.
Nuclear Reactor Physics

2nd Edition
2007
ISBN: 978-3-527-40679-1

Stock, R. (ed.)
Encyclopedia of Applied High Energy and Particle Physics

2009
ISBN: 978-3-527-40691-3

Bohr, H.G. (ed.)
Handbook of Molecular Biophysics
Methods and Applications

2009
Print ISBN: 978-3-527-40702-6

Andrews, D.L. (ed.)
Encyclopedia of Applied Spectroscopy

2009
ISBN: 978-3-527-40773-6

Krücken, R., Dilling, J.
Experimental Methods for Exotic Nuclei

2014
ISBN: 978-3-527-41038-5

Encyclopedia of Nuclear Physics and its Applications

Edited by
Reinhard Stock

WILEY-VCH
Verlag GmbH & Co. KGaA

Editor

Reinhard Stock
Universität Frankfurt
stock@ikf.uni-frankfurt.de
Germany

Cover
Outer vessel of fusion experiment Wendelstein 7-X (Copyright: IPP, Wolfgang Filser).

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at ⟨http://dnb.d-nb.de⟩.

Print ISBN: 978-3-527-40742-2
ePDF ISBN: 978-3-527-64927-3
ePub ISBN: 978-3-527-64926-6
Mobi ISBN: 978-3-527-64925-9
oBook ISBN: 978-3-527-64924-2

Cover Design Adam Design, Weinheim
Composition Laserwords Private Ltd., Chennai, India
Printing and Binding Markono Print Media Pte Ltd, Singapore

Printed in Singapore
Printed on acid-free paper

Contents

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

Preface

With this volume, WILEY-VCH continues a series of reeditions of the ''Wiley Encyclopedia of Applied Physics'' that was edited by Professor G. L. Trigg in the early 1980s. It united more than 600 monographic articles that embraced physics in the widest conceivable way, from ''Accelerators'' via ''Oceanography'' to ''Xerography.'' The individual articles were addressed at a broad scientific/technical readership, written in a didactical style combined with a focus on new and exciting developments.

About 25 years have elapsed since this monumental undertaking, and the fields of physics have undergone substantial evolution. In this volume, we revisit Nuclear Physics, an extremely broad field of multifaceted fundamental research, and a wealth of application.

Nuclear Physics phenomena present a universal testing ground for processes and structures that are related to the three fundamental forces/interactions: strong, electromagnetic, and weak. In a certain sense, it could play the role of a paradigmatic realization of the entire Standard Model of fundamental interaction, if it was not for the fact that the strong force manifestations in nuclei and nuclear interactions represent the remotest realizations of the elementary strong force that is described by the quantum chromodynamics (QCD) theory – in the sense that nuclear forces are a strictly nonperturbative QCD effect of higher order that is essentially inaccessible to analytic or even numerical methods. This is why we do not have, even today, a ''nuclear structure QCD approach.'' This QCD approach exists, however, in the consideration of nucleon structure. The many-nucleon patterns of nuclear properties and reactions can be confronted with effective nucleon–nucleon forces that have been familiar for many decades but can now be, at least qualitatively, understood as a QCD analogy to, for example, van der Waals-like forces. At the limit of ultrarelativistic accelerator energies, nuclear collisions come under explicit governance of QCD, creating the primordial state of matter consisting of quarks and gluons: the so-called quark–gluon plasma. This is one of the present day fundamental research themes of the field, which, furthermore, features some of the most intricate many-body problems known to science.

Applications of Nuclear Physics could fill many such books as this. The selection of topics presented here is guided loosely by their importance in the intuitive view of the scientifically and technically interested public. We have selected topics ranging

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

from genuinely fundamental interdisciplinary science (such as nuclear and neutrino astrophysics, neutron stars and supernovae) via nuclear medicine and matter analysis finally to the techniques of nuclear power generation and its future. We have omitted nuclear weapons, not for the sake of political correctness (we build them, so we should in principle also understand them) but for the reason that classification forbids reasonable monographic public presentation. In any case, the presentation chosen here cannot be all-encompassing. However, to accomplish a certain level of comprehensiveness, we have included a number of updated articles from the previous volume edited in the same style (Encyclopedia of Applied High Energy and Particle Physics, Wiley-VCH 2009), with a closely related dedication.

I wish to thank all the authors for their outstanding devotion. Special thanks go to Dr. A. Mueller of IN2P3 for his very essential help with the entire section on nuclear power, which resulted in a unique spectrum of presentations.

Special thanks go to my editors at WILEY, Mrs. Vera Palmer and Anja Tschörtner, for their unbending encouragement and patience.

Frankfurt *Reinhard Stock*
December 2012

List of Contributors

Hamid Aït Abderrahim
SCK CEN
Boeretang 200
2400 Mol
Belgium

Jacques Barbet
GIP Arronax
1 rue Arronax
44817 Saint-Herblain
France

and

Nantes-Angers Cancer Research Center
UMR 892 INSERM and
UMR 6299 CNRS
8 quai Moncousu
44007 Nantes
France

Carlos A. Bertulani
Texas A&M University-Commerce
Department of Physics and Astronomy
2600 South Neal Street
Commerce, TX 75428
USA

Joseph Bisognano
University of Wisconsin-Madison
Engineering Physics Department and
Synchroton Radiation Center
3731 Schneider Drive
Stoughton, WI 53589
USA

Christoph Blume
Goethe-Universität Frankfurt
Institut für Kernphysik
Max-von-Laue-Str. 1
60438 Frankfurt am Main
Germany

Rudolf Bock
GSI Helmholtzzentrum für
Schwerionenforschung GmbH
Planckstraße 1
64291 Darmstadt
Germany

Allen Caldwell
Max Planck Institute for Physics
(Werner Heisenberg Institute)
Föhringer Ring 6
80805 Munich
Germany

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

Jean-François Chatal
GIP Arronax
1 rue Arronax
44817 Saint-Herblain
France

Nicolas Chouin
LUNAM Université
Oniris
AMaROC
44307 Nantes
France

François Davodeau
Nantes-Angers Cancer Research Center
UMR 892 INSERM and UMR 6299 CNRS
8 quai Moncousu
44007 Nantes
France

Didier De Bruyn
SCK CEN
Boeretang 200
2400 Mol
Belgium

Hans Geissel
GSI Helmholtzzentrum für Schwerionenforschung
Planckstr. 1
64291 Darmstadt
Germany

and

Justus-Liebig-Universität
Gießen
Physics Department
Heinrich-Buff Ring 14
35392 Gießen
Germany

Dietrich Habs
Ludwig-Maximilians-Universität München
Fakultät für Physik
Am Coulomb-Wall 1
85748 Garching
Germany

and

Max-Planck-Institute of Quantum Optics
Hans-Kopfermann-Strasse 1
85748 Garching
Germany

Ferid Haddad
GIP Arronax
1 rue Arronax
44817 Saint-Herblain
France

and

SUBATECH
Université de Nantes
Ecole des Mines de Nantes
CNRS/IN2P3
4 rue A. Kastler
44307 Nantes
France

Wick C. Haxton
University of California
Department of Physics
MC-7300
Berkeley, CA 94720
USA

and

Lawrence Berkeley National Laboratory
Nuclear Science Division
Berkeley, CA 94720
USA

Ragnar Hellborg
Lund University
Department of Physics
Division of Nuclear Physics
Sölvegatan 14
22362 Lund
Sweden

Ingo Hofmann
GSI Helmholtzzentrum für
Schwerionenforschung
Planckstraße 1
64291 Darmstadt
Germany

Sigurd Hofmann
GSI Helmholtzzentrum für
Schwerionenforschung
Department of Nuclear Structure,
Astrophysics and Reactions
Planckstraße 1
64291 Darmstadt
Germany

Mark Huyse
University of Leuven (K.U.Leuven)
Instituut voor Kern-en Stralingsfysica
Departement Natuurkunde en
Sterrenkunde
Celestijnenlaan 200 D
3001 Leuven
Belgium

Robert Jameson
Goethe University Frankfurt
Institute of Applied Physics
Max-von-Laue-Str. 1
60438 Frankfurt am Main
Germany

Jan Jolie
University of Cologne
Institute for Nuclear Physics
Zülpicher Straße 77
50937 Cologne
Germany

Ulli Köster
Institut Laue-Langevin
6 rue Jules Horowitz
38042 Grenoble
France

Françoise Kraeber-Bodéré
Nantes-Angers Cancer Research Center
UMR 892 INSERM and UMR 6299
CNRS
8 quai Moncousu
44007 Nantes
France

and

Nuclear Medicine Department
University Hospital-ICO-INSERM UMR
892
place Alexis Ricordeau
44093 Nantes
France

Gerhard Kraft
GSI Helmholtzzentrum für
Schwerionenforschung
Planckstraße 1
64291 Darmstadt
Germany

Pierre Lapostolle†
Neuilly-sur-Seine
France

Hartmut Lauer
RWE Power AG
Biblis Nuclear Power Plant
vice president rtd.
Biblis
Germany

Matthias Liebendörfer
University of Basel
Department of Physics
Klingelbergstrasse 82
4056 Basel
Switzerland

Gottfried Münzenberg
GSI Helmholtzzentrum für
Schwerionenforschung
Planckstr. 1
64291 Darmstadt
Germany

and

Manipal University
MARG
Madhav Nagar
Manipal 576104
Karnataka
India

Sidney Michiels
SCK CEN
Boeretang 200
2400 Mol
Belgium

Mark Riley
Florida State University
Department of Physics
214 Keen Building
Tallahassee, FL 32306-3016
USA

Jürgen Schaffner-Bielich
Johann Wolfgang Goethe Universität
Institut für Theoretische Physik
Max von Laue-Straße 1
60438 Frankfurt am Main
Germany

Hendrik Schatz
Michigan State University
National Superconducting Cyclotron
Laboratory and Department of Physics
and Astronomy
640 S. Shaw Lane
East Lansing, MI 48824-1321
USA

Klaus Schreckenbach
Technische Universität München
Physik Department E21 and FRM II
James-Franck-Strasse
85747 Garching
Germany

Thomas Schulenberg
Karlsruhe Institute of Technology
Institute for Nuclear and
Energy Technologies
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
Germany

John Simpson
STFC Daresbury Laboratory
Keckwick Lane
Daresbury
Warrington WA4 4AD
UK

Göran Skog
Lund University
Department of Geology
Sölvegatan 12
22362 Lund
Sweden

Weston M. Stacey
Nuclear and Radiological Engineering
Georgia Institute of Technology
770 State Street
Atlanta, GA 30332
USA

Kristina Stenström
Lund University
Department of Physics
Division of Nuclear Physics
Sölvegatan 14
22362 Lund
Sweden

Friedrich-Karl Thielemann
University of Basel
Department of Physics
Klingelbergstrasse 82
4056 Basel
Switzerland

Gert Van den Eynde
SCK CEN
Boeretang 200
2400 Mol
Belgium

Piet Van Duppen
University of Leuven (K.U.Leuven)
Instituut voor Kern-en Stralingsfysica
Departement Natuurkunde en
Sterrenkunde
Celestijnenlaan 200 D
3001 Leuven
Belgium

Friedrich Wagner
Max-Planck-Institut für Plasmaphysik
Wendelsteinstr. 1
17489 Greifswald
Germany

and

St. Petersburg State Polytechnical
University Research Laboratory for
Advanced Tokamak Physics
Polytechnicheskaya 29
19521 St. Petersburg
Russia

David C. Weisser
Australian National University
Department of Nuclear Physics
Research School of Physics and
Engineering
Canberra, ACT 0200
Australia

Part A
Fundamental Nuclear Research

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

1
Nuclear Structure

Jan Jolie

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

1.1 Introduction

The study of nuclear structure today encompasses a vast territory from the study of simple, few-particle systems to systems with close to 300 particles, from stable nuclei to the short-lived exotic nuclei, from ground-state properties to excitations of such energy that the nucleus disintegrates into substructures and individual constituents, from the strong force that hold the atomic nucleus together to the effective interactions that describe the collective behavior observed in many heavy nuclei.

After the discovery of different kinds of radioactive decays, the discovery of the structure of the atomic nucleus begins with the fundamental paper by Ernest Rutherford [1], in which he explained the large-angle alpha (α)-particle scattering from gold that had been discovered earlier by Hans W. Geiger and Ernest Marsden. Indeed, Rutherford shows that the atom holds in its center a very tiny, positively charged nucleus that contains 99.98% of the atomic mass. In 1914, Henry Moseley [2, 3] showed that the nuclear charge number Z equaled the atomic number. Using the first mass separators, Soddy [4] was able to show that one chemical element could contain atomic nuclei with different masses, forming different isotopes. With the availability of α-sources, due to the works of the Curies in Paris, Rutherford [5] was able to perform the first nuclear reactions on nitrogen. The first attempt at understanding the relative stability of nuclear systems was made by Harkins and Majorsky [6]. This model, like many others of the time, consisted of protons and electrons. In 1924, Wolfgang Pauli [7] suggested that the optical hyperfine structure might be explained if the nucleus had a magnetic dipole moment, while later Giulio Racah [8] investigated the effect on the hyperfine structure if the nuclear charge were not spherically symmetric – that is, if it had an electric quadrupole moment.

All of these structure suggestions occurred before James Chadwick [9] discovered the neutron, which not only explained certain difficulties of previous models (e.g., the problems of the confinement of the electron or the spins of light nuclei), but opened the way to a very rapid expansion of our knowledge of the structure of the nucleus. Shortly after the discovery of the neutron, Heisenberg [10] proposed that the proton and neutron are two states of the nucleon classified by a new spin quantum number, the isospin. It may be difficult to believe today, 60 years after Chadwick's discovery, just how rapidly our knowledge of the nucleus increased in the mid-1930s.

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

Hans A. Bethe's review articles [11, 12], one of the earliest and certainly the best known, discuss many of the areas that not only form the basis of our current knowledge but that are still being investigated, albeit with much more sophisticated methods.

The organization here will begin with general nuclear properties, such as size, charge, and mass for the stable nuclei, as well as half-lives and decay modes (α, β, γ, and fission) for unstable systems. Binding energies and the mass defect lead to a discussion of the stability of systems and the possibility of nuclear fusion and fission. Then follow details of the charge and current distributions, which, in turn, lead to an understanding of static electromagnetic moments (magnetic dipole and octupole, electric quadrupole, etc.) and transitions. Next follows the discussion of single-particle and collective levels for the three classes of nuclei: even–even, odd-*A*, and odd–odd (i.e., odd *Z* and odd *N*). With these mainly experimental details in hand, a discussion of various major nuclear models follows. These discussions attempt, in their own way, to categorize and explain the mass of experimental data.

1.2 General Nuclear Properties

1.2.1 Properties of Stable Nuclei

The discovery of the neutron allowed each nucleus to be assigned a number, *A*, the mass number, which is the sum of the number of protons (*Z*) and neutrons (*N*) in the particular nucleus. The atomic number of chemistry is identical to the proton number *Z*. The mass number *A* is the integer closest to the ratio between the mass of a nucleus and the fundamental mass unit. This mass unit, the unified atomic mass unit, has the value $1\ \mathrm{u} = 1.660538921(73) \times 10^{-27}\ \mathrm{kg} = 931.494061(21)\ \mathrm{MeV\,c^{-2}}$. It has been picked so that the atomic mass of a $^{12}C_6$ atom is exactly equal to 12 u. The notation here is $^{A}X_{N}$, where X is the chemical symbol for the given element, which fixes the number of electrons and hence the number of protons *Z*. This commonly used notation contains some redundancy because $A = Z + N$ but avoids the need for one to look up the *Z*-value for each chemical element. From this last expression, one can see that there may be several combinations of *Z* and *N* to yield the same *A*. These nuclides are called *isobars*. An example might be the pair $^{196}Pt_{118}$ and $^{196}Au_{117}$. Furthermore, an examination of a table of nuclides shows many examples of nuclei with the same *Z*-value but different *A*- and *N*-values. Such nuclei are said to be *isotopes* of the element. For example, oxygen (O) has three stable isotopes: $^{16}O_8$, $^{17}O_9$, and $^{18}O_{10}$. A group of nuclei that have the same number of neutrons, *N*, but different numbers of protons, *Z* (and, of course, *A*), are called *isotones*. An example might be $^{38}Ar_{20}$, $^{39}K_{20}$, and $^{40}Ca_{20}$. Some elements have but one stable isotope (e.g., $^{9}Be_5$, $^{19}F_{10}$, and $^{197}Au_{118}$), others, two, three, or more. Tin ($Z = 50$) has the most at 10. Finally, the element technetium has no stable isotope at all. A final definition of use for light nuclei is a *mirror pair*, which is a pair of nuclei with *N* and *Z* interchanged. An example of such a pair would be $^{23}Na_{12}$ and $^{23}Mg_{11}$.

The nuclear masses of stable isotopes are determined with a mass spectrometer, and we shall return to this fundamental property when we discuss the nuclear binding energy and the mass defect in Section 1.3. After mass, the next property

of interest is the size of a nucleus. The simplest assumption here is that the mass and charge form a uniform sphere whose size is determined by the radius. While not all nuclei are spherical or of uniform density, the assumption of a uniform mass/charge density and spherical shape is an adequate starting assumption (more complicated charge distributions are discussed in Section 1.4 and beyond). The nuclear radius and, therefore, the nuclear volume or size is usually determined by electron-scattering experiments; the radius is given by the relation

$$R = r_0 A^{1/3} \tag{1.1}$$

which, with $r_0 = 1.25$ fm, gives an adequate fit over the entire range of nuclei near stability. An expression such as Eq. (1.1) implies that nuclei have a density independent of A, that is, they are incompressible. A somewhat better fit to the nuclear sizes can be obtained from the Coulomb energy difference of mirror nuclei, which covers but a fifth of the total range of A. This yields $r_0 = 1.22$ fm. Even if the charge and/or mass distribution is neither spherical nor uniform, one can still define an equivalent radius as a size parameter.

Two important properties of a nuclide are the spin J and the parity π, often expressed jointly as J^π, of its ground state. These are usually listed in a table of isotopes and give important information about the structure of the nuclide of interest. An examination of such a table will show that the ground state and parity of all even–even nuclei is 0^+. The spin and parity assignments of the odd-A and odd–odd nuclei tell a great deal about the nature of the principal parts of their ground-state wave functions. A final property of a given element is the relative abundance of its stable isotopes. These are determined again with a mass spectrograph and listed in various tables of the nuclides.

1.2.2 Properties of Radioactive Nuclei

A nucleus that is unstable, that is, it can decay to a different or *daughter* nucleus, is characterized not only by its mass, size, spin, and parity but also by its lifetime τ and decay mode or modes. (In fact, each level of a nucleus is characterized by its spin, parity, lifetime, and decay modes.) The law of radioactive decay is simply

$$N(t) = N(0)e^{-\lambda t} = N(0)e^{-t/\tau} \tag{1.2}$$

where $N(0)$ is the number of nuclei initially present, λ is the decay constant, and its reciprocal τ is the lifetime. Instead of the lifetime, often the half-life $T_{1/2}$ is used. It is the time in which half of the nuclei decay. By setting $N(T_{1/2}) = N(0)/2$ in Eq. (1.2), one obtains the relation

$$T_{1/2} = \ln(2)\tau = 0.693\tau \tag{1.3}$$

The decay mode of ground states can be α, β, or spontaneous fission. Excited states mostly decay by γ-emission. More exotic decays are observed in unstable nuclei far from stability where nuclei decay takes place by emission of a proton or neutron.

In α-decay, the parent nucleus emits an α-particle (a nucleus of $^4He_2$), leaving the daughter with two fewer neutrons and protons:

$$^A X_N \rightarrow {}^{A-4}Y_{N-2} + {}^4He_2 \tag{1.4}$$

The α-particle has zero spin, but it can carry off angular momentum. In β-decay the weak interaction converts neutrons into protons (β^--decay) or protons into neutrons (β^+-decay). Which of the two

decays takes place depends strongly on the masses of the initial and final nuclei. Because a neutron is heavier than a proton, the free neutron is unstable against β^--decay and has a lifetime of 878.5(10) s. The mass excess in β^--decay is released as kinetic energy of the final particles. In the case of the free neutron, the final particles are a proton, an electron, and an antineutrino, denoted by $\bar{\nu}$. All of these particles have spin 1/2 and can also carry off angular momentum. In the case of β^+-decay, the final particles are a bound proton, an antielectron or positron, and a neutrino. Finally, as an alternative to β^+-decay the initial nucleus can capture an inner electron. In this so-called electron capture decay, only a neutrino, ν, is emitted by the final nucleus. In general, the decays can be written as

$$\beta^-\text{-decay}: \quad {}^{A}\mathrm{X}_N \rightarrow {}^{A}\mathrm{Y}_{N-1} + \mathrm{e}^- + \bar{\nu} \tag{1.5a}$$

$$\beta^+\text{-decay}: \quad {}^{A}\mathrm{X}_N \rightarrow {}^{A}\mathrm{Y}_{N+1} + \mathrm{e}^+ + \nu \tag{1.5b}$$

$$\beta^+\text{-decay (ec)}: {}^{A}\mathrm{X}_N + \mathrm{e}^- \rightarrow {}^{A}\mathrm{Y}_{N-1} + \nu \tag{1.5c}$$

One very rare mode of decay is double β-decay, in which a nucleus is unable to β-decay to a $Z+1$ daughter for energy reasons but can emit two electrons and make a transition to a $Z+2$ daughter. An example is ${}^{82}\mathrm{Se}_{48} \rightarrow {}^{82}\mathrm{Kr}_{46}$ with a half-life of $(1.7 \pm 0.3) \times 10^{20}$ years. Double β-decay is observed under the emission of two neutrinos. Neutrinoless double β-decay is intensively searched for in ^{76}Ge because it is forbidden for massless neutrinos with definite helicities. Enriched Ge is hence used as it allows the use of a large single crystal as source and detector (for a review see [13]).

In spontaneous fission, a very heavy nucleus simply breaks into two heavy pieces. For a given nuclide, the decay mode is not necessarily unique. If more than one mode occurs, then the branching ratio is also a characteristic of the radioactive nucleus in question.

An interesting example of a multimode radioactive nucleus is ${}^{242}\mathrm{Am}_{147}$. Its ground state ($J^\pi = 1^-$, $T_{1/2} = 16.01$ h) can decay either by electron capture (17.3% of the time) to ${}^{242}\mathrm{Pu}_{148}$ or by β^- decay (82.7% of the time) to ${}^{242}\mathrm{Cm}_{146}$. On the other hand, a low-lying excited state at 0.04863 MeV ($J^\pi = 5^-$, $T_{1/2} = 152$ years) can decay either by emitting a γ-ray (99.52% of the time) and going to the ground state or by emitting an α-particle (0.48% of the time) and going to ${}^{238}\mathrm{Np}_{145}$. There is an excited state at 2.3 MeV with a half-life of 14.0 ms that undergoes spontaneous fission [14]. The overall measured half-life of ${}^{242}\mathrm{Am}_{147}$ is then determined by that of the 0.04863 MeV state. Such long-lived excited states are known as *isomeric states*. From this information on branching ratios, one easily finds the several partial decay constants for ${}^{242}\mathrm{Am}_{147}$. For the ground state, $\lambda_{\mathrm{ec}} = 2.080 \times 10^{-6}\ \mathrm{s}^{-1}$ and $\lambda_{\beta-} = 9.944 \times 10^{-6}\ \mathrm{s}^{-1}$, while for the excited state at 0.04863 MeV, $\lambda_\gamma = 1.439 \times 10^{-10}\ \mathrm{s}^{-1}$ and $\lambda_\alpha = 6.639 \times 10^{-13}\ \mathrm{s}^{-1}$ and for the excited state at 2.3 MeV, $\lambda_{\mathrm{SF}} = 49.5\ \mathrm{s}^{-1}$.

1.3 Nuclear Binding Energies and the Semiempirical Mass Formula

1.3.1 Nuclear Binding Energies

One of the more important properties of any compound system, whether molecular,

atomic, or nuclear, is the amount of energy needed to pull it apart, or, alternatively, the energy released in assembling it from it constituent parts. In the case of nuclei, these are protons and neutrons. The binding energy of a nucleus ${}^{A}X_{N}$ can be defined as

$$B(A,Z) = ZM_{\mathrm{H}} + NM_{\mathrm{n}} - M_{\mathrm{X}}(Z,A) \quad (1.6)$$

where M_{H} is the mass of a hydrogen atom, M_{n} the mass of a neutron, and $M_{\mathrm{X}}(Z,A)$ the mass of a neutral atom of isotope A. Because the binding energy of atomic electrons is very much less than nuclear binding energies, they have been neglected in Eq. (1.6). The usual units are atomic mass units, u. Another quantity that contains essentially the same information as the binding energy is the *mass excess* or the *mass defect*, $\Delta = M(A) - A$. (Another useful quantity is the *packing fraction* $P = [M(A) - A]/A = \Delta/A$.) The most interesting experimental quantity $B(A,Z)/A$ is the binding energy per nucleon, which varies from somewhat more than 1 MeV nucleon^{-1} (1.112 MeV nucleon^{-1}) for deuterium (${}^{2}H_{1}$) to a peak near ${}^{56}Fe_{30}$ of 8.790 MeV nucleon^{-1} and then falls slowly until, at ${}^{235}U_{143}$, it is 7.591 MeV nucleon^{-1}. Except for the very light nuclei, this quantity is roughly (within about 10%) 8 MeV nucleon^{-1}. A strongly bound light nucleus is the α-particle, as for ${}^{4}He_{2}$ the binding energy is 7.074 MeV nucleon^{-1}. It is instructive to plot, for a given mass number, the packing fraction as a function of Z. These plots are quite accurately parabolas with the most β-stable nuclide at the bottom. The β^- emitters will occur on one side of the parabola (the left or lower two side) and the β^+ emitters on the other side. For odd-A nuclei, there is but one parabola, the β-unstable nuclei proceeding down each side of the parabola until the bottom or most stable nucleus is reached. For the even-A nuclei, there are two parabolas, with the odd–odd one lying above the even–even parabola. The fact that the odd–odd parabola is above the even–even one indicates that a pairing force exists that tends to increase the binding energy of the even–even nuclei. See Figure 1.1 for the $A = 100$ mass chain. Other indications of the importance of this pairing force are the before-mentioned 0^+ ground states of all even–even nuclei and the fact that only four stable odd–odd nuclei exist: ${}^{2}H_{1}$, ${}^{6}Li_{3}$, ${}^{10}B_{5}$, and ${}^{14}N_{7}$. For even-A nuclei, the β-unstable nuclei zig-zag between the odd–odd parabola and the even–even parabola until arriving at the most β-stable nuclide, usually an even–even one. If the masses for each A are assembled into a three-dimensional plot (with N running along one long axis, Z along a perpendicular axis, and $M(A,Z)$ mutually perpendicular to these two), one finds a ''landscape'' with a deep valley running from one end to the other. This valley is known as the *valley of stability*.

The immediate consequence of the behavior of $B(A,Z)/A$ is that a very large amount of energy per nucleon is to be gained from combining two neutrons and two protons to form a helium nucleus. This process is called *fusion*. The release of energy in the fission process follows from the fact that $B(A,Z)/A$ for uranium is less than for nuclei with more or less half the number of protons. Finally, the fact that the binding energy per nucleon peaks near iron is important to the understanding of those stellar explosions known as *supernovae*. In Figure 1.2, the packing fraction, $P = \Delta/A$, is plotted against A for the most stable nuclei for a given mass number. Note that P has a broad minimum near iron ($A = 56$) and rises slowly until lawrencium ($A = 260$). This shows most clearly the

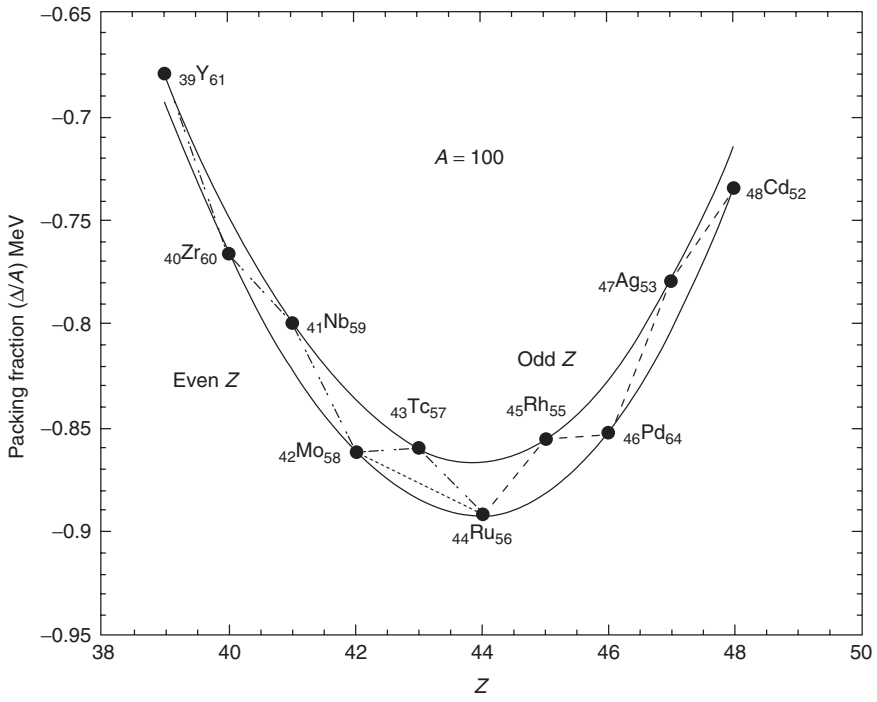


Figure 1.1 The packing fraction Δ/A plotted against the nuclear charge Z for nuclei with mass number $A = 100$. Note that the odd–odd nuclei lie above the even–even ones. The β^- transitions are indicated by - • -, the β^+ transitions by ---, while the double β-decay $^{100}_{42}Mo_{58} \rightarrow ^{100}_{44}Ru_{56}$ is denoted by• • •. Data from [14]. The double β-decay from [15].

energy gain from the fission of very heavy elements.

1.3.2 The Semiempirical Mass Formula

The semiempirical mass formula may be looked upon as simply the expansion of $B(A,Z)$ in terms of the mass number. Because $B(A,Z)/A$ is nearly constant, the most important term in this expansion must be the term in A. From Eq. (1.1) relating the nuclear radius to $A^{1/3}$, we see that a term proportional to A is a volume term. However, this term overbinds the system because it assumes that each nucleon is surrounded by the same number of neighbors. Clearly, this is not true for surface nucleons, and so a surface term proportional to $A^{2/3}$ must be subtracted from the volume term. (One might identify this with the surface tension found in a liquid drop.) Next, the repulsive Coulomb forces between protons must be included. As this force is between pairs of protons, this term will be of the form $Z\{Z-1\}/2$, the number of pairs of Z protons, divided by a characteristic nuclear length or $A^{1/3}$. Two other terms are necessary in this simple model. One term takes into account that, in general, $Z \sim A/2$, clearly true for stable light nuclei, and less so for heavier stable

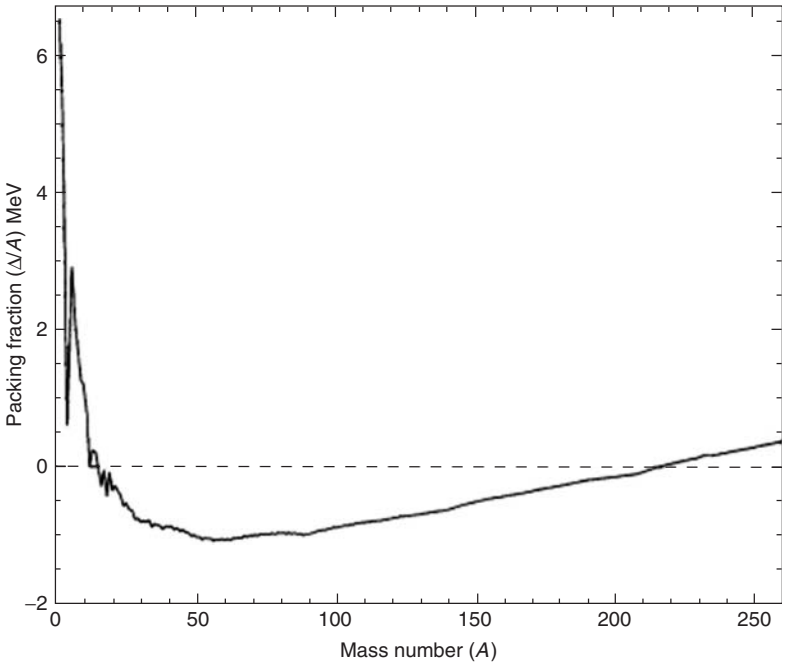


Figure 1.2 The packing fraction Δ/A plotted against the mass number A for all nuclei from ${}^{2}_{1}D_{1}$ to ${}^{260}_{103}Lr_{157}$. Data from [14].

nuclei where more neutrons are needed to overcome the mutual repulsion of the protons. This term is generally taken to be of the form $a_{sym}\ (N-Z)^2/A$. The other term takes into account the fact, noted in Section 1.3.1, that even–even nuclei are more tightly bound than odd–odd nuclei because all of the nucleons of the former are paired off. This is done by adding a term $\delta/2$ that is positive for even–even systems and negative for odd–odd systems and zero for odd-A nuclei. Thus, the two parabolas for even A are separated by δ. From Eq. (1.6), the semiempirical Bethe–Weisäcker mass formula then becomes $M(A,Z) = ZM_H + NM_n - B(A,Z)$ with

$$B(A,Z) = a_v A - a_s A^{2/3} - a_c Z(Z-1)A^{-1/3} - a_{sym}(N-Z)^2 A^{-1} + \frac{\delta}{2} \qquad (1.7)$$

Originally, the constants were fixed by the measured binding energies and adjusted to give appropriate behavior with the mass number [16]. Myers and Swiatecki [17] (see also [18]) have included other terms to account for regions of nuclear deformation, as well as an exponential term of the form $-a_a A\exp(-\gamma A^{1/3})$, for which they provide no physical explanation beyond the fact that it reduces the deviation from experiment. Their model evolved into the macroscopic-microscopic global mass formula, called the *finite-range droplet model* (see [19]) and the *DZ-model* proposed by Duflo and Zuker [20], and more microscopic models, called *HFB* [21]. The many adjustable parameters of the available mass formulas are then fitted to masses of 1760 atomic nuclei [22]. The formulas fit binding energies quite well with errors below 1%, but still

have problems to predict masses far from stability. As those are important for nuclear astrophysics, the measurement of masses of exotic nuclei is an important field today.

A number of consequences flow from even a superficial examination of Eq. (1.7). The fact that the binding energy per nucleon, B/A, is essentially constant with A implies that the nuclear density is constant and, thus, the nuclear force saturates. That is, nucleons interact only with a small number of their neighbors. This is a consequence of the very short range of the strong force. If this were not so, then each nucleon would interact with all others in the given nucleus (just as the protons interact with all other protons), and the leading term in $B(A,Z)$ would be proportional to the number of pairs of nucleons, which is $A(A-1)/2$ or roughly A^2. This would imply that B/A would go as A. Thus, not only does the nuclear force saturate (the Coulomb force does not) but it is also of very short range (that of the Coulomb force is infinite) as the sizes of nuclei are of the order of 3.0 fm (recall Eq. (1.1)).

1.4 Nuclear Charge and Mass Distributions

1.4.1 General Comments

In his 1911 paper, Rutherford was able to conclude that the positive charge of the atom was concentrated within a sphere of radius $<10^{-14}$ m (10 fm). This result came from α-particle scattering. However, for energetic enough α-particles, the scattering result will contain a component due to nuclear interactions of the α-particle, as well as the Coulomb interaction. For probing the structure of nuclei, electrons have the advantage that their scattering is purely Coulombic; however, to determine details of the internal nuclear structure, electron energies must be well over 100 MeV for their de Broglie wavelengths to be less than nuclear dimensions. Well before the existence of such high-energy electron beams, nuclear structure effects were extracted from information provided by optical hyperfine spectra. In particular, nuclear charge distributions (electric quadrupole moments) and current distributions (magnetic dipole moments) were deduced from very accurate optical measurements (see the following section). A result involving the innermost electrons of heavy atoms is the *isotope shift*, which can be observed in atomic X-rays. This arises because the nuclear radii for two different isotopes of the same atom will produce slightly different binding energies of their K-shell electrons. Thus, the K X-rays of these isotopes will be very slightly different in energy. As an example, the isotopic pair $^{203}\mathrm{T1}_{122}$ and $^{205}\mathrm{T1}_{124}$ have an isotope shift of about 0.05 eV. Another early method to determine the charge radius is to take the difference between the binding energies of two mirror nuclei (cf. Section 1.2.1). This leads to an expression that only involves a_c and, thus, the nuclear radius. This is useful for light nuclei for which mirror pairs occur.

With the advent of copious beams of negative muons, much more accurate optical-type hyperfine spectrum studies could be made. The process is quite simple, and the advantages obvious. By stopping negative muons in a target, an exotic atom is formed in which the muon replaces an orbital electron and transitions to the muonic K-shell follow. These transitions of the muon to the $\mathrm{ls}_{1/2}$ state emit photons of the appropriate (but high) energies. (As the muon is more than 200 times as massive as

the electron, the radii of the muon orbits are reduced by that amount, so that electron-shielding problems are much reduced.) The energies of the photons are such that the $2p_{1/2}-2_{3/2}$ splitting is easily measured (in $^{116}Sn_{66}$, it is 45.666 keV). Thus, both nuclear charge radii and isotope shifts are quite accurately determined. In some recent experiments, root mean square (RMS) charge radii have been measured with a precision of 2×10^{-18} m. As electron scattering and muonic atoms are the two methods of measuring characteristics of the nuclear charge radius most susceptible of the greatest accuracy, they will be discussed in turn.

1.4.2 Nuclear Charge Distributions from Electron Scattering

In any scattering experiment, what is measured is the differential cross section $(d\sigma/d\Omega)$. Rutherford developed an expression for α-particle scattering that can be used for low-energy, spinless particles incident on a spinless target. Both incident and target particles are assumed to be point particles. The differential cross section for scattering relativistic electrons off point-charged particles leads to the expression for Mott scattering, while, if the target particle has nonzero spin (there is then a magnetic contribution), one obtains the Dirac scattering formula for $(d\sigma/d\Omega)$. However, real nuclei are not point particles, so one needs to make use of the charge form factor $F(\vec{q})$ with $\vec{q}$ the transferred momentum. The form factor $F(\vec{q})$ is then the Fourier transform of the charge density $\rho_{ch}(\vec{r})$:

$$F(\vec{q}) = \int \rho_{ch}(\vec{r}) e^{i\vec{q}\cdot\vec{r}} d\tau \qquad (1.8)$$

If one restricts the problem to spherically symmetric distributions, the angular integration of the Fourier integral follows at once, so that

$$F(q) = \frac{4\pi}{q} \int \rho_{ch}(r) \sin(qr) r dr \qquad (1.9)$$

If the target nucleus has zero spin (applicable to all even–even nuclei), then the differential cross sections for a point target and a finite-sized target are related by

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_{Ruth.} F(q)^2 \qquad (1.10)$$

With the charge form factor determined experimentally, the inverse transform yields the radial charge density

$$\rho_{ch}(r) = \frac{1}{2\pi^2 r} \int F(q) \sin(qr) q dq \qquad (1.11)$$

If the target nucleus is not of spin zero, then an additional term containing the so-called transverse form factor, $F_T(q)$, is needed. (The form factor defined in Eq. (1.9) is sometimes called the *longitudinal form factor*.) In any event, the charge distribution must be normalized to the number of protons (Z) in the target nucleus.

At this point, there are two ways to proceed. The first is a model-independent analysis of the form factor, or, second, one can assume a model with several parameters and fit these to the data. Limiting oneself to small momentum transfers, one can obtain the form factor as a power series in q^2 by expanding $\sin(qr)$ in Eq. (1.9) in a power series of its argument. Keeping only the lowest term of order q^2, one obtains

$$F(q) = Z\left(1 - \frac{1}{6} q^2 \langle r^2 \rangle\right) \qquad (1.12)$$

with

$$\langle r^2 \rangle = \int r^2 \rho_{\mathrm{ch}}(r) 4\pi r^2 \mathrm{d}r \tag{1.13}$$

the RMS radius of the charge distribution. It should be noted that this is not the nuclear radius R, which is usually taken as the radius of the constant-density sphere. This yields

$$\langle r^2 \rangle = \frac{3}{5} R^2 \tag{1.14}$$

Data compilations [23] show that for most stable nuclei,

$$R \approx 1.25 A^{1/3} \tag{1.15}$$

As was stated in Section 1.2.1, a better way to describe the charge distribution is to use a Fermi distribution which takes account of the constant charge density, ρ_0, at the center of the nucleus and the gradual decrease near the surface. This is achieved by

$$\rho_{\mathrm{ch}}(r) = \rho_0 \frac{1}{1 + \mathrm{e}^{\frac{r - R_{1/2}}{a}}} \tag{1.16}$$

with $R_{1/2}$ the radius at half density and a the diffuseness parameter indicating the distance at which the density falls from 90 to 10% of the constant density ρ_0. For heavy nuclei, the following parameterization holds:

$$\begin{aligned} \rho_0 &= 0.17 \frac{Ze}{A}\,\mathrm{fm}^{-3} \\ a &= 2.4\,\mathrm{fm} \\ R_{1/2} &= 1.128 A^{1/3} - 0.89 A^{-1/3}\,\mathrm{fm} \end{aligned} \tag{1.17}$$

There are enough experimental electron-scattering data available throughout all regions of the stable nuclei that quite accurate charge parameters exist for almost all of the systems. The compendium by de Vries *et al*. [23] lists these parameters fitted to the data for several distribution functions in addition to the two-parameter Fermi functions.

1.4.3 Nuclear Charge Distributions from Atomic Transitions

During the last decades, tremendous progress was obtained in the study of atomic transitions using high-precision laser spectroscopy. This allows the measurement of nuclear charge radii and also of nuclear moments for stable and even unstable isotopes. This is because the difference between a point nucleus and a finite-size nucleus causes a very small change in the Coulomb potential the atomic electrons feel. A small energy difference on the atomic levels results when we assume that the nucleus is a sphere with constant charge density. For 1s electrons one obtains

$$\Delta E_{1\mathrm{s}} = \frac{2}{5} \frac{Z^4 e^2}{4\pi \varepsilon_0} \frac{R^2}{a_0^3} \tag{1.18}$$

with a_0 the Bohr radius. Because no point nucleus exists and the theoretical calculations are not accurate enough to calculate the small shift exactly, one generally measures isotope shifts as the frequency difference of atomic transitions measured in two isotopes of a given element. This then yield the differences in the nuclear radius. Starting from known radii of stable isotopes, it is then possible to determine the radii of unstable nuclei on which one cannot perform electron scattering. It is also possible to measure isotope shifts using optical transitions. Because these are caused by the outermost electrons, the shifts are very small in the order of parts per million. As indicated above, they are still within reach of modern laser techniques.

The small shift for 1s binding energies is related to the large difference between

the Bohr radius and the nuclear radius. On replacing one electron by a muon, the muonic orbits shrink by a factor of 207, the mass difference between the heavy muon and the light electron. At the same time, the muon binding energy is increased by a similar factor, making the transition energies in the mega electron volt region. The energies are so high (in $^{238}U_{146}$, the measured 2p–1s transition is about 6.1 MeV) that one must generate Dirac solutions for the muon moving in a Coulomb potential generated by a non-point-charge distribution. To these initial Dirac solutions, one must add corrections, which, in order of size, are vacuum polarization, nuclear polarization, the Lamb shift, and relativistic recoil. Electronscreening corrections are often included, but they are very tiny (for the ls muonic state in $^{238}U_{146}$, this correction has the value of 11 eV).

As many nuclei are not spherical, several studies have used as the appropriate charge distribution a slightly modified form of Eq. (1.16), which includes the deformations

$$\rho_{\mathrm{ch}}(r) = \rho_0 \left\{ 1 + \exp\left(\frac{r - R_{1/2}\left(1 + \sum_{n=1} \beta_n Y_{n0}(\theta, \phi)\right)}{a} \right) \right\}^{-1} \tag{1.19}$$

Here the β_n are deformation parameters that determine the nuclear shape. As an example, the nuclear mean square radius may be expressed as

$$\langle r^2 \rangle_{\mathrm{deformed}} \approx \langle r^2 \rangle_{\mathrm{sph}} \left(1 + \frac{5}{4\pi}\left(\beta_2^2 + \beta_4^2 + \cdots\right)\right) \tag{1.20}$$

Experiments to fit a, c, and, in deformed regions, β_n have been made throughout the periodic table with results consistent with the electron data. However, to combine the results of electron-scattering experiments with those from muonic atoms, it is necessary to use the so-called Barrett moment

$$\left\langle r^k \mathrm{e}^{-\alpha r} \right\rangle = \frac{4\pi}{Z} \int_0^\infty \rho(r) \mathrm{e}^{-\alpha r} r^{k+2} \mathrm{d}r \tag{1.21}$$

where k and α are fitted to the experimental data. The muonic data are equivalent to data from electron-scattering experiments at low momentum transfer. The inclusion of the muonic Barrett moment improves the overall fit by reducing normalization errors. This then reduces the uncertainties over what would be obtained by fitting either the electron-scattering or the muonic atom data alone. Extensive tables of data fitted by various charge distribution models as well as model independent analyzes can be found in de Vries *et al.* [23].

1.4.4 Nuclear Mass Distributions

While the measurement of the charge distribution can be made using electromagnetic probes, this is not possible for the mass distribution because of the uncharged neutron. Instead, the nuclear strong force has to be used. This is more

complicated as mostly both Coulomb force and strong force are present. Nevertheless, from α-scattering experiments, information of the mass distribution is obtained. There are also indirect ways in which one can get information on nuclear mass radii. One example is the dependence of α-decay rate on the nuclear radius that defines the Coulomb barrier. In deformed nuclei, this causes an anisotropy because the Coulomb barrier is lower in the direction of the longest axis, making the tunnel probability enhanced. A second way is to use pions instead of muons. These interact with the nucleus through both the Coulomb force and strong force, which, in comparison to muonic atoms, causes an extra shift that allows the determination of the mass radius. The result of these experiments on stable nuclei finds that the charge and mass radii are equal to within about 0.1 fm. This somewhat surprising result can be understood as a balance between the proton Coulomb repulsion that tends to push the protons to the outside and a strongly attractive neutron–proton strong force that tends to pull the extra neutrons to the inward.

Recently, the common opinion that the radii scale with $A^{1/3}$ was found to be heavily violated in more exotic nuclei.

Especially in light nuclei with a large neutron number, so-called halo-nuclei, strong deviations were observed (an early review is given in [24]). Using the radioactive beam techniques, very neutron-rich He, Li, and Be isotopes can be created and studied in the laboratory. It turned out that these loosely bound nuclei show very extended neutron radii whereby two neutrons are moving at radii similar to the radii of Pb isotopes. Moreover, as is the case of $^{11}\mathrm{Li}_8$, the bound system consists of three entities: two neutrons and a $^{9}\mathrm{Li}_6$ that cannot exist two by two, as the dineutron and $^{10}\mathrm{Li}_7$ are unbound. The research on exotic nuclei is still in its infancy and more exotic features such as proton halos or neutron skins are expected. They are of importance as they may influence the creation of the elements under astrophysical conditions.

1.5 Electromagnetic Transitions and Static Moments

1.5.1 General Comments

Static electromagnetic nuclear moments played an important role in the unscrambling of the detailed measurements of atomic optical hyperfine structure well before the gross components of atomic nuclei were in hand. Almost a decade before the discovery of the neutron, Pauli [7] suggested that the optical hyperfine splitting might be due, in part, to the interaction with a nuclear magnetic moment (μ). This suggestion lay fallow until 1930, when Goudsmit and Young, using the spectroscopic data of Schiller and of Granath, deduced the nuclear magnetic moment of ^{7}Li to be $\mu = 3.29\mu_{\mathrm{N}}$, where the nuclear magneton equals

$$\mu_{\mathrm{N}} = \frac{e\hbar}{2M_{\mathrm{p}}} = 5.050789 \times 10^{-27}\,\frac{\mathrm{J}}{\mathrm{T}} \tag{1.22}$$

This value is quite close to the currently accepted value ($\mu = 2.327\mu_{\mathrm{N}}$). Because of the existence, by then, of extensive hyperfine optical spectroscopic data, Goudsmit, in 1933, was able to publish a table of some 20 nuclear magnetic moments ranging from ^{7}Li to ^{209}Bi. In 1937, Schmidt published a simple, single-particle model of nuclear magnetic moments and supported it with the experimental moments of 32 odd-proton nuclei and 15 odd-neutron

nuclei. This simple model yields what is now known as the *Schmidt limits*, within which almost all nuclear magnetic moments lie (see the following).

The suggestion that the nuclear electric quadrupole moments (Q) might also play an important role in optical hyperfine structure was again made before the discovery of the neutron. Racah [8] was the first to work out the theory associated with ''nuclear charge asymmetry'' and the interaction with the atomic electrons. Casimir [25], sometime later, developed the theory of nuclear electric quadrupole hyperfine interaction and applied it to ^{151}Eu and ^{153}Eu. In this paper, Casimir mentions work by Schiller and Schmidt, who determined Q for ^{175}Lu. A short time later, Gollnow [26] obtained $Q = 5.9$ b for this nucleus, quite close to the currently accepted value of 5.68 b. This very large quadrupole moment (very much larger than can be accounted for by the single-particle shell model) was to provide, 20 years later, strong impetus for the development of the collective model of the nucleus. In 1954, Schwartz [27] extended the theory of nuclear hyperfine structure to examine the magnetic octupole hyperfine interaction and calculated the first four nuclear magnetic octupole moments (O) from data of the hyperfine structure of the nuclear ground states. The next nuclear moment is the hexadecapole (H); however, no direct measurements of such static moments exist. What is known about these moments comes mainly from electromagnetic transitions of electrons and negative muons. For an in-depth theoretical study of all of these moments and how they can be used to test various nuclear models see, in particular, the text by Castel and Towner [28].

Nowadays, the measurement of moments is still very important to assess the single-particle structure of exotic nuclei and several powerful techniques have been developed in this domain [29]. Most information is, however, gathered via the determination of electromagnetic transitions by γ-ray spectroscopy. This is to a large extent due to the availability of large-volume semiconductor detectors for γ-ray detection and the high computing power that allows one to analyze more and more complex measured spectra using coincidence conditions. The recent development of γ-ray tracking detectors out of segmented Ge-detectors offers very high perspectives in the field of exotic nuclei [30].

1.5.2 Electromagnetic Transitions and Selection Rules

Without going into a detailed discussion on how matrix elements are calculated, we review here the calculation of electromagnetic transitions. The interested reader can find more details in Heyde [31]. The calculation of transitions and also moments involves the wavefunctions of nuclear states and forms a very sensitive probe for nuclear structure research. On the other hand these transitions and moments are electromagnetic in nature making the interaction very well understood. Using the long-wavelength approximation $\lambda \gg R$ and a multipole expansion of the electromagnetic operators, the transition rates per unit of time can be expressed as

$$T(L) = \frac{2}{\varepsilon_0 \hbar} \frac{L+1}{(L[(2L+1)!!]^2)} \left(\frac{\omega}{c}\right)^{2L+1} B(L) \tag{1.23}$$

with L the multipolarity, ω the angular frequency of the radiation such that $\hbar\omega \cong E_i - E_f$ up to a small nuclear recoil correction, and $B(L)$ the reduced transition probability

$$B(J_i \to J_f; L) = \sum_{M_i, M_f} \left| \left\langle \alpha_f; J_f M_f \right| O(LM) \left| \alpha_i; J_i M_i \right\rangle \right|^2 \tag{1.24}$$

in units of $e^2 b^L$ and $\mu_N b^{L-2}$ for electric and magnetic $B(LM)$ values. The labels α identify the initial and final states and $O(LM)$ is the electric or magnetic multipole operator of rank LM. As all states have good angular momentum, one can now use for the transition rates the Wigner–Eckart theorem to remove all reference to the M projections. This yields

$$B(J_i \to J_f; L) = \frac{1}{2J_i + 1} \left| \left\langle \alpha_f; J_f \| O(L) \| \alpha_i; J_i \right\rangle \right|^2 \tag{1.25}$$

The electric multipole operator for a number of point charges becomes

$$O(E; LM) = \sum_i e_{\text{eff}}(i) r_i{}^L Y_{L,M}(\theta_i, \phi_i) \tag{1.26}$$

with $e_{\text{eff}}(i)$ the effective charge of the ith nucleon. Here it is anticipated that owing to core polarization effects and truncations of the model space, other values than the free charges $+e(0)$ for proton (neutron) need to be used. Instead of the operator (Eq. (1.25)), one can also use a similar operator but using the nuclear electric charge density ρ_{ch} and an integration over the nuclear volume. These are used as several models describe the nucleus as a droplet (Section 1.7). For the magnetic multipole operator, we have

$$O(M; LM) = \sqrt{L(2L+1)} \mu_N \sum_i r_i{}^{L-1} \times \left\{ \frac{e_{\text{eff}}(i)}{e} \frac{2}{L+1} [Y_{L-1} \otimes j_i]^{(L)} + \left(g_s(i) - \frac{e_{\text{eff}}(i)}{e} \frac{1}{L+1} \right) \times \left[Y_{L-1} \otimes s_i \right]^{(L)} \right\} \tag{1.27}$$

with the effective gyromagnetic ratio g_s which also may differ from the free ones.

The multipole expansion and the fact that states in atomic nuclei have good angular momentum and parity leads to several selection rules. The first one is related to the vector coupling of the angular momentum and states $|J_i - J_f| \le L \le J_i + J_f$. The second is due to the parity of the operators, which clearly is $(-1)^L$ for the electric and $(-1)^{L-1}$ for the magnetic operator. Owing to this, electric and magnetic transitions of order L cannot take place at the same time between states, and moments such as the electric dipole moment are forbidden. While the selection rules allow the determination of the spin and parities of nuclear excited states, they are also (at a higher level) invaluable to test nuclear models (Section 1.7).

Weisskopf has estimated the so-called single-particle values or Weisskopf units (W.u.) by assuming that a single-particle makes a transition with multipole L from a state with spin $L + 1/2$ toward a state with spin $1/2$, that the radial part of the wavefunction can be approximated by a constant value up to the radius R, and that certain values for the effective charges hold. This leads to the following estimates for the half-lives corresponding to the single-particle values:

$$T_{1/2}(E1) = \frac{6.764 \times 10^{-6}}{E_\gamma^3 A^{2/3}} \text{ (s)}$$

$$T_{1/2}(M1) = \frac{2.202 \times 10^{-5}}{E_\gamma^3} \text{ (s)}$$

$$T_{1/2}(E2) = \frac{9.527 \times 10^{6}}{E_\gamma^5 A^{4/3}} \text{ (s)}$$

$$T_{1/2}(M2) = \frac{3.102 \times 10^{7}}{E_\gamma^5 A^{2/3}} \text{ (s)}$$

$$T_{1/2}(E3) = \frac{2.045 \times 10^{19}}{E_\gamma^7 A^{2}} \text{ (s)}$$

$$T_{1/2}(M3) = \frac{6.659 \times 10^{19}}{E_\gamma^7 A^{4/3}} \text{ (s)} \quad (1.28)$$

One notices that transition rates of the lower multipolarities are faster than the higher by orders of magnitude and that for a given multipolarity, the electric ones are about 100–1000 faster as the magnetic ones. Owing to the selection rules and enhanced quadrupole collectivity, only $E2$ and $M1$ transitions happen on similar timescales. In this case, one has a transition of mixed multipolarity.

The Weisskopf estimates are very crucial to determine whether a transition is caused by a single nucleon changing orbits or by several nucleons acting in a collective way. The measurement of transition rates of excited states delivers very important information on nuclear structure, but is also quite involved. One needs to measure the lifetime of a state, the (mixed) multipolarity, and the energy and intensities of the transitions deexiting a given state. To this end, γ-ray arrays consisting of several Ge-detectors are appropriate. Besides this, electromagnetic decay can also take place with the emission of conversion electrons. These electrons allow one to determine the multipolarity as well as to observe the by γ-emission forbidden $E0$ transitions (due to the fact that the photon has spin 1 with projection $+1$ and -1).

1.5.3 Static Moments

In contrast to the transition rates, the *multipole moments* are generally defined as the matrix element of the $M = 0$ component of the moment operator for a single state with magnetic projection $M = +J$. Of the moments, the two lowest are the most important.

1.5.3.1 Magnetic Dipole Moments

If one assumes that the magnetic properties are associated with the individual nucleons, then the magnetic moment is defined as

$$\mu = \langle \alpha, JJ | \sum_i g_l(i) l_{z,i} + g_s(i) s_{z,i} | \alpha; JJ \rangle \quad (1.29)$$

where the sum extends over all of the A nucleons. Generally, this will not be needed; for instance, the magnetic moment of an odd-A nucleus will be generated by the last neutron and proton as the adjacent even–even ground state has no magnetic moment. In Eq. (1.29), g_l and g_s are the orbital and spin gyromagnetic ratios. They are often chosen as $0.7 g_{\text{free}}$ where the free-particle values are $g_l = 1$, $g_s = 5.587$ for protons and $g_l = 0$, $g_s = -3.826$ for neutrons (all in μ_N). It is instructive to calculate the magnetic moment for a single nucleon (which, as explained earlier, is a good approximation for the ground state of odd-A nuclei). Using the total angular momentum and the Wigner–Eckart theorem, one deduces the single-particle magnetic moments for aligned and antialigned orbital momentum and spin:

$$\mu(j = l + 1/2) = j \left(g_l + \frac{(g_s - g_l)}{2} \right) \quad (1.30a)$$

$$\mu(j = l - 1/2) = \left(j g_l - \frac{j}{j+1} \frac{(g_s - g_l)}{2} \right) \quad (1.30b)$$

The magnetic moments plotted as a function of spin form the so-called Schmidt lines. It is interesting that, when plotted, almost all of these moments lie between the Schmidt lines. The moments that

lie outside these limits occur mainly for some very light nuclei. One may conclude that the single-particle model does possess some validity. Another set of limits, the Margenau–Wigner (M-W) limits, is obtained by replacing the free-particle values for the orbital gyromagnetic ratios, g_l, by the uniform value Z/A. The justification for this is that one is in effect averaging over all states that lead to the correct nuclear spin. This calculation represents an early attempt to account for core contributions to the dipole-moment operator. Figure 1.3 and Figure 1.4 show plots of a number of the ground-state magnetic moments for odd-Z (Figure 1.3) and odd-N (Figure 1.4) nuclei.

Besides the ground state, excited states can have magnetic moment and their measurement is often used to extract information on the underlying single-particle structure. Common in nuclear structure physics is the use of g-factors that are analogous to the single-particle gyromagnetic ratios, except that they are dimensionless. They are defined as

$$\mu = gI\mu_N \tag{1.31}$$

One way to determine the g-factor is to measure the Lamor frequency when excited nuclei are placed in an external magnetic field B. Then,

$$\omega_L = \frac{gB\mu_N}{\hbar} \tag{1.32}$$

The main problem hereby is to align the spins of an ensemble of atomic nuclei.

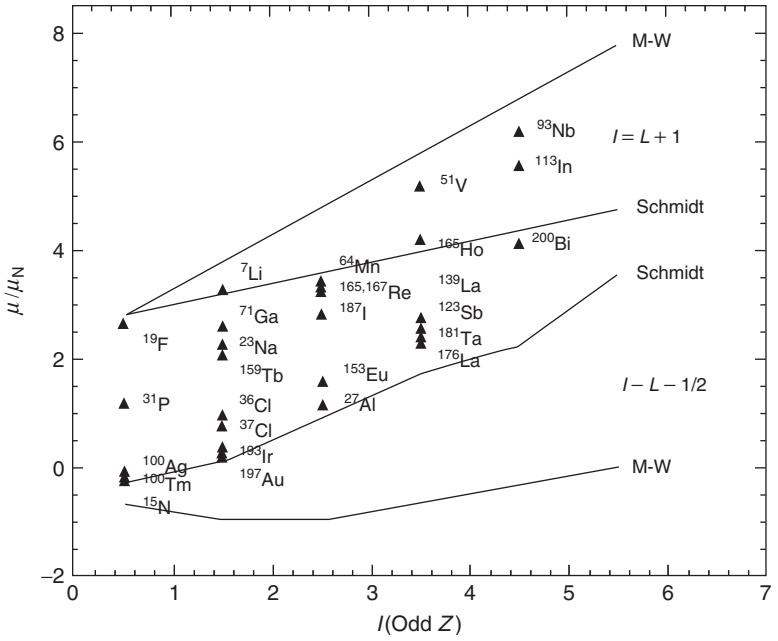


Figure 1.3 Nuclear magnetic moments in units of the nuclear magneton (μ/μ_N) plotted against the nuclear spin (I) for a number of odd-Z nuclei. The Schmidt limits, as well as the Margenau–Wigner (M-W) limits, are shown as solid lines. Data from [14].

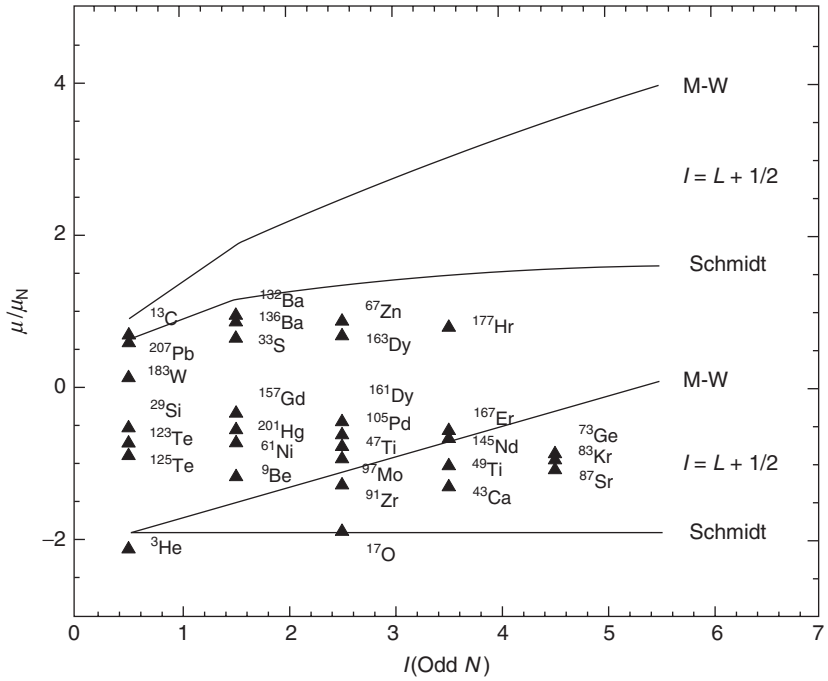


Figure 1.4 Nuclear magnetic moments in units of the nuclear magneton (μ/μ_N) plotted against the nuclear spin (I) for a number of odd-N nuclei. The Schmidt limits, as well as the Margenau–Wigner (M-W) limits, are shown as solid lines. Data from [14].

This has to be done by the nuclear reaction used, cooling in an external magnetic field, or via the observation of changes in angular correlations.

1.5.3.2 Electric Quadrupole Moments

As a general definition of the quadrupole moment, we have the expectation value of $(3z^2 - r^2)$. Using the proportionality of the quantity with $Y_{20}(\theta,\phi)$ one gets

$$Q(J) = \sqrt{\frac{16\pi}{5}} \times \left\langle \alpha; JJ \left| \sum_i \frac{e_{\text{eff}}(i)}{e} r_i^2 Y_{2,0}(\theta_i, \phi_i) \right| \alpha; JJ \right\rangle \tag{1.33}$$

or, using the Wigner–Eckart theorem,

$$Q(J) = \sqrt{\frac{16\pi}{5}} \sqrt{\frac{J(2J-1)}{(2J+1)(2J+3)(J+1)}} \times \left\langle \alpha; J \left\| \sum_i \frac{e_{\text{eff}}(i)}{e} r_i^2 Y_2(\theta_i, \phi_i) \right\| \alpha; J \right\rangle \tag{1.34}$$

The quadrupole moment can also be calculated for a single nucleon in an orbit $j = J$. This yields

$$Q(j) = -\frac{(2j-1)}{(2j+2)} \frac{e_{\text{eff}}(i)}{e} \langle r^2 \rangle \tag{1.35}$$

One thus obtains a negative quadrupole moment for a single nucleon. If the orbit

is filled up to a single hole, a quadrupole moment as in Eq. (1.35) but with a positive sign is expected. We would like to illustrate the application of Eq. (1.35) with examples near the doubly magic nucleus $^{16}O_8$ (see also Section 1.7). The orbit that the ninth nucleon can occupy has $j = 5/2$ and we can use Eqs. (1.13) and (1.14) to estimate $<r^2>$. This yields for $^{17}F_8$ using the free-proton charge $Q = -5.9\,\mathrm{fm}^2$, which is in excellent agreement with the experimental absolute value given in Firestone *et al.* (1996) of $|Q| = 5.8(4)\,\mathrm{fm}^2$. For the odd neutron nucleus $^{17}O_9$, one finds experimentally $Q = -2.578\,\mathrm{fm}^2$, which shows the need for effective charges and can be reproduced using $0.44e$ as neutron effective charge. Finally, if we place five neutrons in the $j = 5/2$ orbit, we have the $N = 13$ isotones and expect moments of $Q = +2.6\,\mathrm{fm}^2$. Experimentally, one finds $Q = 20.1(3)\,\mathrm{fm}^2$ in $^{25}Mg_{13}$ and $Q = 10.1(2)\,\mathrm{fm}^2$ in $^{23}Ne_{13}$. This observation of much larger quadrupole moments occurs for most atomic nuclei having several nucleons in an orbit or in several orbits, indicating that the model is too simple and that all of the electric charges must be considered. This holds especially if the core is not spherically symmetric and the motion becomes collective.

1.6 Excited States and Level Structures

1.6.1 The First Excited State in Even–Even Nuclei

The most obvious characteristic of the various nuclei is that all of the even–even nuclei have ground-state spins and parities of 0^+. This not only categorizes one large group of nuclei but also indicates that the nuclear force is such that it couples, preferentially, pairs of like nucleons to angular momentum zero. The second observation is that the first excited state in even–even nuclei is almost always a 2^+ excitation. This can be understood by the combination of good total angular momentum and the Pauli principle. If one couples two nucleons in orbit j and implies the antisymmetrization, one obtains

$$\varphi(j,j;JM) = \frac{1}{2}\sum_{mm'}(jmjm'|JM)(\psi_1(jm)\,\psi_2(jm') - \psi_2(jm)\psi_1(jm')) \tag{1.36}$$

with the Clebsch–Gordan coupling coefficient for the angular momentum coupling. This can, however, be rewritten as the m-values are in the summation. Using the symmetry properties, one arrives at

$$\varphi\,(j,j;JM) = \frac{1}{2}\left(1-(-1)^{2j-J}\right) \times \sum_{mm'}(jmjm'|JM)\,\psi_1(jm)\psi_2(jm') \tag{1.37}$$

The phase factor, and the fact that $2j$ is an odd number, imply that states with odd values of J do not exist. One might wonder whether this is an essential property of fermions, but surprisingly it is not. Consider bosons with integer angular momentum ℓ. Then, owing to symmetrization, Eq. (1.36) becomes

$$\varphi(\ell,\ell;JM) = \frac{1}{2}\sum_{mm'}(\ell m\ell m'|JM) \times(\psi_1(\ell m)\psi_2(\ell m') + \psi_2(\ell m)\psi_1(\ell m')) \tag{1.38}$$

which yields after reordering,

$$\varphi(\ell,\ell;JM) = \frac{1}{2}\left(1+(-1)^{2\ell-J}\right)\sum_{mm'} \times(\ell m\ell m'|JM)\psi_1(\ell m)\psi_2(\ell m') \tag{1.39}$$

but now 2ℓ is even and again all states with odd J-values disappear. The fact that both bosons and fermions yield similar results is very important for nuclear models. It means that one can often describe even–even atomic nuclei using either fermionic nucleons or collective bosons such as phonons.

1.6.2 Regions of Different Level Structures

The existence of a low-lying 2^+ excitation can be explained in different ways, that is, short-range interaction, a collective quadrupole vibration, or as a first excited state of a rotational band. Which interpretation is right or better and which mixture of interpretations is right depends a lot on where on the Segre chart the nucleus is located and how the higher excitations behave. This is illustrated in Figure 1.5, which shows the known energies of the first excited states in the even–even Cd and Hf isotopes. One notices huge differences in the absolute excitation energies but also a quite smooth behavior. If one now also considers the next excited state, which is a 4^+ excitation and plots the ratio $R4/2$,

$$R4/2 = \frac{E_{\mathrm{exc}}(4^+)}{E_{\mathrm{exc}}(2^+)} \tag{1.40}$$

(see also Figure 1.5) more information can be obtained. In the Hf isotopes, a clear saturation near $R4/2 = 3.33$ occurs indicating that a rotational band, with its typical $I(I+1)$ energy dependence, is built on the ground state. For Cd, we observe very high 2^+ excitation energies and low $R4/2$ ratios at the beginning and end of the isotope chain. This is a clear evidence that at the extremes, a shell closure occurs, when $N = 50$ and 82. In the middle, one finds, compared to Hf, much higher 2^+ excitation energies and $R4/2$ ratios slightly above 2, indicating an anharmonic quadrupole vibrational nature at mid shell.

While the strong pairing of like nucleons can explain the ground state of even–even nuclei, it does not yield predictions for the ground-state spins and parities of odd-A and odd–odd nuclei. Nevertheless, it simplifies this task a lot, as in most cases one can conclude that the ground-state spin and parity of an odd-A nucleus is equal to the one of the last odd nucleon. This forms an important testing ground for the shell model, which predicts the single-particle energy, spin, and parity of the subsequent orbits.

1.6.3 Shell Structures

Various sets of nuclear data indicate that certain numbers of nucleons, either neutrons or protons, correspond to the filling of angular momentum "shells." These are similar to the atomic shells often denoted as the K-shell, L-shell, and so on. In nuclei, the shell filling is similar but not identical, and the principal shell closings occur at experimentally observed numbers. These are for either *N or Z* equal to 2, 8, 20, 28, 50, 82, and 126. These are the *magic numbers*. They occur where there are drastic changes in neutron cross sections, nucleon separation energies, and so on. An interesting, but small, group of nuclei comprises the doubly magic ones – that is, nuclei with both neutron and proton numbers magic. The five stable ones are $^{4}\mathrm{He}_{2}$, $^{16}\mathrm{O}_{8}$, $^{40}\mathrm{Ca}_{20}$, $^{48}\mathrm{Ca}_{28}$, and $^{208}\mathrm{Pb}_{126}$. Note that no stable doubly magic nuclei with 50 and 82 exist. What is notable about the doubly magic nuclei is that their first excited states are at a very

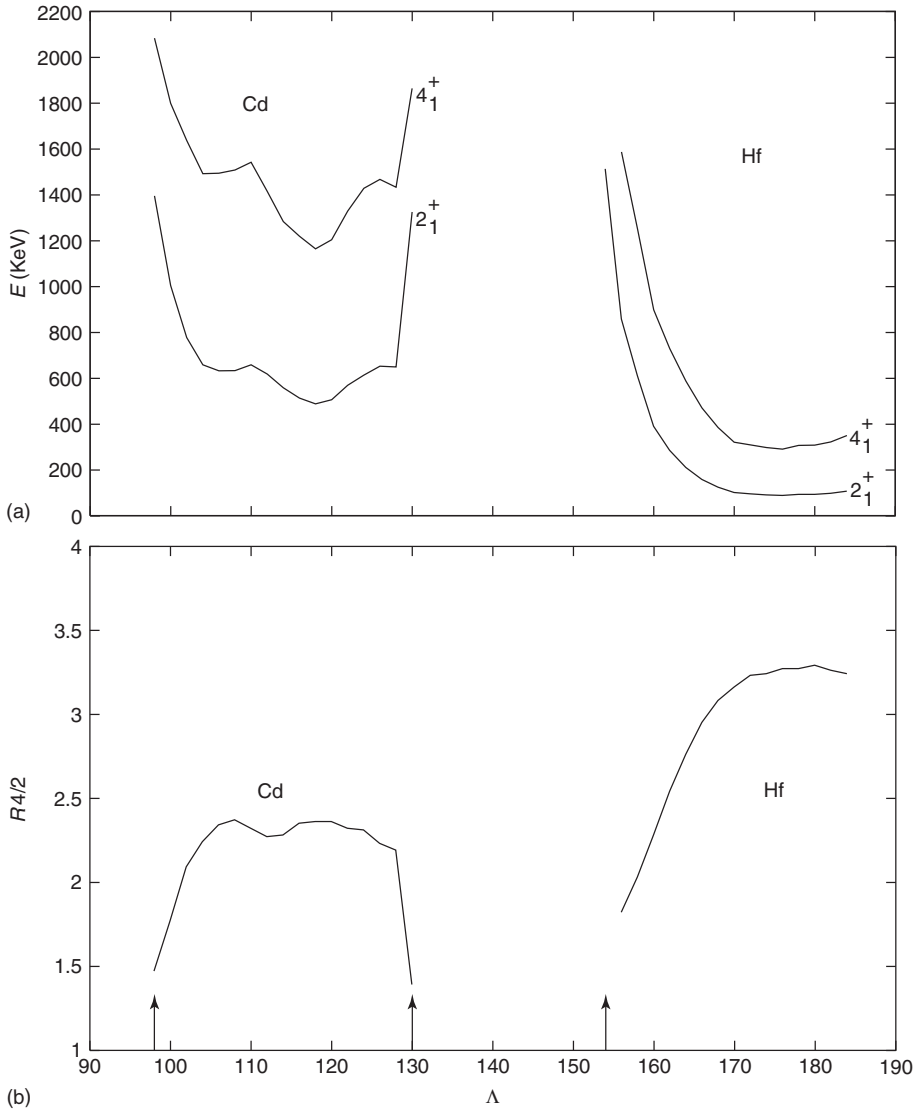


Figure 1.5 The known energies of the first excited states in the even–even Cd and Hf isotopes (a). The $R4/2$ ratios are also given (b). The left side represents the Cd isotopes and the right side the Hf isotopes. Arrows mark the closed shells at $A = 98$, 130, and 154.

high energy compared with their non-doubly magic neighbors. The energy in megaelectronvolts, spin, and parity of these first excited states is 20.1, 0^+; 6.05, 0^+; 3.35, 0^+; 3.83, 2^+; and 2.61, 3^-, respectively. Also the observed spins (and parities) of the first excited states is very different from that of all other even–even nuclei.

Because of these large gaps or changes in nuclear properties, they divide the low-lying excited states of nuclei into roughly three regions. These are, broadly, the nuclei in the neighborhood of the magic numbers that are dominated by shell or single-particle structures and those further away, which show collective behavior. Of particular interest are the ones found between these two, as they may give clues to the onset of collective motion in atomic nuclei and other systems. We mentioned earlier that 126 is a magic number, but strictly speaking, this holds only for neutrons as the unstable element with $Z = 126$ has not yet been observed. Nowadays, there is a strong interest to study whether magic numbers still stay valid in exotic nuclei far from stability and which is the next magic number for superheavy isotopes.

1.6.4 Collective Structures

Within major divisions (even–even, odd-A, and odd–odd), the excited levels are further divided into groups that are single-particle or collective (vibrational or rotational) in nature. However, because of the strong pairing force, even–even nuclei do not show single-particle excited levels but rather one- or two-particle-hole excitations or two-quasiparticle excitations. Here, we introduce the most general features of the two major classes of collective nuclei. We restrict ourselves to even–even nuclei for the sake of simplicity.

1.6.4.1 Vibrational Levels

The intermediate systems between the tightly bound magic nuclei and the deformed nuclei are the so-called vibrational nuclei. Here the number of nucleons outside of a deformable core is small, and the zero-point energy of the lowest oscillations is greater than the energy of deformation, so that the shape of the core is not stabilized. The motion then is of quantized surface oscillations with angular momentum two. One can introduce creation and destruction phonon operators Q^+_{LM} and Q_{LM} which fulfill the boson commutation rules:

$$[Q_{LM}, Q^+_{LM}] = \delta_{LL'}\delta_{MM'} \tag{1.41}$$

and for which all other commutators are zero. The angular momentum of the phonon is L, its magnetic projection M and its parity $\pi = (-1)^L$. The magnetic projection is mostly not of great importance. In this representation, the number operator counting the number of L phonons is given by

$$\widehat{N}_L = \sum_M Q^+_{LM} Q_{LM} \tag{1.42}$$

Using the number operator, the simplest Hamiltonian becomes

$$\widehat{H} = \sum_L \hbar\omega_L \left(\widehat{N}_L + L + \frac{1}{2}\right) \tag{1.43}$$

where the additional factors are chosen such that the Hamiltonian corresponds to the quantized harmonic oscillators for L phonons. Each of the phonons has energy $\hbar\omega_L$. Because most even–even nuclei have as first excited state a 2^+ state, the $L = 2$ quadrupole phonons are dominant. At higher energy, there is evidence for the presence of $L = 3$ octupole phonon making a 3^- state. Using the energy solution of the Hamiltonian equation (Eq. (1.43)),

$$E(N_L, L) = \sum_L \hbar\omega_L \left(N_L + L + \frac{1}{2}\right) \tag{1.44}$$

and the proper angular momentum coupling and symmetrization given in Eq. (1.39), one arrives at a very simple prediction. The one-quadrupole phonon state has $L=2$ and $E_{\rm exc}=\hbar\omega_2$ and the two-phonon states have $L=0$, 2, 4 and $E_{\rm exc}=2\hbar\omega_2$. This exactly yields the $R4/2$ ratio of 2 observed in Figure 1.5.

A nice example of a "good" vibrational even–even nucleus is $^{110}Cd_{62}$. The first excited 2^+ state is at 0.6578 MeV followed by a 0^+, 2^+, 4^+ states at 1.473, 1.476, and 1.542 MeV. The ratio $R4/2=2.35$ is somewhat larger than the one the simple model gives.

1.6.4.2 Rotational Levels

If the number of nucleons outside the deformable core is such that the zero-point oscillations are much less than the energy of deformation, then the system will have a stable but deformed shape and one must quantize a rigid rotator. From classical considerations, we know that, if the system has a permanent, nonspherical shape, there exists a body-fixed system in which the inertial tensor $\mathfrak{I}$ is diagonal and is related to the laboratory-fixed system by an Euler transformation. We denote these two coordinate systems as (1,2,3) and (x,y,z), respectively. The inertial tensor then has components $\mathfrak{I}_1$, $\mathfrak{I}_2$, and $\mathfrak{I}_3$, so that the Hamiltonian is just

$$\widehat{H}=\frac{1}{2}\left(\frac{\widehat{L}_1^2}{\mathfrak{I}_1}+\frac{\widehat{L}_2^2}{\mathfrak{I}_2}+\frac{\widehat{L}_3^2}{\mathfrak{I}_3}\right) \tag{1.45}$$

with $\widehat{L}_{1,2,3}$ the body-fixed angular momentum. The simplest system is the one for which all moments of inertia are equal: $\mathfrak{I}_1=\mathfrak{I}_2=\mathfrak{I}_3=\mathfrak{I}$. Then the energies are directly found:

$$\widehat{H}=\frac{\hbar^2}{2\mathfrak{I}}(L(L+1)) \tag{1.46}$$

The system specified by Eq. (1.45) possesses the symmetry properties belonging to the point group D_2 for which four representations exist. The Hamiltonian operator (Eq. (1.45)) does not mix different representations. The basis functions are those of a symmetric top, $|LMK>$ being diagonal in $\boldsymbol{L}^2$, $\boldsymbol{L}_z$, and $\boldsymbol{L}_3$ with the usual eigenvalues $L(L+1)$, M, and K, respectively. Here K is the projection of the angular momentum on the body-fixed frame. Applying this formula for the ground-state rotational band with $K=0$, we find now $R4/2=3.33$. However, states with different K-values are degenerated. This is not found in real deformed nuclei.

If the inertia tensor is such that $\mathfrak{I}_1\neq\mathfrak{I}_2\neq\mathfrak{I}_3$, an asymmetric top problem results. In this case, the wave functions are to be expanded in terms of the symmetric-top functions:

$$|LM\rangle=\sum_{K=-L}^{+L}A_K\,|LMK\rangle \tag{1.47}$$

The 1, 2, 3 labels of the momental ellipsoid's semiaxes are quite arbitrary and can be chosen in 24 different ways for right- (left-) handed systems. These 24 relabelings can be produced by three different relabeling transformations T_1, T_2, and T_3 [32], which correspond to label interchanges and operate on the $|LM>$ wave functions. They have the properties

$$T_1^2=T_2^4=T_3^3=1 \tag{1.48}$$

From these relations, one can generate Table 1.1, relating the four representations and the values of angular momentum and parity allowed in each. Note in particular that the A and B_2 representations are associated with positive-parity ($\pi=+$) states, while B_1 and B_3 representations are associated with negative parity ($\pi=-$)

Table 1.1 The representation of the point group D_2 to which the states belong, and the allowed values of the angular momentum and parity associated with them.

Representation	K	Parity	Allowed L-values
A	Even	+	$L = 0,2,2,3,4,4,4,5,5$
B_1	Even	−	$L = 1,2,3,3,4,4,5,5,5$
B_2	Odd	+	$L = 1,2,3,3,4,4,5,5,5$
B_3	Odd	−	$L = 1,2,3,3,4,4,5,5,5$

states. From this table, we note that only the rigid rotator systems belonging to the A representation have a zero angular momentum state. They must be associated with the lowest levels of even–even nuclei and form the $K = 0$ ground-state band with $K = 0$ and $L = 0, 2, 4, 6, \ldots$ followed by the γ-vibrational band with $K = 2$ and $L = 2, 3, 4, 5, \ldots$.

Negative-parity states in even–even nuclei could be associated with either the B_1 or B_3 representations. If one argues that, given two (or more) possibilities, the lowest set of states will be the most symmetric, then because the B_1 representation arises from sums over even K quantum numbers, one should associate this representation with the lowest-lying negative-parity states in even–even systems. Furthermore, in the deformed regions, the negative-parity bands in these nuclei not only lie above the ground-state rotational band but they also have as their band head a 1^- level. This is another indication that these negative-parity rotational bands belong to the B_1 representation.

The problem of solving the asymmetric, rigid rotor is far more involved than the symmetric rotor as all of the latter's eigenvalues can be given in closed form. To proceed along this line, one sets $\mathfrak{I}_0 = \mathfrak{I}_1 = \mathfrak{I}_2 \neq \mathfrak{I}_3$. The Hamiltonian of the system then becomes

$$\begin{aligned}\widehat{H}_{\text{sym}} &= \frac{1}{2}\left(\frac{\widehat{L}_1^2 + \widehat{L}_2^2}{\mathfrak{I}_0} + \frac{\widehat{L}_3^2}{\mathfrak{I}_3}\right) \\ &= \frac{1}{2}\left(\frac{\widehat{L}^2}{\mathfrak{I}_0} + \left(\frac{1}{\mathfrak{I}_3} - \frac{1}{\mathfrak{I}_0}\right)\widehat{L}_3^2\right) \quad (1.49)\end{aligned}$$

and both of the operators appearing are diagonal. Thus, we obtain immediately the eigenvalues

$$E_{\text{sym}} = \frac{h^2}{2}\left(\frac{L(L+1)}{\mathfrak{I}_0} + \left(\frac{1}{\mathfrak{I}_3} - \frac{1}{\mathfrak{I}_0}\right)K^2\right) \quad (1.50)$$

This expression is model independent and only depends upon the ratio of the moments of inertia $\mathfrak{I}_0/\mathfrak{I}_3$. There have been recent observations of "superdeformation" for rare earth nuclei with spheroidal-axes ratios of 2 : 1. For symmetric rotational systems with $K = 0$, the ratio $R4/2$ is 10/3 rather than 2 for the vibrational systems. Thus, this ratio differentiates well between rotational and vibrational systems. In recent years, there was quite some interest in how these shape changes occur [33, 34], as they can be treated as quantum-phase transitions. Atomic nuclei that have an $R4/2$ ratio of around 2.9 lie at the critical point of the spherical-deformed phase transition.

1.6.5 Odd-A Nuclei

1.6.5.1 Single-Particle Levels

Near the magic numbers, the ground and lower excited states appear to be mainly single-particle states, that is, states that do not seem to possess collective properties. For instance, near the doubly magic nucleus $^{16}O_8$, one finds that both $^{15}N_8$ and $^{15}O_7$ have $1/2^-$ ground states and nearly identical lower excited states. This shows that the $N = Z = 8$ shell is filled by a $p_{1/2}$ orbit. Above the closed shell are $^{17}F_8$ and $^{17}O_9$ both of which have the expected $5/2^+$ ground state from the $d_{5/2}$ orbit. These data clearly support the notion that angular momentum shells are filled in a manner similar to the electronic shells of the elements.

1.6.5.2 Vibrational Levels

The nucleons all possess intrinsic spin 1/2 and angular momentum l, so that the simplest model for vibrational odd-A nuclei that one might construct is to add a single nucleon to an even–even core in which the nucleons of each kind pair to zero angular momentum. If the core is a good vibrational nucleus ($R4/2 \sim 2$), then the ground state of the odd-A system will have the angular momentum properties of the odd nucleon. Furthermore, by coupling this nucleon to the excited states of the core, one should expect to be able to determine the angular momenta of the odd-A nucleus's excited states. An example is the coupling of a neutron to the vibrational nucleus $^{110}Cd_{62}$ (Section 1.6.2) to yield the odd-A nucleus $^{111}Cd_{63}$, which has a ground-state spin of $1/2^+$. One might expect that there would be two excited states built on the first excited 2^+ state with spins $3/2^+$ and $5/2^+$. This nucleus does indeed have a $3/2^+$, $5/2^+$ pair as its first excited states, which are at 245.4 and 342.1 keV. Coupling of the $1/2^+$ neutron with the two-phonon states would lead to excited state spins from $1/2^+$ through $9/2^+$. Such a set of states does appear at higher energies. However, it is difficult to pick out which states might belong to the two-phonon core states. Below them there is a first negative parity state with spin $11/2^-$ at 396.22 keV, which indicates that one has the presence of a low-lying $h_{11/2}$ neutron orbit.

1.6.5.3 Rotational Levels

One might expect that a simple model for the odd-A rotational levels would follow that of the vibrational levels. However, the experimental data do not show such a similar structure. As an example we consider $^{157}Gd_{93}$ which has a $3/2^-$ ground state. Following the reasoning for vibrational nuclei, one expects from the coupling with the first excited 2+ state a multiplet with spins $1/2^-$, $3/2^-$, $5/2^-$, and $7/2^-$. Instead, one observes a rotational-like band with the spin sequence: $5/2^-$, $7/2^-$, $9/2^-$, $11/2^-$, ... This is due to the fact that the ground-state band has also a given value of K and we conclude that $K = 3/2$.

1.6.6 Odd–Odd Nuclei

As noted in Section 1.6.1, there are only four stable odd–odd nuclei; however, a great deal of data exist for all of this class of nuclei. Again, one might feel that these nuclei would be well represented by an even–even core plus a neutron–proton pair. Even if these are coupled to a spin-zero core, there is still considerable ambiguity as to how these couple to form the ground-state spin of the system. For instance, if the proton and neutron angular momenta are j_p, j_n, respectively, then their vector sum

can lead to the range of J values

$$\left|j_p - j_n\right| \leq J \leq j_p + j_n \tag{1.51}$$

One can obtain the values of j_p and j_n from the ground-state spins of the neighboring odd-A nuclei with one less neutron or one less proton, respectively. In order to determine more exactly the range of the possible ground-state spins of these nuclei, one must make use of Nordheim's rules, which are the following:Strong rule, for $\eta = 0$:

$$J = \left|j_p - j_n\right| \tag{1.52}$$

Weak rule, for $\eta = \pm 1$:

$$J = \left|j_p - j_n\right| \text{ or } j_p + j_n \tag{1.53}$$

with

$$\eta = (j_p - l_p) + (j_n - l_n) \tag{1.54}$$

In general, these do not work all of the time. To predict the excited states spins is even more difficult, making odd–odd nuclei the most difficult to describe.

1.7 Nuclear Models

1.7.1 Introduction

The nucleus, being a many-body collection of A particles interacting through a short range but strong force, is not easily dealt with. From a theoretical point of view, making use of an accurate many-body calculation is impossible because these are beyond the scope of current theoretical methods. Thus, calculations based upon various simple models are able to provide a good deal of physical insight as well as making predictions of measurable properties, providing a way to categorize the vast amounts of data available from nuclear physics laboratories. While these approaches may seem to be organized in an historical order, all are to greater or less extent being continuously refined.

The nuclear many-body problem is not solved definitely, owing to the large number of strongly interacting particles. To illustrate the difficulties, we take the example of ^{208}Pb. If we consider up to two-body interactions between the nucleons, the general Hamiltonian is given by

$$H = \sum_{i=1}^{208} t_i + \frac{1}{2}\sum_{i \neq j}^{208} V_{ij} \tag{1.55}$$

The radius of this nucleus is 7 fm, while the strong force between two nucleons varies significantly on the scale of 0.2 fm. If one would calculate only the two-body interaction on a mesh, one should foresee for each nucleon at least 100 mesh points in each direction and for all 208 nucleons the number of possible combinations would be $100^{3.208}$. For a fast computer today 10^9 operations per second are feasible, which gives 10^{16} operations per year. Thus to perform the calculation one would need 10^{1232} years, while the age of our universe is only 2×10^{10} years.

As an exact treatment is impossible, one of the major characteristics of nuclear structure calculations is the choice of the right approximations and truncations needed to make a calculation feasible while keeping the essential physics. Of great help thereby is the extensive use of symmetry concepts, which allow one to simplify the many-body problem and make the calculations tractable. Many quite different approaches have turned out to be

successful for certain classes of atomic nuclei. Most theoretical models rely on spherical symmetry, which allow the use of the very powerful "Racah machinery," originally developed for atomic physics by Guilio Racah. The introduction of irreducible tensors and coupling coefficients then provides the necessary technical skills to perform nuclear structure calculations.

The first approach is based on the experimental observation of magic numbers and leads to the nuclear shell model as described in Section 1.7.2. In Section 1.7.3 we introduce the deformed shell model, which allows the description of medium-heavy nuclei that are deformed. Section 1.7.4 is devoted to the collective model in which medium heavy and heavy nuclei are described by a quantum liquid droplet.

In Section 1.7.5, we will deal with a new theoretical approach to many-body problems originating from nuclear physics: the interacting boson approximation (IBA). Although the IBA is a method that was developed to describe the atomic nucleus, it has since been applied to molecules, quarks, and fullerenes. The major assumption of the interacting boson model (IBM) is twofold. First, because of the shell structure evinced by many experiments, only the valence nucleons are considered to be important for the low-energy excited states of the atomic nucleus. Those move to a first approximation in orbits formed by their commonly created average field, which is spherical symmetric. Secondly, between like nucleons, either protons or neutrons, there is a dominant pairing force when they occupy the same orbit. This allows to replace coupled fermion pairs by real bosons that are simpler to deal with.

1.7.2 The Spherical-Shell Model

One of the first and best understood nuclear models is the nuclear shell model. The major assumption of this model is that nucleons up to a good approximation move independently in a spherically symmetric averaged field $U(r_i)$ created by the other nucleons. This means that the many-body Hamiltonian is rewritten as

$$H = H_0 + V_{\text{res}} \tag{1.56}$$

with

$$H_0 = \sum_{i=1}^{N} t_i + \sum_{i=1}^{N} U(r_i) \tag{1.57}$$

and the residual interaction V_{res} is given by

$$V_{\text{res}} = \frac{1}{2}\sum_{i\neq j}^{N} V_{ij} - \sum_{i=1}^{N} U(r_i) \tag{1.58}$$

The shell model uses the ansatz (Eq. (1.57)) in which the single particles move in a potential well provided by the other particles. While it might seem improbable that nuclei made up of a large number of strongly interacting particles could in any way be represented by such a model, one must recall that the Pauli principle prevents most interactions of a nucleon in the nucleus. In general, scattering cannot take place, as most of the states into which a low-energy nucleon can scatter are filled. This implies that the mean free path for nucleon motion is quite long, and a single-particle model is worth exploring.

One obvious and important characteristic of odd-A nuclei is their ground-state spins and parities. Starting with the lightest nuclei, the general order of odd-A ground-state spins and parities is $1/2^+$ for

neutron and proton and after the strongly bound ^{4}He $3/2^-$, $1/2^-$ then, after $^{16}O_8$, it becomes $5/2^+$, $3/2^+$, $1/2^+$ and then, after $^{40}Ca_{20}$, $7/2^-$, $5/2^-$, $3/2^-$, $1/2^-$ an order reminiscent of the atomic shells in which the alternating parities are related to the orbital angular momentum suggesting a sequence: s-p-d and s-f and p to which a spin one-half is coupled. To obtain a model with such an ordering of angular momentum l-values requires solving a single-particle Schrodinger equation with a reasonable potential function. The form of this potential is not too important because only minor changes in l-value order occur using one or another r dependence. The two simplest to calculate are the isotropic harmonic oscillator and the infinite square well. The latter is not realistic because the nucleon separation energy is quite finite and well-known for most nuclei. (One might also consider a potential function related to the charge distributions of Eqs. (1.16) and (1.19); however, these would require extensive numerical computing to extract the eigenvalues.) The angular momentum order of the levels and the number of particles in each for the isotropic harmonic oscillator is 1s (2); 1p (6), 1d, 2s (12); If, 2p (20); lg, 2d, 3s (30); and so on. (The semicolons separate the levels with different quanta of energy $\hbar\omega$, and most are degenerate.) The total number of levels thus becomes 2, 8, 20, 40, 70, 112, 168, which only agrees with the lowest observed ''magic numbers'' 2, 8, 20, 28, 50, 82, and 126 (cf. Section 1.6.1).

It took the genius of Maria Goeppert Mayer [35] and others [36] to realize that by adding a spin–orbit term of the form

$$V_{ls} = -V(r)\,\vec{l}\cdot\vec{s} \tag{1.59}$$

the theoretical order could be brought into agreement with that observed. The minus sign guarantees that the $j = l + 1/2$ level will lie below the $j = l - 1/2$ level, as observed and the dependence on l makes that the effect increases with l leaving the shell closures at 2, 8, 20 unaffected while introducing a new shell closure at 28 due to the lowering of the $1f_{7/2}$ orbit. Further, the observed closures at 50, 82, 126, are obtained in a natural way. Also the spins of the odd-A ground states is mostly correctly predicted. The shell model explains via large gaps between the single-particle orbital energies the large binding energies of nuclei having Z or N a magic number making the number of stable isotopes/isotones particularly large, for example, there are 10 stable Sn isotopes with $Z = 50$. For the excited states, magic nuclei show them at very high excitation energies as one needs to break the closed shell and lift one particle over the large shell gap. As an example in the doubly magic nucleus ^{208}Pb, an energy of 2.6 MeV is needed to create the first excited 3^- state and even 4.085 MeV for the 2^+ state.

The robustness of double-magic nuclei allows a major truncation for shell model calculations. It turns out that they can in first instance be treated as inert cores, whereby the spins of all of the particles in the core pair off to zero. Thereby, only the so-called valence nucleons need to be considered. This is of importance as with increasing number of nucleons and possible orbitals, the shell model allows for many different configurations. As angular momentum and parity are conserved, the residual interaction used can be diagonalized separately for a given spin and parity. Nevertheless, the dimensions become quickly huge.

Although the shell model is one of the older nuclear models, it has become of increasing use in the last decades. This is due to two key aspects. The first is

the tremendous increase in computing power allowing large-scale shell model codes to be developed, including quantum Monte Carlo diagonalization. The second is the construction of effective *in medium* interactions starting from the nucleon–nucleon scattering data. Hereby the elimination of the hard core repulsion of the nucleon–nucleon interaction was crucial and produced a universal, so-called $V^{\text{low-}k}$ interaction. The recent progress opens the direct way from quantum chromodynamics (QCD) to the effective interactions, which is in particular very much needed for the study of exotic nuclei. The most involved calculations are done for light nuclei up to ^{11}B. Here no-core shell model calculations are able to describe accurately binding energies and excited states for very light nuclei [37].

1.7.3 The Deformed Shell Model

As the very large quadrupole moments of the nuclei in the middle of the higher shells cannot be explained by the simple spherical model, their explanation requires large contributions from the even–even core, which carries all (or almost all) of the charge. It can be shown that a single-particle moving in a nonrigid potential well will have a lower energy if the well is not spherical, rather than if it is spherical. As the lowering of the particle energy is proportional to the eccentricity, such a system will assume a deformed equilibrium shape. The nuclear Hamiltonian will now contain a term that produces deviations from spherical:

$$H = H_0 + H_{\text{d}} - V_{\text{ls}}(r) + \text{cl}^2 \qquad (1.60)$$

where H_0 is the isotropic harmonic oscillator Hamiltonian and $V_{\text{ls}}(r)$ is the spin–orbit term, both of which have proved so successful for the spherical-shell model. (The I^2 term helps give the proper level order in the isotropic limit.) Expanding the nuclear surface in spherical harmonics gives

$$S(\theta,\phi) = R_0\left[a_0 + \sum_{\lambda>1,\mu} a_{\lambda\mu} Y_{\lambda\mu}(\theta,\phi)\right] \qquad (1.61)$$

where R_0 is the radius of the spherical nucleus and a_0 is unity to quantities of second order and assures that the volume remains constant for small deformations. Taking the lowest order, one has that

$$S(\theta,\phi) = R_0\left[1 + \sum_{-2}^{+2} a_{2\mu} Y_{2\mu}(\theta,\phi)\right] \qquad (1.62)$$

[This expression should be compared with the nuclear charge distribution of Eq. (1.19).] It is customary to take only the symmetric term a_{20} in this equation to be nonzero. This then leads to the deformation term in the Hamiltonian

$$H_{\text{d}} = -Ka_{20}\omega_0^2 r^2 Y_{20}(\theta,\phi) \qquad (1.63)$$

The full Hamiltonian can now be diagonalized using a set of spherical shell model basis states, and the resulting deformed, single-particle states are called *Nilsson* states after Sven Gosta Nilsson [38], who first carried out this diagonalization. The operators in Eq. (1.60) do not connect oscillator states differing by one principal oscillator quantum number, N, but do connect states differing by two. Nilsson neglected matrix elements not diagonal in N. This is a reasonably good assumption as oscillator states differing by $N = 2$ are rather far apart, except for states belonging to large N and large deformation (i.e.,

nuclei with large A and large quadrupole moments). The Nilsson levels are labeled $\Omega|N,n_3,\Lambda]$. The symbol Ω is the projection of the particle angular momentum j on the three axes of the body-fixed axis system and is a good quantum number. In the limit of very large deformations, the single-particle energy is given by

$$E = \left(n_3 + \frac{1}{2}\right)\hbar\omega_3 + (n_\perp + 1)\hbar\omega_\perp \tag{1.64}$$

where the symbol ω_3 is the oscillator frequency along the symmetry axis and the symbol $\omega_\perp$ is the oscillator frequency in the perpendicular directions. Thus, in the limit of very large deformations, one has $N = n_3 + n_\perp$. In this same limit, the projection of the orbital angular momentum on the symmetry axis becomes the last label,

$$N = \pm n_\perp, \pm n_\perp - 2, \ldots, \pm 1 \text{ or } 0 \tag{1.65}$$

This scheme is used to label not only the ground states of deformed nuclei but also the excited state rotational band heads. The level order will now be L, $L+1$, $L+2$, ..., as a result of core rotation, so that, when a level does not fit into this sequence, it must belong to a different single-particle band. One now proceeds to assign quantum numbers to the states by determining the odd-nucleon number; then, because the Nilsson energy levels are plotted against the core deformation, one assigns the deformation by finding that place where a Nilsson level appears with the measured ground-state spin. The next Nilsson level up should give the excited-band head, and so on. As an example, consider the nucleus $^{173}Lu_{102}$, which has a $7/2^+$ ground state. Following the procedure above, the appropriate level is the 7/2[404], with excited members $9/2^+$, $11/2^+$, $13/2^+$. Just above the $9/2^+$ level is a close doublet with spins $5/2^-$, $1/2^-$, which has been assigned to the 1/2[541] Nilsson level. The $3/2^-$ level belonging to this band lies above the $9/2^+$ state. This nonuniform order of a 1/2 band is an example of Coriolis coupling, which mixes the level order of $\Omega = 1/2$ bands.

1.7.4 Collective Models of Even–Even Nuclei

In Section 1.6.4.2, where the rotational-level structures of even–even nuclei were discussed, it was pointed out that the moments of inertia play a crucial role to describe the atomic nucleus as a rigid-body momental ellipsoid. Bohr [32] proposed a model of the momental ellipsoid and thus of deformed even–even nuclei and the cores of odd-A and odd–odd nuclei. Several important assumptions were made in order to extract these moments of inertia. The first is that the nuclear core is incompressible, so that the nuclear density is constant. Second, the flow is assumed to be irrotational, so that a velocity potential exists and satisfies Laplace's equation. Finally, it must be assumed that the motion is such as to preserve the principal axis system. The moments of inertia for quadrupole deformations then are

$$\mathfrak{I}_k^2 = 4B_2\beta^2\sin^2\left(\gamma - \frac{2\pi}{3}k\right) \tag{1.66}$$

Here B_2 and β are the quadrupole mass and deformation parameters, while γ is an asymmetry parameter whose range is $0 \le \gamma \le \pi/6$. There are similar relations for other deformations such as octupole ($Y3$), hexadecapole ($Y4$), and so on. The Hamiltonian for this system is now

$$H_{\text{quad}} = \frac{1}{2}B_2(\dot{\beta}^2 + \beta^2\dot{\gamma}) + \frac{1}{B_2\beta^2}\sum_{k=1}^{3}\frac{L_k^2}{\sin^2\left(\gamma - \frac{2\pi}{k}\right)} + \frac{1}{2}C_2\beta^2 \tag{1.67}$$

If the system is not rigid, then the deformation and asymmetry parameters are no longer fixed, and the system can execute either deformation (β) vibrations or asymmetry (γ) vibrations, or both. The β-vibrational bands have angular momentum 0 – they are akin to molecular "breathing" modes – and the γ-vibrational bands have angular momentum 2. Thus, nuclei will exhibit ground-state bands where $n_\beta = 0$ and $n_\gamma = 0$, β-bands $n_\beta = 1$, $n_\gamma = 0$), γ-bands ($n_\beta = 0$, $n_\gamma = 1$), and higher order bands. Although the β-vibrations are the simplest, they are not really well established [39], much better established are γ-bands of which double excitations with $n_\gamma = 2$ are found in nature [40].

In his original paper, Bohr [32] argued that the nucleus would stabilize around $\gamma = 0$, so that the momental ellipsoid was symmetric (long and thin, cigarlike), while the surface itself was spheroidal. Then $\Im_3 = 0$. This leads to difficulties as Eq. (1.50) shows that the energy eigenvalues become infinite unless $K = 0$. Thus, the ground-state bands of even–even nuclei have $K = 0$, and the angular momentum sequence becomes 0, 2, 4, 6, ..., which is what is observed. These models automatically conform to the A representation. Again, if the system is not rigid, against β-vibrations, there will be one or more β-bands above the ground-state rotational band that mimic the ground-state band in spin sequence, but because $<\beta^2>$ will be greater than for the ground state, the levels will be closer together.

Finally, in this model if γ is not constant, γ-vibrations can arise with angular momentum 2. (One might think of these as equatorial bulges moving around the nuclear equator. They are akin to the well-known Jacobi ellipsoids of a self-gravitating, rotating fluid.) These γ-vibrations then produce a γ-band that is different in spin sequence from the β- and ground-state bands. This occurs because $K_\gamma = 2$, not 0. The γ-band sequence is 2, 3, 4, 5, ... There can be several γ-bands with different K values and different spin sequences. The lowest one has $K_\gamma = 2$, while the next will have $K_\gamma = 0$ and 4, and so on.

It must be made clear that the Bohr model is only applicable for the low-lying positive-parity levels. The negative-parity levels belonging to the B_1 representation have moments of inertia that are considerably more complicated as they represent an octupole or pearlike shape.

An example of a well-developed rotational system is $^{154}Sm_{92}$ where $R4/2 = 3.25$. This nucleus has a well-developed ground-state rotational band with the spin sequence 0^+, 2^+, 4^+, ..., 12^+ with the first excited state rather low in energy at 81.99 keV. It has a β-band with a 0^+ band head 1099.28 keV, followed by the first excited 2^+ state at 1177.78 keV ($\Delta E = 78.52$ keV). This is a striking example of the increase in the moment of inertia due to the quantum of β-vibrational energy. Furthermore, the $K = 2$ γ-band, with sequence 2^+, 3^+, 4^+, ..., starts at 1440.05 keV.

Also to be seen is an octupole band starting with the 1^- level at 921.39 keV, followed by the 2^-, 3^-, 5^-, ... in the expected order. Thus, one sees four well-developed rotational bands fitting quite nicely into the model scheme.

Further, the collective behavior manifests itself in the quadrupole moments and electric quadrupole transitions. The axially

symmetric core collective model yields for a state with quantum numbers K and I a laboratory quadrupole moment Q:

$$Q = \frac{3K^2 - I(I+1)}{(I+1)(2I+3)} Q_0 \tag{1.68}$$

with an intrinsic quadrupole moment Q_0 given by

$$Q_0 = \frac{3}{\sqrt{5\pi}} Z R_0^2 \beta (1 + 0.16\beta) \tag{1.69}$$

The fact that by definition K has to be lower than I (Eq. (1.74)) produces almost always a change of sign between intrinsic and laboratory quadrupole moments. Related to the intrinsic quadrupole moments are the $B(E2)$ values. In between states belonging to a given K band, they are given by

$$B(E2; I_i \rightarrow I_f) = \frac{5}{16\pi} e^2 Q_0^2 \left\langle I_i K 20 | I_f K \right\rangle^2 \tag{1.70}$$

When comparing Eq. (1.70) directly with the data, one often extracts the transition quadrupole moment Q_t as being the Q_0 value obtained for a given transition. A very useful relation is obtained for $B(E2)$ ratios between states belonging to bands having different K quantum numbers. These relations are called the *Alaga rules* and yield

$$\frac{B(E2; I_i \rightarrow I_f)}{B(E2; I_i \rightarrow I'_f)} = \left(\frac{\left\langle I_i K 2 \Delta K | I_f K_f \right\rangle}{\left\langle I_i K 2 \Delta K | I'_f K_f \right\rangle} \right)^2 \tag{1.71}$$

whereby $\Delta K = K_f - K$ as is expected for an angular momentum projection quantum number. Using the Alaga rules it is possible to determine the K values of rotational bands.

If one does not follow the Bohr assumption that the nucleus will stabilize around $\gamma = 0$ but around some other value within its range, then one has to deal with an asymmetric rotator, which will have a considerably more complicated eigenvalue structure. Davydov and Filippov [41] made an extensive investigation and showed that one obtained better results with such a model than with a symmetric one. However, one loses the computational simplicity of the symmetric model.

1.7.5 Boson Models

The models discussed up until now and that have been used most extensively would seem to be almost mutually exclusive. The shell model and its modifications, such as Nilsson's extension to a deformed system, focus on the single-particle aspects of nuclear structure. On the other hand, the collective model of Bohr and Mottelson, and extensions, place the emphasis on the cooperative, or fluid, aspects of these systems. Thus, they seem to stand apart and be almost unrelated. This has led to the next step, the investigation of models that are called *microscopic models* or *calculations*. These include pairing models (similar to the Bardeen–Cooper–Schriefer theory of superconductivity) and the random-phase approximation, to name but two. A more recent and still developing set of microscopic models is called the *interacting boson models*, which we present here briefly.

The solutions of the nuclear shell model lead to single-particle orbits that can be occupied by the neutrons and protons. Because of the Pauli principle, each nucleon in an atomic nucleus can occupy only different orbits. It turns out that once more the short range of the nuclear force leads to a dominant component when

one wants to take account of the residual interaction. The dominant component is called *pairing interaction* and energetically favors the formation of nucleon pairs made out of two nucleons moving in orbits |njm> and |nj-m>, coupling to a two-nucleon state with total spin $J=0$. To a lesser extent, the formation of nucleon pairs with spin $J=2$, 4, 6 in that sequence is also favored. As the number of protons and neutrons is generally different in not too light nuclei, the single-particle orbits occupied by both will generally be very different and thus the pairing will be effective only between like nucleons (either protons or neutrons). When one is mainly interested in the low-energy excitations of the atomic nucleus, one can treat this structure in terms of interacting, nucleon pairs via the residual interaction. However, a many-body problem composed of paired fermions, although more correct, is still very difficult to solve. Instead by replacing these $n=n_\nu+n_\pi$ fermions by

$$N=\frac{n_\nu+n_\pi}{2} \tag{1.72}$$

real bosons with $L=0$, 2 (and eventually 4, 6, ...), one has enormously simplified the original n-body problem. For instance, one does not need to bother about which orbits are occupied by the protons and neutrons, while keeping the essential physics. This is the essence of what is now called the *interacting boson approximation*, which was proposed and elaborated by Iachello and Arima in the second half of the 1970s [42].

The simplest version of the model is called *IBM-1* and deals with s ($L=0$) and d ($L=2$) bosons. Considering the constraint that the number of bosons is fixed $N=n_\mathrm{s}+n_\mathrm{d}$ and limiting the Hamiltonian to one- and two-body interactions, one obtains the multipole expansion form

$$\begin{aligned}\widehat{H}=\varepsilon\widehat{n}_d+c_1\widehat{L}\cdot\widehat{L}+c_2\widehat{Q}\cdot\widehat{Q}\\ +c_3\widehat{T}_3\cdot\widehat{T}_3+c_4\widehat{T}_4\cdot\widehat{T}_4\end{aligned} \tag{1.73}$$

with the operators

$$\widehat{L}_m=\sqrt{10}(d^\dagger\times\widetilde{d})^{(1)}_m \tag{1.74a}$$

$$\begin{aligned}\widehat{Q}_m=\ &(s^\dagger\times\widetilde{d}+d^\dagger\times s)^{(2)}_m\\ &+\chi(d^\dagger\times\widetilde{d})^{(2)}_m\end{aligned} \tag{1.74b}$$

$$\widehat{T}_{3m}=(d^\dagger\times\widetilde{d})^{(3)}_m \tag{1.74c}$$

$$\widehat{T}_{4m}=(d^\dagger\times\widetilde{d})^{(4)}_m \tag{1.74d}$$

The factor $\sqrt{10}$ is used in order to make the dipole operator an angular momentum operator and the first term in Eq. (1.73) contains the d-boson number operator. The multipole expansion has six free parameters of which two are needed for the quadrupole–quadrupole interaction. The use of the multipole expansion is especially advantageous because numerical studies have shown that the values of the octupole and hexadecapole interaction are generally very small. Thus, up to a good approximation, one can describe a variety of atomic nuclei at low excitation energies by the simple four-parameter Hamiltonian:

$$\widehat{H}^\chi=\varepsilon\widehat{n}_d+c_1\widehat{L}\times\widehat{L}+c_2\widehat{Q}^\chi\times\widehat{Q}^\chi \tag{1.75}$$

The index in Eq. (1.75) makes explicit that the quadrupole operator contains an additional parameter.

Of importance in the development of the IBM is the use of dynamical symmetries providing analytic solutions of the many-body problem. Because there are in total six distinct bosons when considering the magnetic projections, the symmetry of the many-body problem is $U(6)$ while the physical symmetry is given by SO(3) the Lie algebra of the angular momentum. Three

dynamical symmetries are obtained. They can be schematically represented by the chains

$$U(6)\left\{\begin{array}{l}\supset U(5)\\ \supset O(6)\\ \supset SU(3)\end{array}\right\}\quad \supset O(5)\quad \left.\right\}\supset O(3) \tag{1.76}$$

Each of the three chains can be used to provide a complete basis to analytically solve the N s, d boson model. One can also rewrite the general Hamiltonian as a combination of all previously determined Casimir operators:

$$\begin{aligned}\widehat{H} = {} & \varepsilon C_1[U(5)] + \alpha C_2[U(5)] \\ & + \delta C_2[SU(3)] + \eta C_2[O(6)] \\ & + \beta C_2[O(5)] + \gamma C_2[O(3)]\end{aligned} \tag{1.77}$$

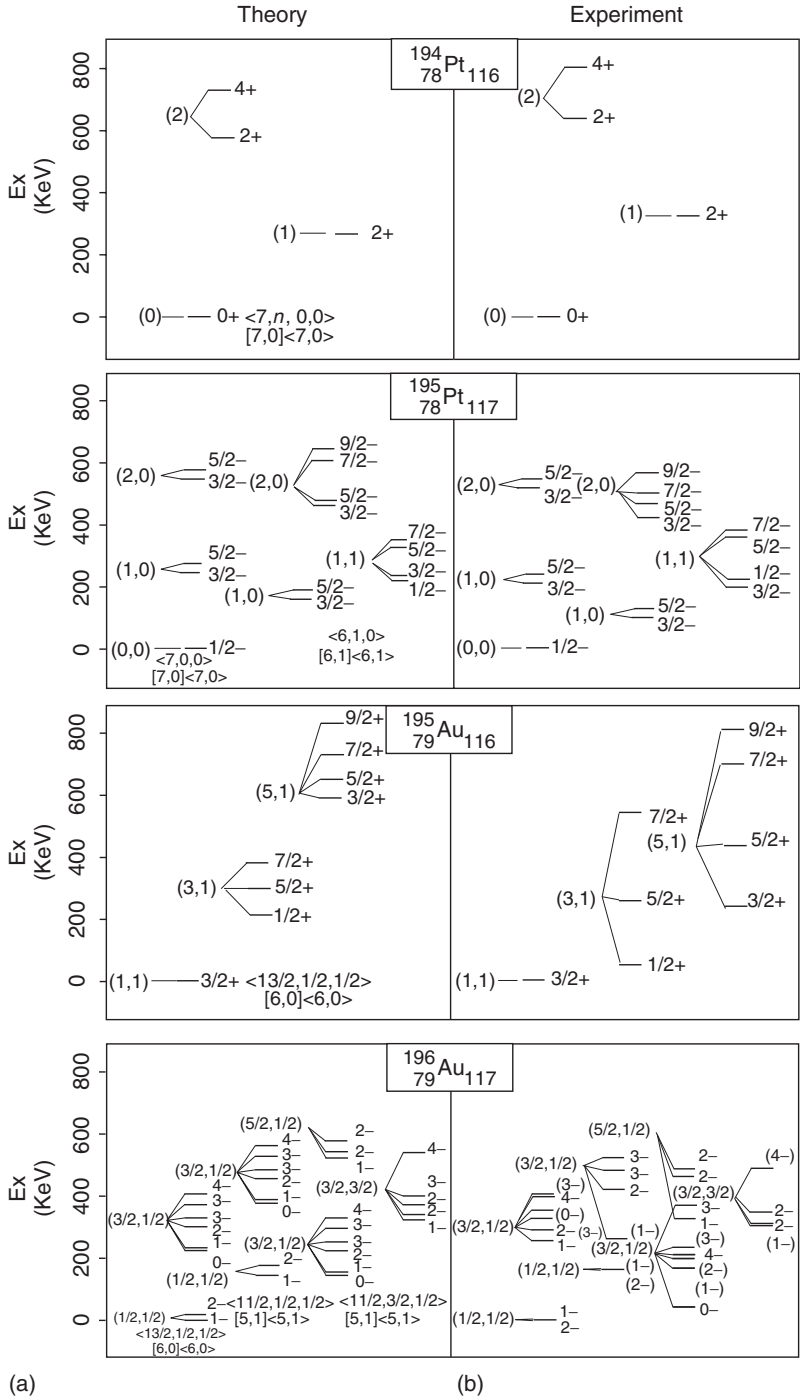
where we have omitted all constant terms that contribute only to the binding energies. The Casimir form Eq. (1.77), like Eq. (1.73), contains six parameters and, once more, all three forms can be transformed into each other. The advantage of the form Eq. (1.77) lies in the fact that the three different limits correspond to three classes of atomic nuclei: vibrational, rotational, and γ-unstable nuclei correspond to the Hamiltonian with $\delta = \eta = 0$ ($U(5)$ limit), $\varepsilon = \alpha = \eta = \beta = 0$ (SU(3) limit) and $\varepsilon = \alpha = \delta = 0$ ($O(6)$ limit) . In these cases, closed analytic expressions for excitation energies and wavefunctions can be obtained [42]. Two of the limits already had known counterparts in the collective model. It was the discovery of the third, the $O(6)$ limit, in the Pt nuclei at the end of the 1970s [43], that boosted the use of the model. Soon it became evident that it formed a good approximation for the nuclear many-body problem in many nuclei. Outside the limits the model needs to be solved numerically and the transitional nuclei can be described.

The IBM-1 model describes only positive-parity levels of even–even nuclei. To describe negative-parity states, one must use negative parity p and f bosons constructed in a similar way to the s and d bosons. Furthermore, the IBM-1 has no explicit neutron–proton degree of freedom. This is introduced in the IBM-2 where a neutron–proton label is introduced for the s and d bosons. As the neutron and proton bosons are distinguishable, the IBM-2 leads to new sets of states, called *mixed symmetry states* (in the neutron–proton degree of freedom). Furthermore, analytical solutions can be derived for the IBM-2 [44]. Mixed symmetry states have been observed in many atomic nuclei, especially in deformed nuclei where the lowest state is a 1^+ state called the *scissor state* [45], and also recently in vibrational nuclei [46].

Finally, if one wants to describe odd-A nuclei, a model denoted interacting boson–fermion model (IBFM) must be used. Odd-A nuclei always have an unpaired nucleon that cannot form an s or d boson because a partner is missing. Therefore, an odd–even nucleus is described in the model with N bosons and one fermion, the unpaired nucleon, which can occupy several orbitals j. In addition, for such systems dynamical symmetries, called *Bose–Fermi symmetries*, can be used, although their elaboration is much more complex [47]. The symmetry starts then from the product $U^{\mathrm{B}}(6) \times U^{\mathrm{F}}(M)$ with

$$M = \sum_j 2j + 1 \tag{1.78}$$

and ends with SU(2). In 1980, a daring extension of the model was proposed [48]. Using supersymmetry, the structure of some odd–even nuclei could be related


Theory
Experiment
194 78 Pt 116
195 78 Pt 117
195 79 Au 116
196 79 Au 117
Ex (KeV)
0
200
400
600
800
(a)
(b)

to that of the much simpler even–even nuclei via the embedding of the dynamical symmetries in a supersymmetry:

$$\begin{array}{ccccc} U(6/M) & \supset & U^{\mathrm{B}}(6) & \times & U^{\mathrm{F}}(M) \\ | & & | & & | \\ [N] & & [\mathrm{N}] & & [1^m] \end{array} \qquad (1.79)$$

connecting the N s, d bosons + 1 fermion problem to the N + 1 boson problem by the supersymmetric reduction rule $N = \mathrm{N} + m$. This embedding when realized in nature would reveal itself through the possibility to describe excited states in an odd–even nucleus and in the adjacent even–even nuclei using a common set of quantum numbers and parameters related via the supergroup. If dynamical symmetries are required, only a limited class of atomic nuclei can be described in this way. During the 1980s, experimental evidence was accumulated, but unfortunately it was not possible to determine completely the structure of the odd–even nucleus starting from the even–even nucleus alone.

Later an extension of supersymmetry was proposed, which incorporates a distinction between neutrons and protons. This extended supersymmetry allows one to describe a quartet of nuclei in a common framework. This quartet, called a *magic square*, consists of nuclei having the same total number of bosons (paired nucleons) and fermions (unpaired nucleons). It consists of an even–even nucleus, two odd–even nuclei and an odd–odd nucleus. The latter is interesting in two respects: its energy spectrum can be predicted from those of the other three members and heavy odd–odd nuclei are the most complex objects found in low-energy nuclear structure. If dynamical supersymmetry is able to describe these nuclei, which could not be described by other theoretical models, strong evidence for its existence can be obtained. This development is of importance not only for nuclear physics but for all other applications of supersymmetry in physics.

The odd–odd nucleus ^{196}Au is considered to be the ultimate test of supersymmetry in nuclear physics for three reasons. It is situated in a region of nuclei that was known to exhibit dynamical symmetries and supersymmetries. At the same time, it is the most difficult mass region in which to describe odd–odd nuclei. Finally, when its excitations with negative parity were predicted in 1989 on the basis of the three other nuclei, none of them were experimentally known. In a detailed spectroscopic study the excited states of ^{196}Au were determined and a good agreement with the supersymmetric model was obtained [49]. The results of this combined effort are illustrated in Figure 1.6. They reveal good agreement with the analytical prediction:

Figure 1.6 The lowest observed states (b), are compared to the theoretical result (a) for the four members of the quartet (note that one is dealing with holes in the doubly closed shell nucleus ^{208}Pb). The even–even nucleus ^{194}Pt, the negative parity states in the odd-neutron nucleus ^{195}Pt, the positive parity states in the odd-proton nucleus ^{195}Au and the negative-parity states in the odd–odd nucleus ^{196}Au up to 500 keV can be described using the formula (Eq. (1.88)) based on dynamical supersymmetry. Not shown are higher-lying states of the even–even and odd–even nuclei, which are also described by supersymmetry. Together with the spin and parities, the quantum numbers $[N_1,N_2]$, $<\Sigma_1,\Sigma_2>$, $<\sigma_1,\sigma_2,\sigma_3>$, and (τ_1,τ_2) are also indicated. (Based on [49].)

$$E = A(N_1(N_1+5) + N_2(N_2+3)) + B(\Sigma_1(\Sigma_1+4) + \Sigma_2(\Sigma_2+2)) + B'(\sigma_1(\sigma_1+4) + \sigma_2(\sigma_2+2) + \sigma_3^2) + C(\tau_1(\tau_1+3) + \tau_2(\tau_2+1)) + DL(L+1) + EJ(J+1) \qquad (1.80)$$

obtained using dynamical supersymmetry.

Glossary

Bosons: Quantum particles obeying Bose–Einstein statistics, so having integral spin.

Daughter: In a radioactive decay in which the radioactive nucleus A decays into nucleus B the nucleus B is said to be the daughter of nucleus A.

de Broglie wavelength: The wavelength λ of a particle given by the relation $\lambda = hp$, where h is Planck's constant and p the momentum of the particle.

Decay constant λ: If N nuclei are present at time t, then the decay constant is given by the number of decaying nuclei $-\mathrm{d}N/\mathrm{d}t$ divided by N.

Electron capture: A radioactive decay process, sometimes called *K capture*, where the decay energy Q is greater than the binding energy of one of the atom's cloud of electrons ($Q > BEe$). The nucleus absorbs one of the atomic electrons to undergo β^+ decay and emits thereby a neutrino. If $Q > 2m_e c^2$, then electron capture competes with normal β^+ decay, whereby a positron and an electron neutrino are created.

Fermions: Half integer spin particles obeying Fermi–Dirac statistics.

Half-life $T_{1/2}$: In radioactive decay, the time in which half of the nuclei initially present decay.

Isobars: Nuclei with the same mass number A but with different numbers of protons (Z) and neutrons (N).

Isomeric state: An excited nuclear state whose half-life for γ-emission is quite long; similar to a metastable state of an atomic system.

Isotones: Nuclei with the same number of neutrons (N) but different numbers of protons (Z) and thus, different values for A.

Isotope shift: Small changes in the wavelengths of X-ray optical, and especially muonic transitions in going from one isotope to the next, which give a measure of the change in the nuclear radius as A changes by one unit.

Isotopes: Nuclei with the same number of protons (Z) but different numbers of neutrons (N) and thus, different values of A.

Lamb shift: A small quantum electrodynamic effect that is principally due to the radiative coupling of the orbiting electron (or muon) with the vacuum field and the finite nucleus.

Lifetime τ: The reciprocal of the decay constant. The lifetime is related to the half-life by $T_{1/2} = \ln(2)\tau$.

Mass defect: The difference between the mass of a nucleus of mass number A in atomic mass units less A, $\Delta = M(A) - A$.

Mass excess: Negative of *mass defect*.

Mirror pair: Two light nuclei with the same mass number A but with numbers of protons (Z) and neutrons (N) interchanged.

Muon: A member of the lepton family of elementary particles with spin one-half. The muon is 207 times as heavy as an electron and has a lifetime of

2.197×10^{-6} s. The muon, like the electron, has a negative charge.

Nuclear polarization: A small reduction in the Coulomb potential caused by the penetration of the nuclear volume by the bound orbiting particle (an electron or muon).

Packing fraction: The mass defect per unit mass number $P = \Delta/A$.

Parity π: The behavior of a state function upon reflection of the coordinates through the origin, $r \rightarrow -r$, then either $\phi(-r) = +\phi(r)$ and the state is said to have *positive parity* or $\phi(-r) = -\phi(r)$ and the state is said to have *negative parity*.

Pauli principle: The requirement that in a system of like particles obeying Fermi–Dirac statistics (half integer spin particles), no two particles can have the same set of quantum numbers. More generally, the Pauli principle requires wavefunctions for identical particles to be antisymmetric for fermions and symmetric for bosons upon interchange of two particles.

Phonon: In nuclear physics a phonon is a quantized surface vibration of a quantum fluid.

Vacuum polarization: A small radiative correction to the Coulomb potential of an atom arising from the emission and reabsorption of virtual positron–electron pairs.

Valley of stability: In a three-dimensional plot of the nuclear masses $M(Z,N)$ with the proton number Z as abscissa neutron number N as ordinate and $M(Z,N)$ normal to the Z–N-plane, the stable nuclei will be found along a region where $N \sim Z$, which appears to form a deep valley – *the valley of stability*.

Woods–Saxon potential: For a real potential function, this is written with an adjustable surface diffuseness parameter and a radius parameter R. For a complex potential, one adds a similar term, V_{Im}.

References

1. Rutherford, E. (1911) *Philos. Mag.*, **21**, 669–688.
2. Moseley, H.G.J. (1913) *Philos. Mag.*, **26**, 1024–1034.
3. Moseley, H.G.J. (1914) *Philos. Mag.*, **27**, 703–713.
4. Soddy, F. (1913) *Nature*, **92**, 399–400.
5. Rutherford, E. (1919) *Philos. Mag.*, **37**, 581–586.
6. Harkins, W.D. and Majorsky, S.L. (1922) *Phys. Rev.*, **19**, 135–156.
7. Pauli, W. (1924) *Naturwissenschaften*, **12**, 741–743. reprinted in Kronig, R., Weisskopf, V. F. (eds) (1964) *Collected Scientific Papers*, Vol. 2, New York: John Wiley & Sons, Inc., pp. 198–200.
8. Racah, G. (1931) *Z. Phys.*, **71**, 431–434.
9. Chadwick, J. (1932) *Nature*, **129**, 312.
10. Heisenberg, W. (1932) *Z. Phys.*, **77**, 1.
11. Bethe, H.A. and Bacher, R.F. (1936) *Rev. Mod. Phys.*, **8**, 82–229.
12. Bethe, H.A. (1937) *Rev. Mod. Phys.*, **9**, 69–244.
13. Avignone, F.T. III, Elliott, S.R., and Engel, J. (2008) *Rev. Mod. Phys.*, **80**, 481–516.
14. Lederer, C.M. and Shirley, V.S. (eds) (1978) *Table of Isotopes*, 7th edn, John Wiley & Sons, Inc., New York.
15. Kudomi, N., Ejiri, H., Nagata, K., Okada, K., Shibata, T., Shima, T., and Tanaka, J. (1992) *Phys. Rev.*, **46**, R2132–R2135.
16. Bohr, A. and Mottelson, B. (1969) *Nuclear Structure*, Vol. I, Benjamin, New York.
17. Myers, W.D. and Swiatecki, W.J. (1966) *Nucl. Phys.*, **81**, 1–60.
18. Myers, W.D. (1977) *Droplet Model of Atomic Nuclei*, IFI/Plenum, New York, and references therein.
19. Möller, P., Nix, J.R., Myers, W.D., and Swiatecki, W.J. (1995) *Atomic Data Nucl. Data Tables*, **39**, 185–225.
20. Duflo, J. and Zuker, A.P. (1995) *Phys. Rev. C*, **52**, R23–R27.
21. Goriely, S., Samyn, M., and Pearson, J.M. (2007) *Phys. Rev. C*, **75**, 064312.

22. Lunney, D., Pearson, J.M., and Thibault, C. (2003) *Rev. Mod. Phys.*, **75**, 1021–1082.
23. de Vries, H., de Jager, C.W., and de Vries, C. (1987) *Atomic Data Nucl. Data Tables*, **36**, 495–536.
24. Tanihata, I. (1995) *Prog. Part. Nucl. Phys.*, **35**, 505–574.
25. Casimir, H.B.G. (1935) *Physica*, **2**, 719–723.
26. Gollnow, G. (1936) *Z. Phys.*, **103**, 443–453.
27. Schwartz, C. (1955) *Phys. Rev.*, **97**, 380–395.
28. Castel, B. and Towner, I.S. (1990) *Modern Theories of Nuclear Moments*, Clarendon Press: Oxford, and references therein.
29. Neyens, G. (2003) *Rep. Prog. Phys*, **66**, 633–689.
30. Eberth, J. and Simpson, J. (2008) *Prog. Part. Nucl. Phys.*, **60**, 283–337.
31. Heyde, K.L.G. (1990) *The Nuclear Shell Model*, Springer Series in Nuclear and Particle Physics, Springer(A).
32. Bohr, A. (1952) *Kgl. Dan. Vid. Selsk. Mat. Fys. Medd.*, **26** (14).
33. Jolie, J., Cejnar, P., Casten, R.F., Heinze, S., Linnemann, A., and Werner, V. (2002) *Phys. Rev. Lett.*, **89**, 182502.
34. Cejnar, P., Jolie, J., and Casten, R.F. (2010) *Rev. Mod. Phys.*, **82**, 2155–2212.
35. See also: Mayer, M. G. (1948) *Phys. Rev.* **74**, 235–239; *Phys. Rev*, (1949), **75**, 1969–1970.
36. Haxel, O., Jensen, J.H.D., and Suess, H.E. (1949) *Phys. Rev.*, **75**, 1766.
37. Pieper, S.C., Varga, K., and Wiringa, R.B. (2002) *Phys. Rev.*, **C66**, 044310.
38. Nilsson, S.G. (1955) *Kgl. Dan. Vid. Selsk. Mat. Fys. Medd.*, **29**, 16.
39. Garrett, P.E. (2001) *J. Phys.*, **G27**, R1.
40. Börner, H.G. *et al.* (1991) *Phys. Rev. Lett.*, **66**, 691 and 2837.
41. Davydov, A.S. and Filippov, G.F. (1958) *Nucl. Phys.*, **8**, 237–249.
42. Iachello, F. and Arima, A. (1987) *The Interacting Boson Model*, Cambridge University Press, Cambridge.
43. Casten, R.F. and Cizewski, J.A. (1978) *Nucl. Phys.*, **A309**, 477.
44. Van Isacker, P., Heyde, K., Jolie, J., and Sevrin, A. (1986) *Ann. Phys.*, **171**, 253–296.
45. Richter, A. (1995) *Prog. Particle Nucl. Phys.*, **34**, 261.
46. Pietralla, N., von Brentano, P., and Lisetskij, A.G. (2008) *Prog. Particle Nucl. Phys.*, **60**, 225–282.
47. Iachello, F. and Van Isacker, P. (1991) *The Interacting Boson-Fermion Model*, Cambridge University Press, Cambridge.
48. Iachello, F. (1980) *Phys. Rev. Lett.*, **44**, 772–775.
49. Groeger, J. *et al.* (2000) *Phys. Rev. C*, **62**, 064304.

Further Readings

In this list, the symbols E, I, and A following the references indicate the level of difficulty: elementary, intermediate, and advanced, respectively.

Bethe, H.A. (1936) *Rev. Mod. Phys.*, **8**, 82–229 (I).
This and the preceding form the best review of the early state of the subject: Bethe, H.A. and Bacher, R.F. (1937) *Rev. Mod. Phys.*, **9**, 69–244(I).
Bohr, A. and Mottelson, B. (1969) *Nuclear Structure*, Vol. 1, Benjamin, New York(A).
This and the preceding, while old, are comprehensive and still useful:
Bohr, A. and Mottelson, B. (1975) *Nuclear Structure*, Vol. 2, Benjamin, New York(A).
Bonatsos, D. (1988) *Interacting Boson Models of Nuclear Structure*, Clarendon, Oxford(A).
Castel, B. and Towner, I.S. (1990) *Modern Theories of Nuclear Moments*, Clarendon, Oxford(A).
An excellent text, more advanced than Pal (1983):
Casten, R.F. (1990) *Nuclear Structure from a Simple Perspective*, Oxford University Press, New York(I).
An overview of the use of symmetry concepts in nuclear structure with many experimental examples:
Frank, A., Jolie, J., and Van Isacker, P. (2008) *Symmetries in Atomic Nuclei: From Isospin to Supersymmetry*, Springer Tracts in Modern Physics, Springer (A).
A nice introduction into the practical aspects of the nuclear shell model:
Heyde, K.L.G. (1990) *The Nuclear Shell Model*, Springer Series in Nuclear and Particle Physics, Springer (A).
An excellent introduction for students of which each new edition is substantially updated:
Heyde, K.L.G. (1994) *Basic Ideas and Concepts in Nuclear Physics*, Fundamental and Applied Nuclear Physics, IOP (I).

Discusses many aspects, both theoretical and experimental, of this fascinating area of nuclear structure:
Janssens, R.V.F. and Khoo, T.L. (1991) Superdeformed nuclei. *Annu. Rev. Nucl. Part. Sci.*, **41**, 321–355 (A).

An ideal textbook for students:
Krane, K.S. (1987) *Introductory Nuclear Physics*, John Wiley & Sons, Inc., New York (I).

A compendium of nuclear data, of great value:
Lederer, C.M. and Shirley, V.S. (eds) (1978) *Table of Isotopes*, 7th edn, John Wiley & Sons, Inc., New York (A).

Nice treatments of several of the older microscopic models:
Pal, M.K. (1983) *Theory of Nuclear Structure*, Van Nostrand Reinhold, New York (E).

Very deep and detailed introduction in the algebraic methods used in the nuclear shell model and the interacting boson model:
Talmi, I. (1993) *Simple Models of Complex Nuclei, Contemporary Concepts in Physics*, Harwood Academic Publishers, Chur (A).

Another excellent, more advanced text:
Wong, S.S.M. (1990) *Introductory Nuclear Physics*, Prentice-Hall, Englewood Cliffs, NJ (I).

2
Nuclear Reactions

Carlos A. Bertulani

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

2.1 Introduction

With the collision of two nuclei, a nuclear reaction can take place and, as in the case of a chemical reaction, the final products can be different from the initial ones. This process occurs when a target is bombarded by particles coming from an accelerator or a radioactive substance. It was in the latter way that Rutherford observed, in 1919, the first nuclear reaction produced in the laboratory,

$$\alpha + {}^{14}_{7}\mathrm{N} \rightarrow {}^{17}_{8}\mathrm{O} + \mathrm{p} \qquad (2.1)$$

using α-particles from a ^{214}Bi sample.

As in Eq. (2.1), other reactions were induced using α-particles, the only projectile available initially. With the development of accelerators around 1930, the possibilities multiplied by changing the energy and mass of the projectile. At present, it is possible to bombard a target with protons of energy greater than 1 TeV ($1\ \mathrm{TeV} = 10^{12}\ \mathrm{eV} = 1.602 \times 10^{-7}\ \mathrm{J}$), and beams of particles as heavy as uranium are available for study of reactions with heavy ions.

Sometimes we can have more than two final products in a reaction, as in

$$\begin{aligned} \mathrm{p} + {}^{14}\mathrm{N} &\rightarrow {}^{7}\mathrm{Be} + 2\alpha, \\ \mathrm{p} + {}^{23}\mathrm{Na} &\rightarrow {}^{22}\mathrm{Ne} + \mathrm{p} + \mathrm{n} \end{aligned} \qquad (2.2)$$

or just one, as in the capture reaction $\mathrm{p} + {}^{27}\mathrm{Al} \rightarrow {}^{28}\mathrm{Si}^{*}$, where the asterisk indicates an excited state that usually decays, emitting γ-radiation. Under special circumstances, more than two reactants are possible. Thus, for example, the reaction $\alpha + \alpha + \alpha \rightarrow {}^{12}\mathrm{C}$ can take place in the overheated plasma of stellar interiors.

The initial and final products can also be identical. This case characterizes a process that can be elastic, as in $\mathrm{p} + {}^{16}\mathrm{O} \rightarrow \mathrm{p} + {}^{16}\mathrm{O}$ where there is only transfer of kinetic energy between the projectile and the target, or inelastic, as in $\mathrm{n} + {}^{16}\mathrm{O} \rightarrow \mathrm{n} + {}^{16}\mathrm{O}^{*}$, where part of the kinetic energy of the system is used in the excitation of ^{16}O.

Naturally, nuclear reactions are not limited to nuclei. They could involve any type of particle and also radiation. Thus, the reactions

$$\begin{aligned} \gamma + {}^{63}\mathrm{Cu} &\rightarrow {}^{62}\mathrm{Ni} + \mathrm{p}, \\ \gamma + {}^{233}\mathrm{U} &\rightarrow {}^{90}\mathrm{Rb} + {}^{141}\mathrm{Cs} + 2\mathrm{n} \end{aligned} \qquad (2.3)$$

are examples of nuclear processes induced by γ-radiation. In the first case a γ-ray knocks a proton off ^{63}Cu and in the second it induces nuclear fission of ^{233}U, with the production of two fragments and emission of two neutrons.

Unlike a chemical reaction, the resulting products of a nuclear reaction are not

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

determined univocally: starting from two or more reactants there can exist dozens of final products with an unlimited number of available quantum states. As an example, the collision of a deuteron with ^{238}U can result in, among others, the following reactions:

$$\begin{aligned} \mathrm{d} + {}^{238}\mathrm{U} &\rightarrow {}^{240}\mathrm{Np} + \gamma, \\ &\rightarrow {}^{239}\mathrm{Np} + \mathrm{n}, \\ &\rightarrow {}^{239}\mathrm{U} + \mathrm{p}, \\ &\rightarrow {}^{237}\mathrm{U} + \mathrm{t} \end{aligned} \tag{2.4}$$

In the first one, the deuteron is absorbed by uranium, forming an excited nucleus of ^{240}Np that de-excites by emitting a γ-ray. The second and third are examples of *stripping reactions*, in which a nucleon is transferred from the projectile to the target. The last one exemplifies the inverse process: the deuteron captures a neutron from the target and emerges as ^{3}H (tritium). This is denoted as a *pickup reaction*. Another possibility would be, in the first reaction, that ^{240}Np fissions instead of emitting a γ-ray, contributing with dozens of possible final products for the reaction.

Each reaction branch, with well-defined quantum states of the participants, is known as a *channel*. In Eq. (2.4), for the *entrance channel* $\mathrm{d} + {}^{238}\mathrm{U}$, there are four possible *exit channels*. Note that a different exit channel would be reached if some of the final products were in the excited state. The probability that a nuclear reaction takes place through a certain exit channel depends on the energy of the incident particle and is measured by the *cross section* of that channel.

Nuclear reactions proceed through many possible distinct mechanisms. For instance, in *direct reactions*, the projectile and the target have an interaction of short duration, with possible exchange of energy or particles between them. Another mechanism involves the *fusion* of the projectile with the target, the available energy being distributed among all nucleons, forming a highly excited *compound nucleus*. The decay of the compound nucleus leads to the final products of the reaction.

In high-energy collisions, the nuclei fragment and particles that were not initially present are produced (for instance, pions, kaons, etc.). The reactions proceed through an intermediate phase in which the nuclear matter is compressed. At very high energies, the quarks and gluons inside nucleons become ''free'' for a short time, forming the *quark–gluon plasma*. The study of high-energy reactions with nuclei is very important for a better understanding of what happens during spectacular stellar explosions (supernovae) and in the interior of compact stars, as for instance, neutron stars. The study of nuclear reactions at high energies allows to obtain information on the *equation of state* of nuclear matter.

2.2 Basic Principles

2.2.1 Conservation Laws

Several conservation laws contribute to restrict the possible processes in a nuclear reaction.

1. *Baryonic number*. There is no experimental evidence of processes in which nucleons are created or destroyed without the creation or destruction of corresponding antinucleons. The application of this principle to low-energy reactions is still more restrictive. Below the threshold

for the production of mesons (∼140 MeV), no process related to the nuclear forces is capable of transforming a proton into a neutron and vice versa, and processes governed by the weak force (responsible for the β-emission of nuclei) are very slow in relation to the times involved in nuclear reactions ($\sim 10^{-22}$–10^{-16} s). In this way, we can speak separately of proton and neutron conservation; this should show up with same amounts in both sides of a nuclear reaction.

2. *Charge*. This is a general conservation principle in physics, valid in any circumstance. In purely nuclear reactions, it is computed making the sum of the atomic numbers, which should be identical, on both sides of the reaction.
3. *Energy and linear momentum*. These are two of the most applied principles in the study of the kinematics of reactions. Using them, angles and velocities are related to the initial parameters of the problem.
4. *Total angular momentum*. It is always a constant of motion. In the reaction

$$^{10}\mathrm{B} + {}^{4}\mathrm{He} \rightarrow {}^{1}\mathrm{H} + {}^{13}\mathrm{C} \qquad (2.5)$$

^{10}B has $I = 3$ in the ground state, whereas the α-particle has zero angular momentum. If it is captured in an s-wave ($l_i = 0$), the intermediate compound nucleus is in a state with $I_c = 3$. Both final products have intrinsic angular momenta equal to 1/2. Hence, their sum is 0 or 1. Therefore, the relative angular momentum of the final products will be $l_f = 2$, 3, or 4.

5. *Parity*. It is always conserved in reactions governed by nuclear interaction. In the previous example, ^{10}B, ^{4}He, and the proton have equal parities, while ^{13}C has odd parity. Therefore, if $l_i = 0$, we necessarily have $l_f = 3$. Thus, the orbital momentum of the final products of Eq. (2.5) is determined by the combined conservation of the total angular momentum and of the parity.
6. *Isospin*. This is an approximate conservation law that is applied to light nuclei, where the effect of the Coulomb force is small. A nuclear reaction involving these nuclei not only conserves the *z*-component of the isospin (a consequence of charge and baryonic number conservation) but also conserves the total isospin **T**. Reactions that populate excited states not conserving the value of **T** are strongly inhibited. An example is the reaction $\mathrm{d} + {}^{16}\mathrm{O} \rightarrow \alpha + {}^{14}\mathrm{N}$, where the excited state 0^+, with 2.31 MeV, of ^{14}N is about a 100 times more populated than the 1^+ ground state. Conservation of energy, angular momentum, and parity does not impose any prohibition for that channel, whose low occurrence can only be justified by isospin conservation: the ground states of the four participant nuclei in the reaction have all $\mathbf{T} = 0$ and the state 1^+ of ^{14}N has $\mathbf{T} = 1$.

2.2.2 Kinematics

We consider a typical reaction, in which the projectile a and the target A result in two products, b and B. This can also be expressed in the notation that we used so far, $\mathrm{a} + \mathrm{A} \rightarrow \mathrm{b} + \mathrm{B}$, or even in a more compact notation, A(a,b)B.

Often, a and b are light nuclei and A and B the heavy ones; the nucleus b is emitted

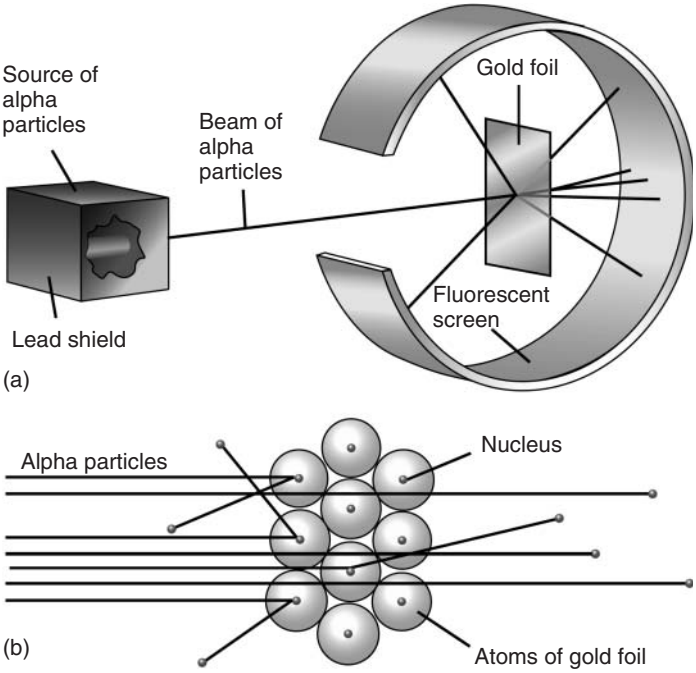


Figure 2.1 (a) Rutherford's scattering experiment with α-particles scattering on a gold foil. (b) Microscopic interpretation of the experimental results.

at an angle θ and its energy registered in the laboratory system. The recoiling nucleus B has usually a short range and cannot leave the target. It is convenient to introduce the *Q-value of the reaction*, which measures the energy gained (or lost) because of the difference between the initial and final masses:

$$Q = (m_a + m_A - m_b - m_B)c^2 \quad (2.6)$$

Using energy and momentum conservation in the reaction, one gets

$$Q = E_b\left(1 + \frac{m_b}{m_B}\right) - E_a\left(1 - \frac{m_a}{m_B}\right) - \frac{2}{m_B}\sqrt{m_a m_b E_a E_b}\cos\theta \quad (2.7)$$

From this relation, one concludes that, when Q is negative, the incident particle has an energy threshold, E_t, which is a function of the angle θ, below which nuclei b are not observed:

$$E_t = \frac{-Q m_B (m_B + m_b)}{m_a m_b \cos^2\theta + (m_B + m_b)(m_B - m_a)} \quad (2.8)$$

This example shows the power of conservation laws in the analysis of nuclear reactions. Similar to this example, very useful relations can be derived in the relativistic regime, using other conserved quantities.

2.2.3 Cross Sections

Figure 2.1 shows schematically a typical scattering experiment. In fact, this is the sketch of Rutherford's experiment

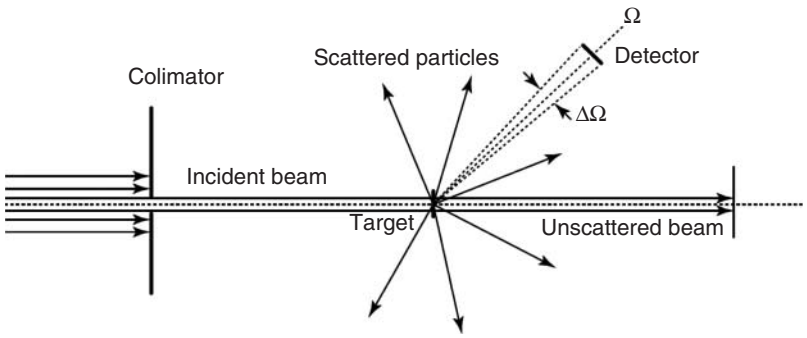


Figure 2.2 Schematic representation of a scattering experiment.

in 1910 [1]. Projectiles (here, α-particles) from a source pass through a collimator and collide with a target (gold foil). Some projectiles are scattered by the target and reach the detector (here, a fluorescent screen). Rutherford expected them to go straight through the target, with perhaps minor deflections. Most did go straight through, but, to his surprise, some particles bounced straight back! Rutherford hypothesized that the positive α- particles had hit a heavy mass of positive particles, which he called *the nucleus*.

Let us consider an experiment such as the one represented in Figure 2.2 that measures the count rate of events leading to the population of channel-α, $N_\alpha(\Omega,\ \Delta\Omega)$. Assuming that interactions between beam particles can be neglected, the count rate $N_\alpha(\Omega,\ \Delta\Omega)$ should be proportional to the incident flux J, to the fraction of solid angle $\Delta\Omega$, where the particles are scattered to, and to the number of target particles per unit volume. Hence, $N_\alpha(\Omega,\ \Delta\Omega)$ can then be written as $N_\alpha(\Omega,\ \Delta\Omega) = (\Delta\Omega \cdot n \cdot J)\ \mathrm{d}\sigma_\alpha(\Omega)/\mathrm{d}\Omega$. The constant of proportionality,

$$\frac{\mathrm{d}\sigma_\alpha(\Omega)}{\mathrm{d}\Omega} = \frac{N_\alpha(\Omega, \Delta\Omega)}{\Delta\Omega \cdot n \cdot J} \tag{2.9}$$

is called the *differential cross section* for channel-α. This quantity is very useful, as it does not depend on experimental details (detector size, incident flux, and target thickness). It depends exclusively on the physics of the projectile and the target particles.

One is often interested in the angle-integrated *total cross section* given by

$$\sigma_\alpha = \int \mathrm{d}\Omega \left[\frac{\mathrm{d}\sigma_\alpha(\Omega)}{\mathrm{d}\Omega}\right] \tag{2.10}$$

In nuclear physics, cross sections are usually measured in units of *barns*. One barn corresponds to an area of a circle of diameter approximately equal to $8\ \mathrm{fm} = 8 \times 10^{-15}\ \mathrm{m}$. For comparison between theory and experiment, it is necessary to have cross sections in the same reference frame. Of course, the measured cross section is obtained in the laboratory frame, where the target is at rest. From the theoretical point of view, however, it is important to take advantage of the translational invariance of the projectile–target Hamiltonian, introducing the center of mass (CM) frame. The transformation is done using the laws of conservation of energy and momentum.

The Rutherford experiment has a simple classical description in which one assumes that the α-particles follow hyperbolic trajectories in the Coulomb field of the heavy target. Scattering at an angle θ can be

easily calculated and with that one is able to calculate the Coulomb, or Rutherford cross section. The result is

$$\frac{d\sigma_R}{d\Omega} = \left(\frac{Z_P Z_T e^2}{4E}\right)^2 \text{cosec}^4\left(\frac{\theta}{2}\right) \tag{2.11}$$

where Z_P (Z_T) is the projectile (target) charge number and E is the CM energy.

2.2.4 Elastic Scatteringa

When a beam of particles – represented in quantum mechanics by a plane wave – hits a nucleus, the wavefunction is modified by the presence of a scattering potential $V(r)$, responsible for the appearance of a phase in the outgoing part of the wave. Elastic scattering is just one of the channels along which the reaction can proceed and is known as *elastic channel*. Inelastic scattering and all the other channels are grouped in the *reaction channel*.

The occurrence of a nuclear reaction through a given reaction channel leads to a modification of the outgoing part of the wavefunction not only by a phase factor but also by a change in its magnitude, indicating that there is a loss of particles in the elastic channel. For a projectile with momentum $p = \hbar k$, this can be expressed by

$$\Psi \sim \frac{1}{2i}\sum_{l=0}^{\infty}(2l+1)i^l P_l(\cos\theta) \times \frac{\eta_l e^{i\left(\frac{kr-l\pi}{2}\right)} - e^{-i\left(\frac{kr-l\pi}{2}\right)}}{kr} \tag{2.12}$$

where the complex coefficient η_l is the factor mentioned above. If $\eta = 1$, the sum in Eq. (2.12) can be done analytically, leading to $\Psi \sim \exp(i\mathbf{k}.r)$, that is, a plane wave. But if $\eta_l = \exp[i\delta l]$, with δ_l real, the incoming and outgoing waves have the same magnitude, that is, the scattering is elastic.

The sum in Eq. (2.12) is known as the *partial wave expansion* of the scattering wave. The label $l = 0, 1, 2, \ldots =$ (s, p, d, ... waves) denotes the contribution of a particular angular momentum (in units of $\hbar$) to the total wavefunction. Classically, the angular momentum of an incident particle is given by $l = kR$, where R is known as the *impact parameter*, which is the perpendicular distance to the target if the projectile were undeflected. In quantum mechanics l is not continuous, varying in steps of one.

We can rewrite Eq. (2.12) as a sum of a plane wave and a scattering outgoing wave in the form $\Psi \sim \exp(i\mathbf{k}.r) + f(\theta)e^{ikr}/r$, where $f(\theta)$ accounts for the distortion of the outgoing wave at the scattering angle θ. It is known as the *scattering amplitude*:

$$f(\theta) = \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)(\eta_l - 1)P_l(\cos\theta) \tag{2.13}$$

One can now calculate the cross section by counting the number of particles that are scattered through an angle θ. This can be done by calculating the particle current associated with the wavefunction Ψ. One obtains the differential scattering cross section

$$\frac{d\sigma_e}{d\Omega} = |f(\theta)|^2 = \frac{1}{4k^2}\left|\sum_{l=0}^{\infty}(2l+1) \times (1-\eta_l)P_l(\cos\theta)\right|^2 \tag{2.14}$$

The total scattering cross section, Eq. (2.10), becomes

$$\sigma_e = \pi\lambda\!\!\!^{-2}\sum_{l=0}^{\infty}(2l+1)|1-\eta_l|^2 \tag{2.15}$$

with $\lambda\!\!\!^{-} = \lambda/2\pi = 1/k$.

2.2.5 Reaction Cross Sections

To calculate the reaction cross section, it is necessary to compute the number of particles that disappear from the elastic channel, which is measured by the flux of the probability vector current through a spherical surface of large radius centered at the target, calculated with the total wavefunction of Eq. (2.12). One finds

$$\sigma_r = \pi \lambda\!\!\!^{-2} \sum_{l=0}^{\infty} (2l+1)(1-|\eta_l|^2) \qquad (2.16)$$

From Eqs. (2.15) and (2.16), when $|\eta_l| = 1$, the reaction cross section is zero and we have pure scattering. The contrary, however, cannot happen, as the vanishing of σ_e also implies the vanishing of σ_r. In general, there is a region of allowed values of η_l for which the two cross sections can coexist.

The maximum of σ_r happens for $\eta_l = 0$, which corresponds to total absorption. Let us suppose that the absorption potential is limited to the surface of a nucleus with radius $R \gg \lambda\!\!\!^{-}$, that is, that all the particles with impact parameter smaller than the radius R are absorbed. This is equivalent to saying that all particles are absorbed for $l \leq R/\lambda\!\!\!^{-}$. In this case,

$$\sigma_r = \pi \lambda\!\!\!^{-2} \sum_{l=0}^{R/\lambda\!\!\!^{-}} (2l+1) = \pi (R+\lambda\!\!\!^{-})^2 \qquad (2.17)$$

This is the value that would be intuitively adequate for the total cross section, that is, equal to the geometric cross section (the part $\lambda\!\!\!^{-}$ can be understood as an uncertainty in the position of the incident particle). However, we saw above that the presence of scattering is always obligatory. For $\eta_l = 0$, the scattering and reaction cross sections are identical, yielding the total cross section

$$\begin{aligned} \sigma &= \sigma_r + \sigma_e \\ &= \pi (R+\lambda\!\!\!^{-})^2 + \pi (R+\lambda\!\!\!^{-})^2 \\ &= 2\pi (R+\lambda\!\!\!^{-})^2 \end{aligned} \qquad (2.18)$$

that is twice the geometric cross section!

The presence of the scattering part, which gives an apparently unexpected result in Eq. (2.18), can be interpreted as the effect of diffraction of the plane waves at the nuclear surface. This effect leads to a "shadow" behind the nucleus, decreasing its apparent diameter so that, at a certain distance, the perturbation caused by the presence of the nucleus disappears and the plane wave is reconstructed. In this situation, we can say that the part of the beam that is diffracted has to be the same as the part that is absorbed, justifying the equality of σ_r and σ_e. The diffraction phenomenon appears clearly in the elastic scattering or inelastic angular distribution (differential cross section as a function of the scattering angle): Figure 2.3 exhibits angular distributions for the elastic scattering of 30 MeV protons on ^{40}Ca,^{120}Sn, and ^{208}Pb. The oscillations in the cross sections are characteristic of a *Fraunhofer diffraction*, similar to light scattering by an opaque disk. The angular distance $\Delta\theta$ between the diffraction minima closely follows the expression $\Delta\theta = \hbar/pR$, characteristic of diffraction phenomena.

2.2.6 Excitation Functions

When the projectile has a very low energy, $k \to 0$, and in particular, $l = kR \ll 1$. As an example consider the scattering of neutrons with $l = 0$ and ignore the spins of the neutron and of the target. In this case, the *Schrödinger equation* for the radial

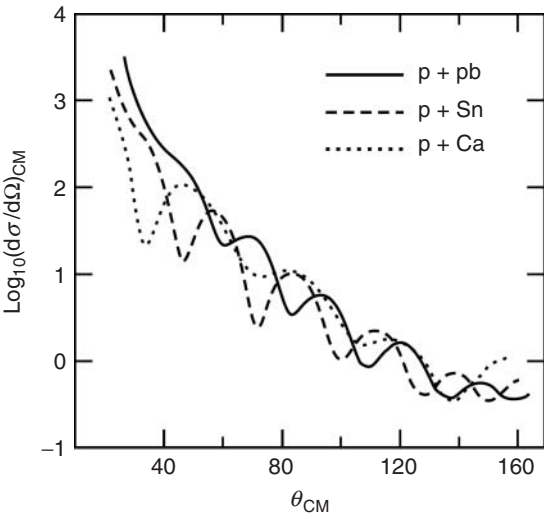


Figure 2.3 Elastic scattering angular distribution of protons of 30 MeV on Ca, Sn, and Pb. The curves are adjusted to the experimental data of Ref. 2

motion of the neutron with respect to the target is

$$\frac{d^2u_0}{dr^2} + k^2u_0 = 0 \quad (r \geq R) \tag{2.19}$$

This is valid for the radial wavefunction u_0 at distances r larger than the *channel radius* $R = R_a + R_A$, with R_a and R_A being the radii of the projectile and the target, respectively. The solution of Eq. (2.19) is

$$u_0 = \eta_0 e^{ikr} - e^{-ikr} \quad (r \geq R) \tag{2.20}$$

A radial wavefunction inside the nucleus should connect to the external function (2.20) with a continuous function and its derivative at $r = R$. That is, the function

$$f_l \equiv R\left[\frac{du_l/dr}{u_l}\right]_{r=R} \tag{2.21}$$

must have identical values if calculated with the internal or the external function and this condition creates a relationship between f_l and η_l. Hence, the knowledge of f_l leads to the knowledge of the cross sections. For neutrons with $l = 0$, the application of Eq. (2.20) results in

$$f_0 = ikR\,\frac{\eta_0 + e^{-2ikR}}{\eta_0 - e^{-2ikR}} \tag{2.22}$$

from which we extract

$$\eta_0 = \frac{f_0 + ikR}{f_0 - ikR}e^{-2ikR} \tag{2.23}$$

If f_0 is a real number, then $|\eta_0|^2 = 1$. The reaction cross section, Eq. (2.16), will be zero and we have pure scattering.

Using Eq. (2.23), the scattering cross section, Eq. (2.15), can be written as

$$\sigma_{e,0} = \pi\lambda\!\!\!^{-2}|A_{res} + A_{pot}|^2 \tag{2.24}$$

with $A_{res} = -2ikR/(f_0 - ikR)$ and $A_{pot} = \exp(2ikR) - 1$. The separation of the cross section in two parts has physical justification: A_{pot} does not contain f_0 and therefore does not depend on conditions inside the nucleus. It represents the situation where the projectile does not penetrate the nucleus as in the idealized situation where the nucleus is considered an impenetrable hard sphere. The wavefunction is

zero inside the nucleus and u_0 vanishes at $r = R$, implying $f_0 \to \infty$ and $A_{res} \to 0$. Hence, A_{pot} is only responsible for scattering.

Substituting Eq. (2.23) in Eq. (2.16) and using

$$f_0 = f_R + \mathrm{i}f_I \tag{2.25}$$

we have

$$\sigma_{r,0} = \pi \lambda\!\!\!^{-2} \frac{-4kRf_I}{f_R^2 + (f_I - kR)^2} \tag{2.26}$$

an equation that is useful when we study the presence of resonances in the *excitation function* (cross section as a function of the energy).

2.3 Statistical Reactions

2.3.1 Compound Nucleus

When a low-energy neutron (<50 MeV) enters the range of nuclear forces, it can be scattered or begin a series of collisions with the nucleons. The products of these collisions, including the incident particle, will continue in their course, leading to new collisions and new changes of energy. During this process, one or more particles can be emitted and they form, with the residual nucleus, the products of what is known as a *pre-equilibrium* reaction. At low energies, the largest probability is the continuation of the process so that the initial energy is distributed among all nucleons, with no emitted particle. The final nucleus with $A + 1$ nucleons has an excitation energy equal to the kinetic energy of the incident neutron plus the binding energy the neutron has in the new, highly unstable, nucleus [3]. It can, among other processes, emit a neutron with the same or smaller energy as that of the one absorbed. The de-excitation process is not necessarily immediate and the excited nucleus can live for a relatively long time. We say that there is, in this situation, the formation of a *compound nucleus* as intermediary stage of the reaction. In the final stage, the compound nucleus can evaporate one or more particles, fission, and so on. In our notation, for the most common situation in which two final products are formed (the evaporated particle plus the residual nucleus or two fission fragments, etc.) we write:

$$\mathrm{a} + \mathrm{A} \to \mathrm{C}^* \to \mathrm{B} + \mathrm{b}$$

the asterisk indicating that the compound nucleus C is in an excited state.

The compound nucleus lives long enough to ''forget'' how it was formed and the de-excitation to the final products b and B only depends on the energy, angular momentum, and parity of the quantum state of the compound nucleus. An interesting experimental verification was accomplished by Ghoshal in 1950 [4]. He studied two reactions that take to the same compound nucleus, $^{64}Zn^*$, and measured the cross sections of three different forms of decay, as shown below:

$$\begin{array}{lcl} p + {}^{63}Cu \searrow & & \nearrow \; {}^{63}Zn + n \\ & {}^{64}Zn^* \to & \nearrow \; {}^{62}Cu + n + p \\ a + {}^{60}Ni \nearrow & & \searrow \; {}^{62}Zn + 2n \end{array} \tag{2.27}$$

If the idea of the compound nucleus is valid and if one chooses the energy of the proton and of the incident α-particle to produce the same excitation energy, then the cross section for each one of the three exit channels should be independent of

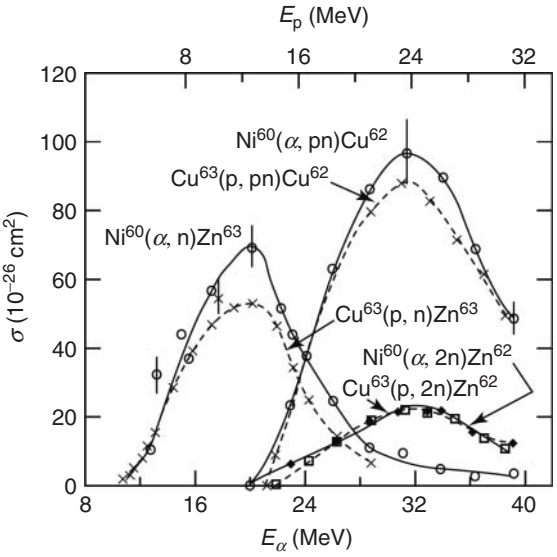


Figure 2.4 Cross sections for the reactions shown in Eq. (2.24). The scales of the upper axis (energy of the protons) and lower axis (energy of the α-particle) were adjusted to correspond to the same excitation energy of the compound nucleus [4].

the way the compound nucleus is formed. That is, the properties of the compound nucleus do not have any relationship with the nuclei that formed it. This is confirmed in Figure 2.4, where one sees clearly that the cross sections depend practically only on the exit channels.

The angular distribution of fragments, or evaporated particles, of a compound nucleus should be isotropic in the CM, and this is verified experimentally. However, the total angular momentum is conserved and cannot be "forgotten." Reactions with large transfer of angular momentum, as when heavy ions are used as projectiles, can show a nonisotropic angular distribution in the CM system.

The occurrence of a nuclear reaction in two stages allows the cross section for a reaction A(a,b)B to be written as the product, $\sigma(\mathrm{a,b}) = \sigma_c(\mathrm{a,A})P(\mathrm{b})$, where σ_c(a,A) is the cross section of formation of the compound nucleus starting from the projectile a and the target A and P(b) is the probability that the compound nucleus emits a particle b leaving a residual nucleus B. If the quantum numbers of entrance and exit channels are well specified, that is, if the reaction begins at an entrance channel α and ends at an exit channel β, one can write

$$\sigma(\alpha, \beta) = \sigma_c(\alpha)P(\beta) \tag{2.28}$$

We can associate the probability $P(\beta)$ to the width Γ_β of the channel β and write

$$P(\beta) = \frac{\Gamma(\beta)}{\Gamma} \tag{2.29}$$

where Γ is the total width, that is, $\tau = \hbar/\Gamma$ is the half-life of disintegration of the compound nucleus. Equation (2.29) just expresses the fact that the decay probability through channel β is the decay rate through that channel divided by the total decay rate. In the competition between the several channels β, the nucleons have clear preference over the γ-radiation whenever there is available energy for their emission and, among the nucleons, the neutrons have preference as they do not have the

Coulomb barrier as an obstacle. Thus, in a reaction where there is no restriction for neutron emission, we can say that

$$\Gamma \cong \Gamma_n \tag{2.30}$$

where Γ_n includes the width for the emission of one or more neutrons.

The study of the function $P(\beta)$ is done in an evaporation model that leads to results in many aspects similar to the evaporation of molecules of a liquid, with the energy of the emitted neutrons having the form of a Maxwell–Boltzmann distribution (MB)

$$I(E) \propto E \exp\left(-\frac{E}{\theta}\right) \mathrm{d}E \tag{2.31}$$

with I measuring the amount of neutrons emitted with energy between E and $E + \mathrm{d}E$. The quantity θ, with dimension of energy, has the role of a *nuclear temperature*. It is related to the density of levels ω of the daughter nucleus B by

$$\frac{1}{\theta} = \frac{\mathrm{d}S}{\mathrm{d}E} \tag{2.32}$$

with

$$S = \ln\ \omega(E) \tag{2.33}$$

where $\mathrm{d}S/\mathrm{d}E$ is calculated for the daughter nucleus B at the maximum excitation energy that it can have after the emission of a neutron, that is, in the limit of emission of a neutron with zero kinetic energy.

The level density $\omega(E)$ is a measure of the number of available energy states for the decay of the compound nucleus in the interval $\mathrm{d}E$ around the energy E. In this sense, the relationship (2.33) is, neglecting the absence of the Boltzmann constant, identical to the thermodynamic relationship between the entropy S and the number of states available for the transformation of a system. Equation (2.32) is the well-known relation between the entropy and the temperature.

A simple expression for the energy dependence of the state density is provided by the equidistant spacing model, which assumes that the one-particle states are equally spaced with spacing d and that the total energy of the nucleus is simply obtained by adding the energies of the constituent nucleons. The solution of this problem can be obtained from statistical mechanics [6]:

$$\rho(E) \sim \exp\ (2\sqrt{aE})$$

Extensive analyzes of experimental data show that for nuclei far from the region of the magic nuclei, a varies linearly with A (or with N and Z), as shown in Figure 2.5:

$$a \cong \frac{A}{k}\ \mathrm{MeV}^{-1} \tag{2.34}$$

It is found that $k \cong 7.5 - 8$.

2.3.2 Energy Spectrum of Neutrons

The energy distribution of neutrons emitted by a compound nucleus has the aspect of the curve shown in Figure 2.6. Only the low-energy part obeys Eq. (2.31) and the reason is simple: the emission of a low-energy neutron leaves the residual nucleus with a large excitation energy, and the level density is very high. The large density of final states makes the problem tractable with the statistical model that leads to Eq. (2.31); in the opposite situation, it is the low-energy states of the residual nucleus. These isolated states appear as peaks in the tail of the distribution. When the emission is of a proton or of another charged particle, the form of Figure 2.6 is distorted, the part of low energy of the spectrum

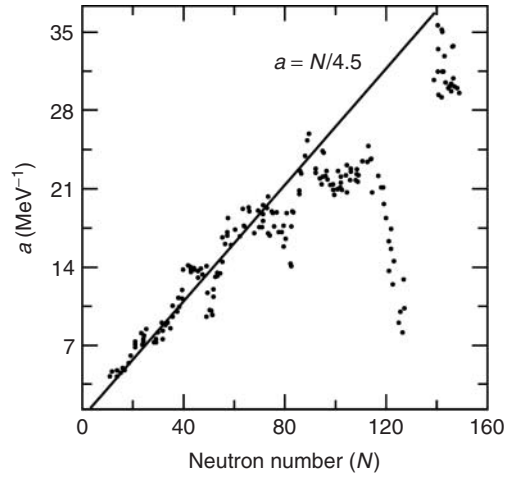


Figure 2.5 Values of the level density parameter a as a function of the neutron number. Far from magic regions, $a \approx N/4.5\ \mathrm{MeV}^{-1}$, which approximately corresponds to $a = A/7.5\ \mathrm{MeV}^{-1}$ [5].

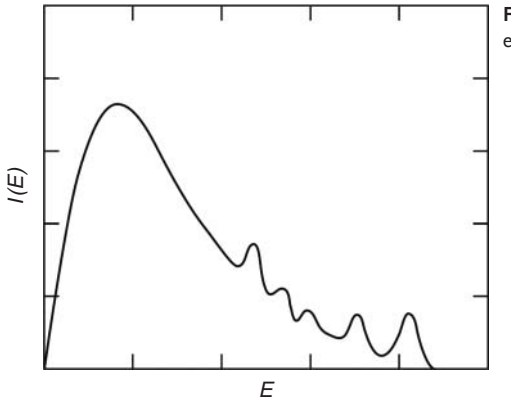


Figure 2.6 Energy spectrum of neutrons evaporated by a compound nucleus.

being suppressed partially by the Coulomb barrier.

Let us now assume that there is only elastic scattering or formation of a compound nucleus. We further assume that the elastic scattering is purely potential, without resonant elastic scattering. That is, there is no reemission of neutrons with energy equal to the incident projectile. This is equivalent to saying that the probability that the exit channel is the same as the entrance channel is very low. According to these assumptions, the cross section $\sigma_c(\alpha)$ of Eq. (2.28) is the reaction cross section of Eq. (2.16). We can still write the wavefunction inside the nucleus as just an incoming wave

$$u_0 \cong \exp(-\mathrm{i}Kr) \quad (r < R) \tag{2.35}$$

where $K = \sqrt{2m(E - V_0)}/\hbar$ is the wavenumber inside the nucleus, and it is assumed that the neutron with total energy E is subject to a negative potential V_0. Equation (2.35) is clearly a crude simplification when the incident neutron interacts in a complicated way with the other nucleons in the nucleus. It allows,

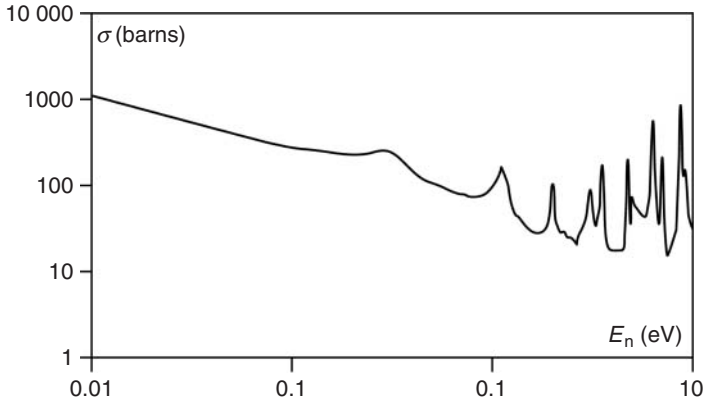


Figure 2.7 Total cross section for low-energy neutrons hitting ^{235}U.

however, to explain the average behavior of the cross sections for low energies. Starting with Eqs. (2.21) and (2.35), one determines the value of f_0:

$$f_0 = -\mathrm{i}KR \tag{2.36}$$

and from Eq. (2.26), we get

$$\sigma_c = \pi \lambda\!\!\!^{-2} \frac{4kK}{(k+K)^2} \tag{2.37}$$

for the cross section of compound nucleus formation for neutrons with $l = 0$. At low energies, $E|V_0|$, thus kK. Under these conditions, $\sigma_c = 4\pi/kK$, since $\lambda\!\!\!^{-} = k^{-1} \gg R$. Thus, σ_c varies with $1/k$. That is,

$$\sigma_c \propto \frac{1}{v} \tag{2.38}$$

where v is the velocity of the incident neutron. This is the well-known $1/v$ *law* that governs the behavior of the capture cross section of low-energy neutrons. Figure 2.7 exhibits the excitation function (cross section as function of the energy) for the reaction n+^{235}U. The cross section decays with $1/v$ up to 0.3 eV, where a series of resonances appear.

2.3.3 Resonances

To understand why there are resonances, we shall use again the simple model of a single particle subject to a square-well potential. We know that inside the well, the Schrödinger equation only admits solution for a discrete group of values of energy, $E_1, E_2, \ldots, E_n$. A particle is confined to the interior of the well by reflections that it has at the surface of the well. In these reflections, the wave that represents the particle should be in phase before and after the reflection and this only happens for a finite group of energies. Outside the well, the Schrödinger equation does not impose restrictions and the energy can have any value. But we know, from the study of the passage of a beam of particles through a potential step, that the discontinuity of the potential at the step provokes reflection even when the total energy of the particles is larger than the step, a situation where classically there would not be any difficulty for the passage of the particles. This reflection is partial and it becomes larger the closer the energy is to the height of the potential step. We

can say that a particle with energy slightly positive is almost as confined as a particle inside the well. From this fact results the existence of almost bound states of positive energy known as *quasi-stationary states* or *resonances*. These resonances appear as peaks in the excitation function, a peak at a given energy meaning that the energy coincides with a given resonance of the nucleus.

The existence of resonances can also be inferred from the properties of the wavefunction. We consider only elastic scattering, with the other channels closed. The external and internal wavefunctions are both sine functions, the first with wavenumber $k = \sqrt{2mE}/\hbar$ and the second with $K = \sqrt{2m(E - V_0)}/\hbar$. If E is small and V_0 is about -35MeV, we have $K \gg k$. The internal and external parts should join at $r = R$ with continuous function and derivatives. As the internal frequency is much larger than the external one, the internal amplitude is quite reduced. There is a perfect matching between both and the internal amplitude is identical to the external one only at the proximity of the situation in which the derivative is zero. The energy for which this happens is exactly the energy of resonance.

Resonances appear in the excitation function at relatively low energies, where the number of open channels is not very large and it is possible to return to the entrance channel. To arrive at an expression of the cross section that describes a resonance, we rewrite Eq. (2.35) as,

$$u_l \sim \exp(-\mathrm{i}Kr) + b\exp(\mathrm{i}Kr)\ , \quad (r < R) \tag{2.39}$$

this time containing a second part, which takes into consideration the part of the wave that returns. This second part allows the existence of resonant scattering, where the incident particle is re-emitted with the same energy that it entered, after forming the compound nucleus. The complex amplitude b is always smaller than one, because there are no creation of particles in the region $< R$ in Eq. (2.26).

The second parenthesis in the denominator of Eq. (2.26) is never zero, since the numerator forces f_l to be always negative. If for f_R vanishes a certain energy, $\sigma_{r,0}$ passes by a maximum in that energy. We can tentatively identify these energies as being the energy of the resonances. Let us take the extreme case of a single resonance at the energy E_R, that is, $f_R = 0$ for $E = E_R$. We can expand f_R in a Taylor series in the neighborhood of a resonance, $f_R(E) = (E - E_R)(\mathrm{d}f_R/\mathrm{d}E)_{E=E_R} + \cdots$. Keeping just the first term of the expansion and using Eqs. (2.24) and (2.26), we get

$$\sigma_{\epsilon,0} = \pi\lambda\!\!\!^{-2} \left|\exp(2\mathrm{i}kR) - 1 + \frac{\mathrm{i}\Gamma_\alpha}{(E - E_R) + \mathrm{i}\frac{\Gamma}{2}}\right|^2 \tag{2.40}$$

$$\sigma_{r,0} = \pi\lambda\!\!\!^{-2}\frac{\Gamma_\alpha(\Gamma - \Gamma_\alpha)}{(E - E_R)^2 + \left(\frac{\Gamma}{2}\right)^2} \tag{2.41}$$

where we define

$$\Gamma_\alpha = -\frac{2kR}{(\mathrm{d}f_R/\mathrm{d}E)_{E=E_R}} \quad \text{and}$$
$$\Gamma = \frac{2kR - 2f_I}{(\mathrm{d}f_R/\mathrm{d}E)_{E=E_R}} \tag{2.42}$$

The energy Γ that appears in Eq. (2.40) is the total width of the resonance, $\Gamma = \Gamma_\alpha + \Gamma_\beta + \cdots$, that is, the sum of the widths for all the possible processes of decay of the nucleus, starting from the resonant state. Γ_α is the entrance channel width, and $\Gamma - \Gamma_\alpha$ is the sum of the widths of all the exit channels except α. If we

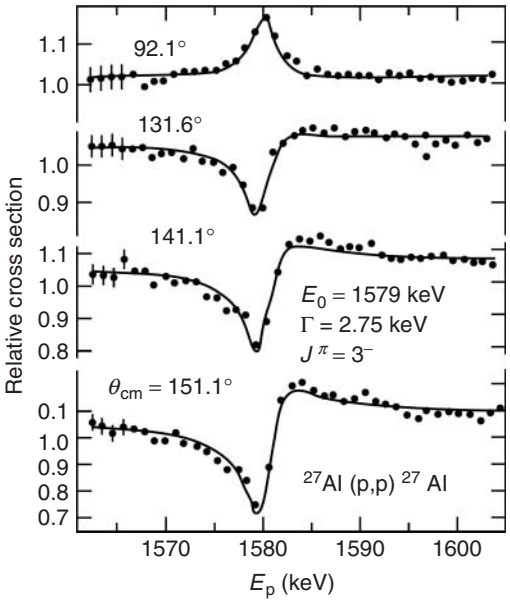


Figure 2.8 Differential cross section at four angles for the elastic scattering of protons off ^{27}Al, in units of the Rutherford cross section, in the neighborhood of the 1579 keV resonance. For larger angles, we have typical interference between the resonant scattering and the potential scattering [7].

restrict the exit channels to a single channel β, or we denote β as the group of exit channels except α, Eq. (2.41) is rewritten as

$$\sigma_{\alpha,\beta} = \pi \lambda\!\!\!^{-2} \frac{\Gamma_\alpha \Gamma_\beta}{(E - E_R)^2 + \left(\frac{\Gamma}{2}\right)^2} \tag{2.43}$$

which is the usual way of presenting the *Breit–Wigner formula*, which describes the form of the cross section close to a resonance. Let us recall that Eq. (2.43) refers to an incident particle of $l = 0$, without charge and without spin. If the spins of the incident and target particles are s_a and s_A, respectively, and the incident beam is described by a single partial wave $l \neq 0$, one can show that the cross section of Eq. (2.43) should be multiplied by the statistical factor $g = (2I + 1)/(2s_a + 1)(2s_A + 1)$, where I is the quantum number of the total angular momentum $I = \mathbf{s}_a + \mathbf{s}_A + \mathbf{l}$ of the compound nucleus. g reduces, naturally, to the unit in the case of zero intrinsic and orbital angular momenta.

If the exit channel is the same as the entrance channel α, the cross section should be obtained from Eq. (2.40) and its dependence in energy is more complicated because in addition to the resonant scattering there is the potential scattering, and the cross section from Eq. (2.40) will contain an interference term between both. The presence of interference results in a peculiar aspect of the scattering cross section, which differs from the simple form in Eq. (2.43) for the reaction cross section. This is seen in Figure 2.8 that shows the forms that a resonance can take in the scattering cross section.

The region of energy where resonances show up can extend to 10 MeV in light nuclei but it ends well before this in heavy nuclei. Starting from this limit, the increase in level density with energy implies that the average distance between levels is smaller than the width of the levels and individual resonances cannot be resolved experimentally. They form a continuum

and this region is known as the *continuum region*. The cross section in this region fluctuates.

2.3.4 Statistical Theory of Nuclear Reactions

The fluctuating behavior of low-energy nuclear reactions is due to the interference of the reaction amplitudes corresponding to the excitation of each of the overlapping states, which vanish in the energy average of the cross section since these amplitudes are complex functions with random modulus and phase. Calling $\sigma_{\mathrm{CN}}(c)$ the cross section for the formation of a compound nucleus in the entrance channel c, and using the *reciprocity theorem*, which relates the cross section $\sigma_{cc'}$ to the cross section for the time-reversed process $c' \to c$, one gets

$$\sigma_{cc'}(E_{c'})\mathrm{d}E_{c'} = \sigma_{\mathrm{CN}}(c) \times \frac{(2I_{c'}+1)\mu_{c'}E_{c'}\,\sigma_{\mathrm{CN}}(c')\omega(U_{c'})\mathrm{d}U_{c'}}{\sum_c \int_0^{E_c^{\max}} (2I_c+1)\mu_c E_c\,\sigma_{\mathrm{CN}}(c)\omega(U_c)\mathrm{d}U_c} \tag{2.44}$$

where I_c is the angular momentum and μ_c is the reduced mass in channel c. Ejectiles with energy in the range $E_{c'}$ to $E_{c'} + \mathrm{d}E_{c'}$ leave the residual nucleus with energy in the range $U_{c'}$ to $U_{c'} + \mathrm{d}U_{c'}$, where $U_{c'} = E_{\mathrm{CN}} - B_{c'} - E_{c'}$ and E_{CN} and $B_{c'}$ are, respectively, the compound nucleus energy and the binding energy of the ejectile in the compound nucleus. Equation (2.44) is the *Weisskopf–Ewing formula* for the angle-integrated cross sections [9]. To a good approximation, the level density $\omega(U) \propto \exp(U/T)$, so the ejectile spectrum given by the Weisskopf–Ewing theory is Maxwellian. It rises rapidly above the threshold energy, attains a maximum, and then falls exponentially.

The Weisskopf–Ewing theory is simple to use, but it has the disadvantage that it does not explicitly consider the conservation of angular momentum and does not give the angular distribution of the emitted particles. This is provided by the *Hauser–Feshbach theory* [10]. This theory takes into account the formation of the compound nucleus in states of different J and parity π. Let us consider the case of a reaction leading from the initial channel c to a final channel c'. If there is no *pre-equilibrium emission*, one may identify the compound nucleus formation cross section $\sigma_{\mathrm{CN}} = \sum_{J,\pi} \sigma_{\mathrm{CN}}^{J\pi}$ with the optical model reaction cross section $\sigma_R = \pi/k^2 \sum_l (2l+1)T_l$, which, if the transmission coefficients $T_l = 1 - |S_l|^2$, do not depend on J. S_l is the average value of the scattering amplitude over several overlapping resonances.

Using the above assumptions, the reciprocity theorem, and following a similar derivation as with the Ewing–Weisskopf theory, the cross section for transition $c \to c'$ is given by the *Hauser–Feshbach* expression

$$\sigma_{cc'} = \frac{\pi}{k^2} \sum_J \frac{(2J+1)}{(2i_c+1)(2I_c+1)} \times \frac{\sum_{s,l} T_l(c) \sum_{s',l'} T_{l'}(c')}{\sum_c \sum_{s,l} T_l(c)} \tag{2.45}$$

The compound nucleus states may be both of positive and negative parity. Since parity is conserved, in evaluating the expression (2.45), one must take into account that the parity of compound nucleus states and the parity of the residual nucleus states may impose restrictions on the values of the emitted particle angular momentum. Thus, positive parity compound nucleus states decay to positive parity states of the residual only by even angular momenta and to negative parity residual nucleus states by odd angular momenta.

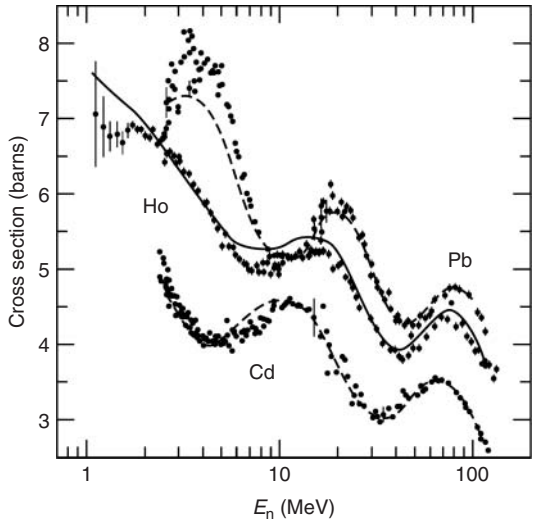


Figure 2.9 Total cross section for neutrons incident on cadmium, holmium, and lead, showing an oscillatory behavior. The curves for cadmium and lead are results of calculations with the optical model [8].

2.3.5 The Optical Model

Expression (2.18) shows that if a particle is always absorbed when it reaches the nucleus, the total cross section falls monotonically with the energy and grows linearly with $A^{\frac{1}{3}}$. In Figure 2.9, we see curves of total cross section of neutron scattering showing an oscillatory behavior for its energy dependence as well as for the mass dependence of the target. Their presence is mainly due to interference phenomena between the part of the incident beam that passes through the nucleus and the part that passes around it [11].

The basis of the *optical model* was established by Feshbach and collaborators in 1953 [12]. In this model, the interaction between the nuclei in a reaction is described by a potential $U(r)$, with r being the distance between the CM of the two nuclei. One replaces the complicated interaction that a nucleon has with the rest of the nucleus with a potential that acts on the nucleon. The potential $U(r)$ includes a complex part that takes into account the absorption effects, that is, the inelastic scattering. The nuclear scattering is treated in similar form as the scattering of light by a glass sphere and the name of the model is derived from this analogy.

In its most commonly used form, the optical potential is written as the sum

$$U(r) = U_R(r) + U_I(r) + U_D(r) + U_S(r) + U_C(r) \qquad (2.46)$$

which contains parameters that can vary with energy and masses of the nuclei and that should be chosen by an adjustment to the experimental data. Obviously, the optical potential $U(r)$ will only make sense if these variations are small for neighboring masses or energies.

The first part of Eq. (2.46), $U_R(r) = -Vf(r, R, a)$, is real and represents a nuclear well with depth $-V$, being multiplied by a Woods–Saxon form factor $f(r, R, a) = [1 + \exp\{(r - R)/a\}]^{-1}$, where R is the radius of the nucleus and a measures the diffuseness of the potential, that is, the width of the region where the function f is sensibly different from 0 or

1. V, R, and a are treated as adjustable parameters.

The absorption effect or, alternatively, the disappearance of particles from the elastic channel is taken into account, including the two following imaginary parts $U_I(r) = -\mathrm{i}Wf(r, R_I, a_I)$ and $U_D(r) = 4\mathrm{i}a_I W_D \mathrm{d}f(r, R_I, a_I)/\mathrm{d}r$. An imaginary part produces absorption. It is easy to see this for the scattering problem of the square well: if an imaginary part is added to the well, $U(r) = -V_0 - \mathrm{i}W_0 (r < r_0)$ and $U(r) = 0 (r > r_0)$, it appears in the value of $K = [2m(E + V_0 + \mathrm{i}W_0)]^{1/2}/\hbar$. This will produce an exponentially decreasing internal wavefunction. Thus, it corresponds to an absorption of particles from the incident beam.

U_I is responsible for the absorption in the whole volume of the nucleus, but U_D, built from the derivative of the function f, acts specifically in the region close to the nuclear surface. These two parts have complementary goals: at low energies, there are no available unoccupied states for nucleons inside the nucleus and the interactions are essentially at the surface. In this situation, $U_D(r)$ is important and $U_I(r)$ can be ignored. On the other hand, at high energies, the incident particle has larger penetration and, in this case, the function $U_I(r)$ is important.

As with the shell model potential, used in nuclear structure, a spin–orbit interaction term is added to the optical potential. This term, which is the fourth part of Eq. (2.46), is usually written in the form

$$U_S(r) = \mathbf{s}\cdot\mathbf{l}\; V_s\; \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} f(r, R_S, a_S) \quad (2.47)$$

$\mathbf{s}$ is the spin operator and $\mathbf{l}$ the angular orbital momentum operator. As with $U_D(r)$, the part $U_S(r)$ is only important at the surface of the nucleus since it contains the derivative of the form factor f. The values of V_S, R_S, and a_S must be adjusted by experiment.

The presence of the term U_S is necessary to describe the effect of *polarization*. Through experiences of double scattering, it can be verified that proton or neutron beams suffer strong polarization at certain angles. This means that the quantity

$$P = \frac{N_\mathrm{c} - N_\mathrm{b}}{N_\mathrm{c} + N_\mathrm{b}} \quad (2.48)$$

where N_c is the number of nucleons in the beam with spin up and N_b with spin down, has a value significantly different from zero at these angles. With the inclusion of U_S, the optical model is able to reproduce in many cases the experimental values for the polarization given in Eq. (2.48).

Finally, a term corresponding to the Coulomb potential is added to Eq. (2.46) whenever the scattering involves charged particles. It has the form

$$U_C(r) = \frac{Z_1 Z_2 e^2}{2R_\mathrm{c}}\left(3 - \frac{r^2}{R_\mathrm{c}^2}\right) \; (r \le R_\mathrm{c}) \quad (2.49)$$

$$= \frac{Z_1 Z_2 e^2}{r} \; (r > R_\mathrm{c}) \quad (2.50)$$

where it is assumed that the nucleus is a homogeneously charged sphere of radius equal to the *Coulomb barrier radius* R_c, which separates the regions of nuclear and Coulomb forces.

Figure 2.10 exhibits the result of the application of Eq. (2.46) to the elastic scattering of 17 MeV protons on several light nuclei. The angular distribution is very well reproduced by the model, which also reproduces correctly the polarization given by Eq. (2.46) for copper as a function of the scattering angle.

The optical model has a limited set of adjustable parameters and is not capable to describe abrupt variations in the

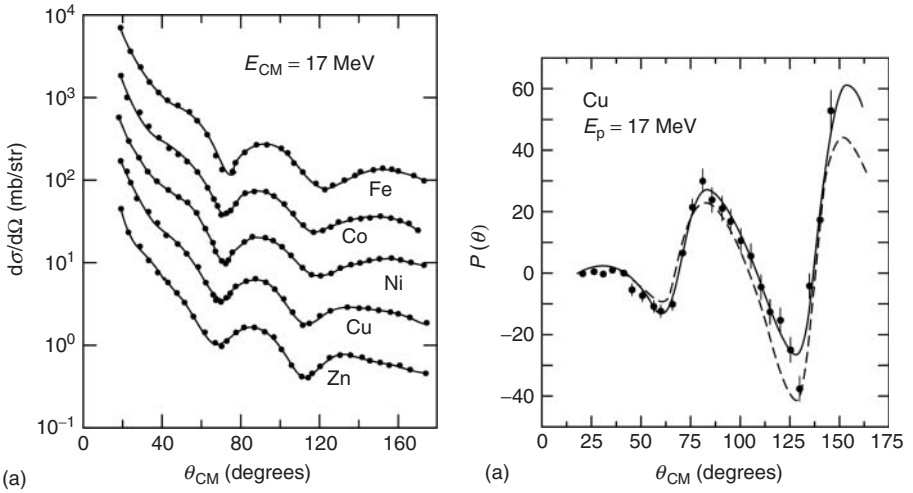


Figure 2.10 (a) Angular distribution of the elastic scattering of 17 MeV protons on nuclei in the region $Z = 26 - 30$. (b) The polarization value Eq. (2.48) in the scattering of 9.4 MeV protons on copper. The curves are, in both cases, obtained with fits from the optical model [13].

cross sections, as it happens for isolated resonances. However, it can provide a good description of the cross sections in the presence of the oscillations of large width in the continuous region, as it treats these as wave phenomena.

2.4 Direct Reactions

Direct reactions become more probable as one increases the energy of the incident particle: the wavelength associated with the particle decreases and localized areas of the nucleus can be "probed" by the projectile. In this context, peripheral reactions, where only a few nucleons of the surface participate, become important. Direct reactions happen during a time of the order of 10^{-22}\$. Reactions with formation of compound nuclei can be up to six orders of magnitude slower. A reaction type at a given energy is not necessarily exclusive; the same final products can be obtained, part of the events in a direct manner, other parts through the formation and decay of a compound nucleus.

There are two characteristic types of direct reactions. In the first, the incident particle scatters inelastically and the transferred energy is used to excite a collective mode of the nucleus. Rotational and vibrational bands can be studied in this way. The second type involves a modification in the nuclear composition. Examples are transfer of nucleons, as *pickup* and *stripping* reactions. An important reaction of the latter kind is a *knockout* reaction where the incident particle knocks out a particle of the target nucleus and continues in its path, resulting in three reaction products. Reactions with nucleon exchange can also be used to excite collective states. An example is a pickup reaction where a projectile captures a neutron from a deformed target and the product nucleus is in an excited state belonging to a rotational band.

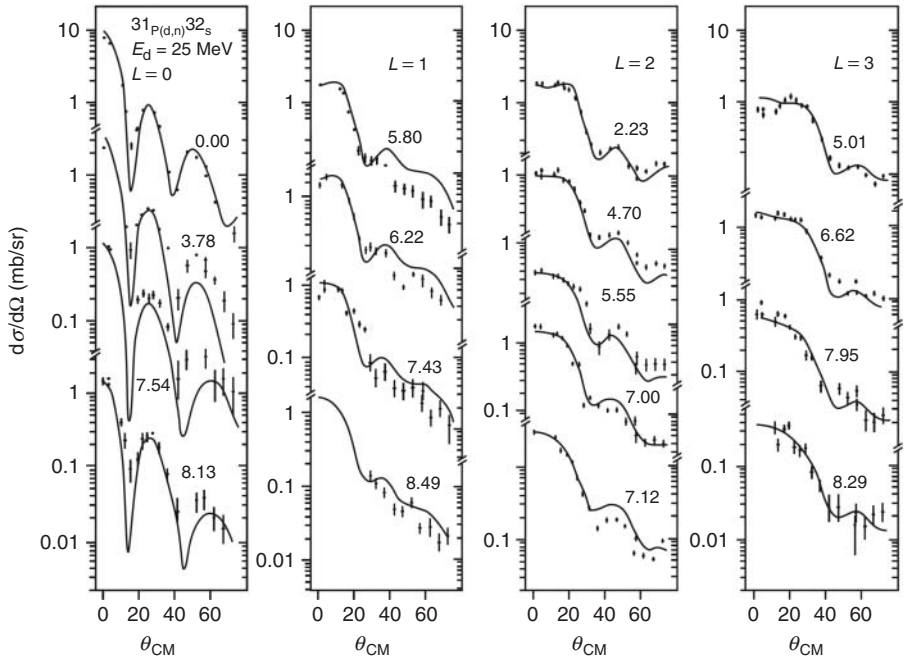


Figure 2.11 Angular distribution of the reaction $^{31}P(d, n)^{32}S$, with the transfer of a proton to several states of ^{32}S. The curves are results of DWBA calculations for the indicated l values [14].

Direct reactions exhibit a peculiar form of angular distribution, which allows us to extract information on the reaction mechanism with the employment of simple models. Typical examples are the stripping reactions (d,n) and (d,p), where the angular distribution of the remaining nucleon presents a forward prominent peak and smaller peaks at larger angles, with the characteristic aspect of a diffraction figure.

Figure 2.11 shows experimental results for the reaction $^{31}P(d, n)^{32}S$ [14]: angular distributions of the detected neutrons corresponding to each energy level of ^{32}S are exhibited for several angular momenta. We see that the behavior of the cross sections is in agreement with qualitative predictions: the curves exhibit a first peak at a value of θ that grows with l. Other smaller peaks occur as θ increases. The increase of θ_l with l is an important characteristic that can be used to identify the value of the transferred momentum in a given angular distribution.

We consider an initial quantum state with particles of mass m_a hitting a target A of mass m_A. The final quantum state are particles of mass m_b moving away from nucleus B of mass m_B. The differential cross section for the process is given by

$$\frac{d\sigma}{d\Omega} = \frac{m_a m_b k_b}{(2\pi\hbar^2)^2 k_a} |V_{fi}|^2 \tag{2.51}$$

involving the matrix element

$$V_{fi} = \int \Psi_b^* \Psi_B^* \Psi_\beta(\mathbf{r}_\beta) V \Psi_a \Psi_A \Psi_\alpha(\mathbf{r}_\alpha) d\tau \tag{2.52}$$

$\Psi_a, \Psi_b, \Psi_A,$ and Ψ_B are the internal wavefunctions of the nuclei a, b, A, and B. Ψ_α, Ψ_β are the wavefunctions of the relative motion in the entrance channel α and in the exit channel β. The integration volume $d\tau$ spans the coordinates of all particles. V is the perturbation potential that causes the "transition" from the entrance to the exit channel.

The use of plane waves for Ψ_α and Ψ_β in Eq. (2.52) is known as the *first Born approximation*. With it, we can arrive at an approximate expression for the behavior of the differential cross section. As the nuclear forces are of short range, one can restrict the integral Eq. (2.52) to regions where $\mathbf{r}_\alpha \cong \mathbf{r}_\beta = \mathbf{r}$. This leads to

$$V_{fi} \cong \int d\mathbf{r} \exp(i\mathbf{k}\cdot\mathbf{r}) \left\{ \int \Psi_b^* \Psi_B^* V \Psi_a \Psi_A \, d\tau' \right\} \tag{2.53}$$

where $\mathbf{k} = \mathbf{k}_\alpha - \mathbf{k}_\beta$. The global variables in $d\tau$ were separated into variables $d\mathbf{r}$ and $d\tau'$.

Expanding the plane wave in a Legendre polynomial series, we obtain

$$V_{fi} \cong \sum_{l=0}^{\infty} i^l (2l+1) \int j_l(kr) P_l(\cos\theta) F(\mathbf{r}) d\mathbf{r} \tag{2.54}$$

where $F(\mathbf{r}) = \int \Psi_b^* \Psi_B^* V \Psi_a \Psi_A d\tau'$ contains all the internal properties and is known as the *form factor* of the reaction. The action of V is restricted to the surface of the nucleus: outside the nucleus the action of V is limited by the short range of the nuclear forces and inside the nucleus there is a strong deviation to the absorption channel. Expression (2.54) becomes $V_{fi} \cong \sum_{l=0}^{\infty} c_l j_l(kR)$, where the coefficients c_l contain information on the form factor $F(\mathbf{r})$. The index l can be identified as the angular momentum transferred and, for a reaction that involves a single value of l, we can write for the differential cross section:

$$\frac{d\sigma}{d\Omega} \propto |j_l(kR)|^2 \tag{2.55}$$

where the dependence in θ is contained in

$$k^2 = k_\alpha^2 + k_\beta^2 - 2k_\alpha k_\beta \cos\theta.$$

We have an oscillatory behavior for the angular distribution, the maxima separated by π from each other in the axis kR.

The Born approximation with plane waves predicts for certain cases the correct place of the first peaks in the angular distribution but without reproducing correctly the intensities. A considerable progress can be made in the perturbative calculations if, instead of plane waves in Eq. (2.52), we use *distorted* waves that contain, besides the plane wave, the part dispersed elastically by the optical potential of the target. The Born approximation with distorted waves, or distorted wave Born approximation (DWBA), became a largely employed tool in the analysis of experimental results of direct reactions. With it, one can try to extract with a certain reliability the value of the angular momentum l transferred to the nucleus in a stripping or pickup reaction. An example of this is the already mentioned stripping reaction, $^{31}P(d, n)^{32}S$, for deuterons of 25 MeV. For the energy levels shown in Figure 2.11, the assignments of the value of l for the level is, in most cases, univocal.

The angular momentum l transferred in a direct reaction generally modifies the value of the total angular momentum of the nucleus. If J_i is the spin of the target nucleus, the spin J_f of the product nucleus is limited to the values

$$\left| |J_i - l| - \frac{1}{2} \right| \leq J_f \leq J_i + l + \frac{1}{2} \tag{2.56}$$

and the initial and final parities obey the relationship $\pi_i\pi_f = (-1)^l$. Equation (2.56) allows, with the knowledge of the target nucleus and of the transferred angular momentum, the determination of the parity of the product state formed and is a tool for the determination of its spin.

The knowledge of the transferred angular momentum value in a direct reaction opens the possibility to test the predictions of the shell model for the structure of nuclei. In a direct reaction, one assumes that the nucleon is located in an orbit of the nucleus with the same angular momentum as the transferred momentum in the reaction. In almost all cases studied with direct reactions, the value of the assigned l corresponds exactly to that predicted by the shell model. We know, however, that the real situation is more complicated, owing to the presence of the residual interactions that give place to configuration mixing. As a result, the cross section for the formation of a state i of the product nucleus is related to that calculated with DWBA for the formation from a single-particle state by

$$\left(\frac{d\sigma}{d\Omega}\right)_{exp} = \frac{2J_B + 1}{2J_A + 1}\mathcal{S}_{ij}\left(\frac{d\sigma}{d\Omega}\right)_{DWBA} \tag{2.57}$$

where the *spectroscopic factor* $\mathcal{S}_{ij}$ measures the weight of the configuration j used in the DWBA calculation, in the final state i, with the sum-rule

$$\sum_j \mathcal{S}_{ij} = n_i \tag{2.58}$$

The sum (2.57) embraces all the nucleons in the configurations j of the product nucleus. The statistical weight $(2J_B + 1)/(2J_A + 1)$ that appears in the DWBA calculation involving the angular momentum of the target nucleus J_A and final nucleus J_B is explicitly given in Eq. (2.58).

2.5 Heavy Ion Reactions

2.5.1 Types and Properties

Heavy ion reactions (with $A > 4$) can be separated into three major categories.

1. Owing to their large charge, two heavy nuclei feel a strong mutual Coulomb repulsion. To produce a nuclear reaction, the projectile needs enough energy to overcome the Coulomb barrier. For a very heavy target, such as ^{238}U, about 5 MeV per nucleon is necessary. Then the wavelength of the projectile is small compared with the dimensions of the nuclei and classical and semiclassical methods become useful in the description of the reaction.
2. The projectile carries a large amount of angular momentum and a good part of it can be transferred to the target in the reaction. Rotational bands with several dozens of units of angular momentum can be created. In fact, heavy ion reactions are the best suited to feed high spin levels.
3. Direct reactions and formation of compound nucleus are also common processes in reactions with heavy ions. But some peculiarities of these are not found in reactions with projectile nucleons. One of these processes can be understood as intermediate between a direct reaction and the formation of a compound nucleus. Fusion does not occur but the projectile and target pass a relatively long time under the mutual action of the nuclear forces. Nuclear matter is exchanged between both and there is a strong heating of the two nuclei, with a large transfer of kinetic

energy to the internal degrees of freedom. These are the *deep inelastic collisions*.

The kind of process that prevails depends on the distance of closest approach d between the projectile and the target. If this distance is sufficiently large only the long-range Coulomb interaction acts and, for a classical hyperbolic trajectory, d is related to the impact parameter b and to the energy E of the projectile by $d = a/2 + [(a/2)^2 + b^2]^{1/2}$, where a is the distance of closest approach in a head-on collision. This is related to E by $a = Z_1 Z_2 e^2/4\pi_0 E$.

Experimentally, the variable under control is the energy E of the projectile and, for E sufficiently large, d can be small enough to enter the range of nuclear forces. Collisions near this limit are called *grazing* collisions and are characterized by values of b_{graz} and d_{graz}. Assuming that there is always reaction when $b < b_{graz}$, the reaction cross section σ_r can be determined geometrically by $\sigma_r = \pi b_{graz}^2$. The experimental determination of σ_r allows to establish the value $d_{graz} = 0.5 + 1.36\ (A_1^{1/3} + A_2^{1/3})$ showing that the distance of grazing collision is somewhat larger than that deduced from two touching spheres ($1.36fm > r_0 = 1.2$ fm).

When the impact parameter is close to b_{graz}, one expects nuclear reactions of short duration, without the contribution of the compound nucleus formation. Such reactions are elastic and inelastic scattering and involve a transfer of few nucleons. When the incident energy is sufficiently high, small values of b can lead to the projectile penetrating the target. Depending on the energy and on the involved masses, the reaction can end in one of the processes below:

1. *Fusion* – this is the preferred process when one has light nuclei and low energy. There is formation of a highly excited compound nucleus that decays by evaporation of particles and γ-radiation emission, leading to a cold residual nucleus. If the energy in the CM is close to the Coulomb barrier energy, the cross section of compound nucleus formation starting from two nuclei is practically equal to the reaction cross section.
2. *Fission* – when the compound nucleus is heavy the fission process competes strongly with the evaporation of particles in each stage of the evaporation process. A very heavy compound nucleus with a large excitation energy has a very small probability of arriving to a cold residual nucleus without fission at some stage of the de-excitation. The role of the angular momentum l transmitted to the target nucleus is also essential. The fission barrier decreases with the increase of l and for a critical value l_{crit} the barrier ceases to exist. A nucleus with angular momentum greater than l_{crit} suffers immediate fission, and this is also a limiting factor in the production of superheavy elements.
3. *Deep inelastic collision (DIC)* – this is a phenomenon characteristic of reactions involving very heavy nuclei ($A > 40$) and with an incident energy of 1–3 MeV above the Coulomb barrier. In DIC, the projectile and the target spend some time under mutual action, exchanging masses and energy but without leading to the formation of a compound nucleus. The projectile escapes after transferring part of its energy and angular momentum to its internal degrees of freedom and to the

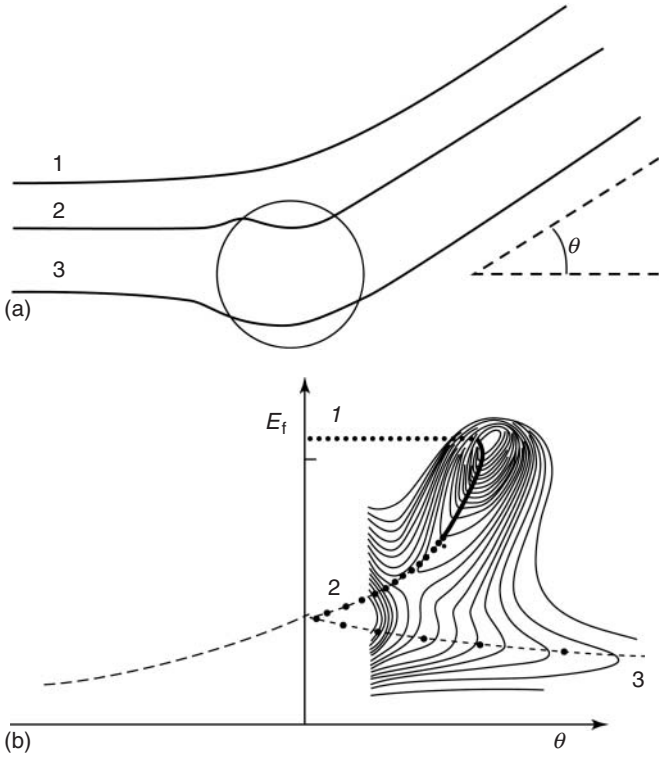


Figure 2.12 (a) Reactions with different impact parameters leading to the same scattering angle θ. (b) Lines of maxima in topographical diagrams of the final energy against the scattering angle.

target, with values reaching 100 MeV and $50\hbar$.

One of the most interesting aspects of DIC is the correlation between the energy dissipated in the collision and the scattering angle in the CM. Let us look at Figure 2.12a. Trajectory 1 shows the projectile with an impact parameter that leads it out of the range of nuclear forces. The Coulomb scattering angle will become larger as the impact parameter decreases. In a graph as in Figure 2.12b, where one plots the final energy against the scattering angle, trajectories of type 1 are located in the upper branch, where there is no dissipation and the initial kinetic energy stays unaffected. This upper branch has a maximum value for the scattering angle. At a smaller impact parameter, the nuclear attractive force and, with that, the effects of dissipation of DIC begin to act. A given Coulomb scattering angle θ can also be reached by the combination of the nuclear and Coulomb forces. Only, now there is loss of energy and the events are located in branch 2 of Figure 2.12b. There still is no one-to-one correspondence between angle and energy because an infinite number of trajectories can lead to the same angle θ. Branch 2 should be understood as a line of maxima in a three-dimensional representation (called

the *Wilczynski diagram*) where the axis perpendicular to the paper is proportional to the cross section $d^2\sigma/dEd\theta$. The same angle θ can also be obtained by trajectory 3, with a longer interaction time between the nuclei and a larger dissipation. As now the projectile is deflected toward the nucleus, the scattering angle would be formally $-\theta$ but, as there is no experimental distinction between θ and $-\theta$, these events appear as an independent ramification in the lower part of the diagram.

2.5.2 Superheavy Elements

The heaviest element found in Nature is $^{238}_{92}$U. It is radioactive, but it has survived since its formation in supernovae explosions because it has a decay half-life of the order of the age of the Earth. Elements with larger atomic number (transuranic) have shorter half-lives and have disappeared. They are created artificially through nuclear reactions using heavy elements as target. Initially, the projectiles used were light particles: protons, deuterons, α-particles and neutrons. The use of neutrons is justified because the β^--emission of the compound nucleus increases the value of Z and it was in this way [15] that the first transuranic element, neptunium, was obtained: $n + {}^{238}U \rightarrow {}^{239}U \rightarrow {}^{239}Np + \beta^-$.

Reactions with light particles can produce isotopes up to mendelevium ($Z = 101$), but it is not possible to go beyond that; the half-lives for α-emission or spontaneous fission become extremely short, making the preparation of a target impracticable. The alternative is to place a heavy element under the flux of very intense neutrons. This can be done using special reactors or using the remaining material of nuclear explosions. The elements einsteinium ($Z = 99$) and fermium ($Z = 100$) were discovered in this way in 1955 but the increasing competition of beta decay with alpha decay and with spontaneous fission prevents this method from being used for larger Z.Starting in 1955, heavy ion accelerators began to deliver beams with high enough intensity and energy to compete in the production of transuranic isotopes. The first positive result was the production of two californium isotopes ($Z = 98$) with the fusion of carbon and uranium nuclei:

$$\begin{aligned} {}^{12}_{6}C + {}^{238}_{92}U &\rightarrow {}^{244}_{98}Cf + 6n, \\ {}^{12}_{6}C + {}^{238}_{92}U &\rightarrow {}^{246}_{98}Cf + 4n \end{aligned} \qquad (2.59)$$

This opened up the possibility of directly reaching the nucleus one wants to create from the fusion of two smaller nuclei. The difficulty of such task is that the cross sections for the production of heavy isotopes are extremely low. For example, the reaction $^{50}_{22}\mathrm{Ti} + ^{208}_{82}\mathrm{Pb} \rightarrow ^{257}_{104}\mathrm{Rf} + \mathrm{n}$, which produces the element rutherfordium, has a cross section of only 5 nb. A small increase in the charges drastically reduces this value. The cross section for the fusion reaction $^{58}_{26}\mathrm{Fe} + ^{208}_{82}\mathrm{Pb} \rightarrow ^{265}_{108}\mathrm{Hs} + \mathrm{n}$ is 4 pb. In comparison, the typical cross sections of DIC for heavy nuclei are in the range 1–2 b.

In spite of the experimental refinement that these low cross sections demand, one is able to produce isotopes with charge as heavy as $Z = 118$. The understanding of the mechanisms that lead to fusion is, however, not fully understood. According to the traditionally accepted model, the fusion of two nuclei proceeds in two stages: the formation of a compound nucleus and the de-excitation of the compound nucleus by evaporation of particles, preferentially neutrons. The difficulties for the materialization of the process in very heavy nuclei reside in both stages [16].

2.6 Electromagnetic Probes

2.6.1 Coulomb Excitation

Coulomb excitation is an inelastic scattering process in which a nucleus excites another nucleus with its electromagnetic field V. This field can be decomposed in terms of a series of multipoles, for example, $E1, E2, M1, \cdots$, which carry well-defined angular momenta and parities. At low bombarding energies, $E2$ (quadrupole) excitations are more common, while at higher energies, $E1$ (dipole) excitations dominate. The Coulomb excitation cross section in first-order perturbation theory is given by

$$\frac{\mathrm{d}\sigma_{i\to f}}{\mathrm{d}\Omega} = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{el}} \frac{16\pi^2 Z_2^2 e^2}{\hbar^2} \times \sum_{\pi\lambda\mu} \frac{B(\pi\lambda, I_i \to I_f)}{(2\lambda+1)^3} |S(\pi\lambda,\mu)|^2 \tag{2.60}$$

where $(\mathrm{d}\sigma/\mathrm{d}\Omega)_{\mathrm{el}}$ is the Rutherford elastic scattering cross section, Z_2 is the charge of the projectile, $B(\pi\lambda, I_i \to I_f)$ is the *reduced matrix element* of the target, $\pi\lambda = E1, E2, M1, \ldots$ is the multipolarity of the excitation, and $\mu = -\lambda, -\lambda+1, \ldots, \lambda$. The orbital integrals $S(\pi\lambda,\mu)$ contain information on the dynamics of the reaction [17].

Because Coulomb excitation is an external process, that is, it occurs when the nucleons from one nucleus are outside the nuclear matter distribution from the other nucleus, the matrix elements for Coulomb excitation are the same as those for excitation by real photons (except for $E0$ Coulomb excitations, which are very small). Therefore, Coulomb excitation cross sections can be written as

$$\frac{\mathrm{d}\sigma_C(E_x)}{\mathrm{d}E_x} = \sum_{E\lambda} \frac{n_{E\lambda}(E_x)}{E_x} \sigma^{\gamma}_{E\lambda}(E_x) + \sum_{M\lambda} \frac{n_{M\lambda}(E_x)}{E_x} \sigma^{\gamma}_{M\lambda}(E_x) \tag{2.61}$$

where $\sigma^{\gamma}_{\pi\lambda}(E_x)$ are the *photonuclear cross sections* for the multipolarity $\pi\lambda$, E_x is the excitation energy, and $n_{M/E\lambda}(E_x)$ are known as *virtual photon numbers* [17].

The photonuclear cross sections are related to the reduced matrix elements, for the excitation energy E_x, through the relation [17]

$$\sigma^{\pi\lambda}_{\gamma}(E_x) = \frac{(2\pi)^3(\lambda+1)}{\lambda[(2\lambda+1)!!]^2} \left(\frac{E_x}{\hbar c}\right)^{2\lambda-1} \times \frac{\mathrm{d}B(\pi\lambda, E_x)}{\mathrm{d}E_x} \tag{2.62}$$

where $\mathrm{d}B/\mathrm{d}E_x$ are the *electromagnetic response functions*, such that

$$B(\pi\lambda, I_i \to I_f) = \int \mathrm{d}E_x \frac{\mathrm{d}B(\pi\lambda, E_x)}{\mathrm{d}E_x}$$

For differential cross sections, an analog expression exists, namely,

$$\frac{\mathrm{d}\sigma_C(E_x)}{\mathrm{d}\Omega} = \frac{1}{E_x} \sum_{\pi\lambda} \frac{\mathrm{d}n_{\pi\lambda}}{\mathrm{d}\Omega}(E_x,\theta) \sigma^{\pi\lambda}_{\gamma}(E_x) \tag{2.63}$$

where Ω denotes the solid scattering angle. This is not different from Eq. (2.60), only rewritten in a more practical way. It helps us to see how we can access information on phenomena induced by real photons by using the field of a charged particle.

In Figure 2.13, we show a calculation (with $E_\gamma \equiv E_x$) of the virtual photons for the $E1$ multipolarity, "as seen" by a projectile passing by a lead target at impact parameters equal to and exceeding $b = 12.3$ fm, for three typical bombarding energies. As the projectile energy increases,

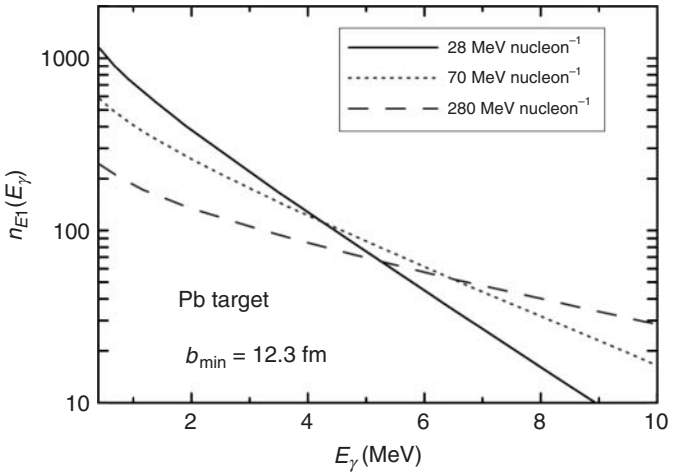


Figure 2.13 Number of virtual photons for the $E1$ multipolarity, "as seen" by a projectile passing by a lead target at impact parameters $b_{min} = 12.3$ fm and larger, for three typical bombarding energies.

more virtual photons of large energy are available for the reaction. This increases the energy of states probed in the excitation process, making it possible to study giant resonances, lepton and meson production, and so on [17].

2.6.2 Photonuclear Reactions and Giant Resonances

A photonuclear reaction is a reaction resulting from the interaction of the electromagnetic radiation with a nucleus. Therefore, one can access information that are complementary to Coulomb excitation (and vice versa).

When the energy of the photon is located above the separation energy of a nucleon, the cross section of photoabsorption reveals the presence of characteristic sharp resonances. However, when the incident energy reaches the range of 15–25 MeV, a new change appears in the cross section, with the presence of a wide and large peak, called *giant electric dipole resonance* (GDR) ($E1$ excitation). Figure 2.14 exhibits the excitation function of photoabsorption of ^{120}Sn at photon energies around the GDR at 15 MeV.

The giant resonance occurs in nuclei along the whole periodic table, with the resonance energy decreasing as $E_{GDR}80/A^{1/3}$ for $A > 20$. Their widths are almost all in the range between 3.5 and 5 MeV. In few cases, they can reach 7 MeV. They are a collective excitation, that is, an excitation simultaneously involving many nucleons . The time-varying electric field of the photon is very effective in inducing collective oscillations of protons against neutrons.

The GDR resonance arises from an excitation that transfers by one unit of angular momentum to the nucleus ($\Delta l = 1$). If the nucleus is even–even it is taken to a 1^- state. The transition also changes the isospin of one unit ($\Delta T = 1$) and, so it is also called *isovector resonance*. The photon can excite giant isoscalar

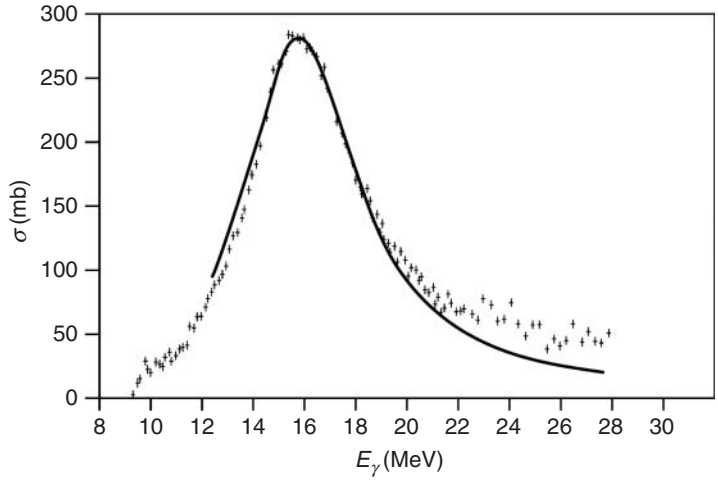


Figure 2.14 Giant resonance in the absorption of photons by ^{120}Sn [18].

resonances less effectively, but still with appreciable cross sections, with $\Delta T = 0$. But electric quadrupole ($\Delta l = 2$) and electric monopole ($\Delta l = 0$) are observed mostly in reactions with charged particles. In a giant electric quadrupole resonance, the nucleus oscillates between a spherical (assuming that this is the form of the ground state) and an ellipsoidal form. If protons and neutrons vibrate in phase, we have an isoscalar resonance ($\Delta T = 0$) and if they are oscillate in opposite phase we have an isovector resonance ($\Delta T = 1$). The giant monopole resonance is a very special case of nuclear excitation where the nucleus contracts and expands radially, maintaining its original form but changing its volume. It is also called *breathing mode*. It can also happen in isoscalar and isovector forms. It is an important way to study the *compressibility of nuclear matter*.

Besides electric giant resonances, associated with a variation in the form of the nucleus, magnetic giant resonances exist, which involve *spin vibrations*. In these, nucleons with spin up move out of phase with nucleons with spin down. The number of nucleons involved in the process cannot be very large because it is limited by the Pauli principle. The magnetic resonances can also separate into isoscalar resonances, where protons and neutrons of the same spin vibrate against protons and neutrons of opposite spin, and isovector, where protons with spin up and neutrons with spin down vibrate against their corresponding partners with opposite spins. The last cases are effectively probed in charge-exchange reactions, that is, when the projectile charge changes to $Z - 1$ or $Z + 1$, as in (d,p) and (d,n). They are known as *giant Gamow–Teller resonances*.

Giant resonances can also be produced in an excited nucleus, a case known as the *Brink–Axel hypothesis* [19]. A special case is when two giant resonances are excited simultaneously in the same nucleus. The double giant dipole resonance (or *multiphonon giant resonance* was observed for the first time in reactions with double charge exchange induced by pions in ^{32}S [20]. But Coulomb excitation proved to be the best probe, as shown theoretically in Refs [17, 21].

2.6.3 Electron Scattering

Electron scattering experiments have provided a rich database on most stable or long-lived nuclei. Electrons are structureless point-like objects that only interact electromagnetically. Therefore, electron scattering avoids the complexity of the strong interaction between the projectile and the target and provides clean information about the charge distribution in the nucleus.

Considering relativity and the spin of the electrons, the differential cross section for elastic scattering by spinless point-like nuclei can be expressed by the *Mott scattering formula* (for simplicity, we use $\hbar = c = 1$ units):

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Mott}} = \frac{Z_T^2 e^4 \cos^2(\theta/2)}{4p_0^2 \sin^4(\theta/2)[1+(2p_0/M)\sin^2(\theta/2)]} \tag{2.64}$$

where Z_{Te} and M are the charge and mass of the target nucleus, respectively, and p_0 is the momentum of the incoming electron. As the nucleus is not a point-like particle, the formula is modified by adding the nuclear electric and magnetic form factors, which contain information of the charge and magnetization distributions inside the nucleus.

The *Rosenbluth formula* accounts for the electron spin and explicitly expresses the cross section for arbitrary nuclei as

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right) = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Mott}} \left\{A_0(q^2) + B_0(q^2) \times \left[\frac{1}{2} + \left(1 + \frac{q^2}{4M^2}\right)\right] \tan^2(\theta/2)\right\}$$

where $q^2 = (\mathbf{p}_\mathrm{f} - \mathbf{p}_\mathrm{i})^2 - (E_\mathrm{f} - E_\mathrm{i})^2$ is the 4-momentum transfer squared, where $\mathbf{p}_\mathrm{i}(\mathbf{p}_\mathrm{f})$ and $E_\mathrm{i}(E_\mathrm{f})$ are the electron incoming (outgoing) momentum and energy. $A_0(q^2)$ and $B_0(q^2)$, functions of q^2, are the form factors associated with the charge and magnetization distribution of the nucleus, respectively. For a spin-0 nucleus,

$$A_0(q^2) = \frac{G_E^2(q^2)}{1 + q^2/4M^2} \quad \text{and} \quad B_0(q^2) = 0 \tag{2.65}$$

where $G_E(q^2)$ is the *Sachs charge form factor*, and the factor $1 + q^2/4M^2$ is the kinematical recoil correction. For $q \to 0$, $q^2 \approx \mathbf{q}^2$, where $\mathbf{q}$ is the three-momentum transfer, and to leading order in powers of $\mathbf{q}$:

$$G_E(q^2) \simeq G_E(\mathbf{q}^2) = \int \rho_T(\mathbf{r}) \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot r} \mathrm{d}^3 r \tag{2.66}$$

One thus sees that electron scattering is related to the Fourier transform of the nuclear charge density, $\rho_T(\mathbf{r})$. If the function $G_E(\mathbf{q}^2)$ is mapped for a sufficient number of momentum transfers $\mathbf{q}^2$, the Fourier transform can be inverted and $\rho_T(\mathbf{r})$ can be mapped with precision. For heavy elements ($A > 10$), this technique shows that the charge radius is closely proportional to $1.12\ A^{1/3}$ fm, where A is the nuclear mass number. This indicates the density saturation of the nuclear matter.

When $\mathbf{q}^2 = 0$, inelastic electron scattering probes the same multipolarity transitions ($E1, E2, \ldots$) as in Coulomb excitation or with real photons. However, in contrast to the later probes, electron scattering also provides information on nuclear excitations for cases in which $\mathbf{q}^2 \neq 0$. This is useful in many aspects. For example, at high momentum transfers, the momentum transfer $\mathbf{q}$ can be absorbed by a nucleon in the nucleus, which acquires an energy of the order of $\mathbf{q}^2/2m_N^*$, thus probing the effective mass m_N^* of a nucleon

in the nucleus. This is of relevance to understand the effects of nucleon–nucleon interactions in the nuclear environment. The q-region where this occurs is termed *quasi-free scattering*. As electrons can penetrate the nuclei, they can also more effectively probe monopole (or breathing mode) excitations of the nuclei, a case in which the nuclear mass distribution vibrates radially. This is of importance for determining the *compressibility modulus* of the nuclear matter. Finally, in high-energy electron scattering, the electron can penetrate deep inside the nucleons and probe the spin and charge distributions of quarks and gluons inside the nucleon [22].

2.7 Relativistic Nuclear Collisions

2.7.1 Transport Theories and Equation of State

As the bombarding energy in nucleus–nucleus collisions increases, the structure aspects of the nuclei become less relevant. Except for the bulk properties of the nuclei (size and number of nucleons), the physics involved is primarily due to the individual, and sometimes collective, hadronic collisions. Several theoretical tools are used to describe these reactions, and, in particular, we quote (i) *time-dependent Hartree–Fock* (TDHF), (ii) *anti-symmetrized molecular dynamics* (AMD), (iii) *Boltzmann–Uehling–Uhlenbeck* (BUU), and so on.

At intermediate energies of $E_{\rm lab} \sim 100 - 1000$ MeV nucleon^{-1}, the nucleons and the products of their collisions can be described individually and their propagation can be described by semiclassical equations. One of such equations, and perhaps the most popular in such studies, is the so-called *BUU equation*:

$$\frac{\partial f}{\partial t} + \left(\frac{\mathbf{p}}{m} + \nabla_{\mathbf{p}} U\right)\cdot\nabla_{\mathbf{r}} f - \nabla_{\mathbf{r}} U\cdot\nabla_{\mathbf{r}} f = \int \mathrm{d}^3 p_2 \int \mathrm{d}\Omega\, \sigma_{\rm NN}(\Omega)|\mathbf{v}_1 - \mathbf{v}_2| \times\left\{f_1' f_2' \left[1 - f_1\right]\left[1 - f_2\right] - f_1 f_2 \left[1 - f_1'\right] \times \left[1 - f_2'\right]\right\}. \quad (2.67)$$

If $\mathrm{d}N$ is the number of particles in the volume element $\mathrm{d}^3 r$ and whose momenta fall in the momentum element $\mathrm{d}^3 p$ at time t, then the distribution function $f(\mathbf{r}, p, t)$ is given by $\mathrm{d}N = f(\mathbf{r}, p, t)\mathrm{d}^3 r \mathrm{d}^3 p$. Thus, the BUU equation is an equation for the distribution function $f(\mathbf{r}, p, t)$. To account for the effect of each particle interacting with all others, one introduces the concept of *mean-field*, $U(\mathbf{r}, p, t)$. This mean-field exerts a force on each particle, given by $-\nabla_{\mathbf{r}} U(\mathbf{r}, p, t)$. In addition, the momentum dependence of the potential introduces a dependence through the derivative $-\nabla_{\mathbf{p}} U(\mathbf{r}, p, t)$.

Owing to the nucleon–nucleon collisions, the distribution function within $\mathrm{d}^3 r \mathrm{d}^3 p$ can also be modified by nucleons leaving (or entering) this volume. This is taken care by of the *collision term*, that is, the right-hand side of the BUU equation. $\sigma_{\rm NN}$ is the nucleon–nucleon differential scattering cross section, $\mathbf{v}_1$ and $\mathbf{v}_2$ are the velocities of two colliding nucleons. The first factor inside braces are for collisions populating the volume element and the second term for those depleting it. The factors $(1 - f)$ account for Pauli blocking of final occupied states. The integrals average over scattering angle and over all collisions within $\mathrm{d}^3 r \mathrm{d}^3 p$. The BUU equation falls in the category of what one calls *quantum transport theories*. Hadronic transport theories have been quite successful in applications, describing a multitude of measured particle spectra.

Equation (2.67) needs as basic ingredients the mean-field U and the cross section σ_{NN}. Because these two quantities are related to each other, one should in principle derive them in a self-consistent microscopic approach, as in the *Brueckner theory*.

An important ingredient in the transport theory calculations is the compressibility K of nuclear matter, which refers to the second derivative of the compressional energy E with respect to the density [1)]

$$K = 9\rho^2 \frac{\partial^2}{\partial \rho^2} \left(\frac{E}{A} \right) \tag{2.68}$$

This is an important quantity, for example, for nuclear astrophysics. Supernova models might or not lead to explosions depending on the value of K. The central collisions of heavy nuclei are one of the few probes of this quantity in the laboratory. The dependence of the calculations on K follow from the dependence of the mean-field potential U ($U \sim E/A +$ kinetic energy terms) on the particle density ρ. A typical parametrization for U is the Skyrme parametrization $U = a\rho/\rho_0 + b(\rho/\rho_0)^\sigma$.

The output of Eq. (2.67) is the distribution function % $f(\mathbf{r}, p, t)$, which allows one to calculate many properties of heavy ion collisions. Let us quote *collective flows*, proton and neutron production rates (subthreshold and above threshold), pion and kaon yields, and so on. Combining Eq. (2.67) with a *phase–space coalescence model*, one can also calculate such quantities as exclusive flows and intermediate fragment formations.

The dynamics of the central high-energy reactions can be broken down into several stages. Baryon-density contour plots are shown in Figure 2.15 for 400 MeV nucleon^{-1} Au + Au collisions at $b = 0$, which will serve to illustrate our points.

Following an initial interpenetration of projectile and target densities, the NN collisions begin to thermalize matter in the overlap region making the momentum distribution there centered at zero momentum in the CM. The density in the overlap region rises above normal and a disk of excited and compressed matter forms at the center of the system. More and more matter dives into the region with compressed matter that begins to expand in transverse directions. At late stages, when the whole matter is excited, transverse expansion predominates.

2.7.2 Kinematics

In relativistic nucleus–nucleus collisions, other definitions are best suited for discussing relations between energies, momenta, and angles. The *rapidity* is a variable frequently used to describe the behavior of particles in inclusively measured reactions. It is defined by

$$y = \frac{1}{2} \ln \left(\frac{E + p_\parallel}{E - p_\parallel} \right) \tag{2.69}$$

which corresponds to $\tanh y = p_\parallel / E$, where y is the rapidity, $p_\parallel$ is the longitudinal momentum along the direction of the incident particle, E is the energy, both defined for a given particle. The accessible range of rapidities for a given reaction is determined by the available CM energy and all participating particles' rest masses. One usually gives the limit for the incident particle, elastically scattered at zero angle:

$$|y_{max}| = \ln \left[\frac{E + p}{m} \right] = \ln (\gamma + \gamma\beta) \tag{2.70}$$

1) The factor of 9 is a historical artifact from the convention originally defining K.

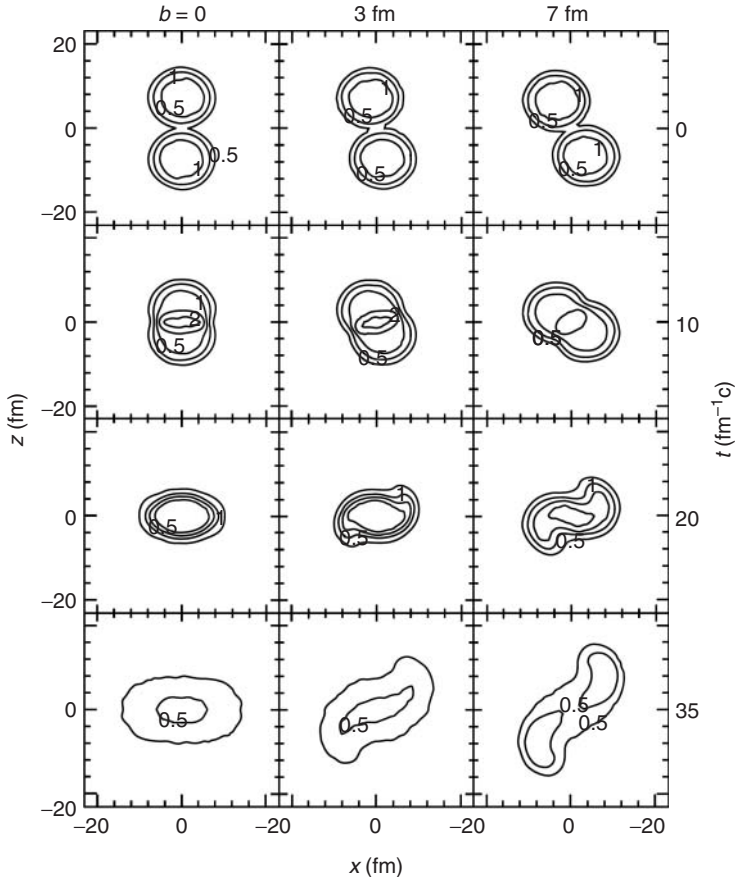


Figure 2.15 Contour plots of baryon density in the reaction plane in Au + Au collisions at 400 MeV nucleon^{-1}. The displayed contour lines are for the densities $\rho/\rho_0 = 0.1$, 0.5, 1, 1.5, and 2 [23].

where $\beta \equiv v$ is the velocity and all variables referring to the through-going particle given in the desired frame of reference (i.e., in the CM).

Note that $\partial y/\partial p_{\parallel} = 1/E$. A Lorentz boost β along the direction of the incident particle adds a constant, $\ln(\gamma + \gamma\beta)$, to the rapidity. Rapidity differences, therefore, are invariant to a Lorentz boost. Statistical particle distributions are flat in y for many physics production models. Frequently, the simpler variable *pseudorapidity* η is used instead of rapidity.

The pseudorapidity is a handy variable to approximate the rapidity if the mass and momentum of a particle are not known. It is an angular variable defined by

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{2.71}$$

whose inverse function is $\theta = 2\arctan(e^{-\eta})$, where θ is the angle between the particle being considered and the undeflected beam. η is the same as the rapidity y if one sets $\beta = 1$ (or $m = 0$).

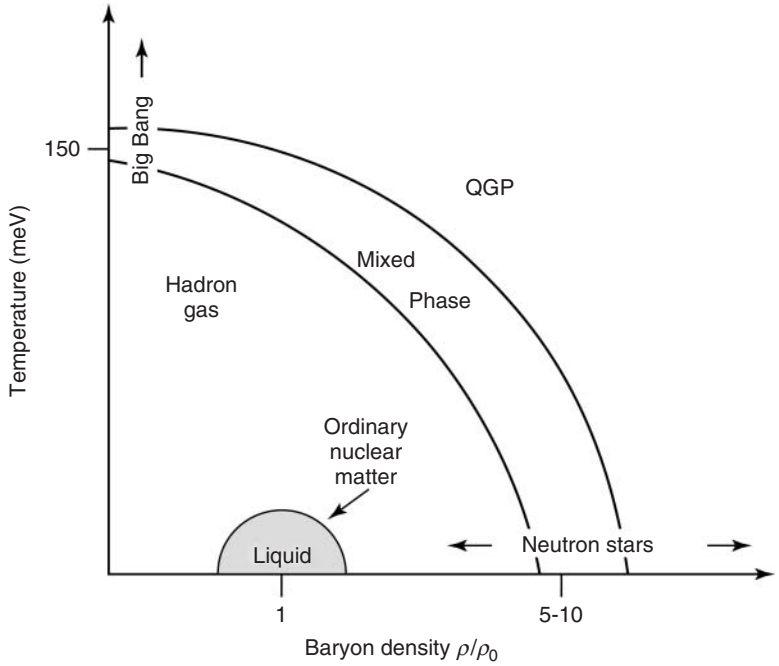


Figure 2.16 As water comes in different phases (solid, liquid, and gas), so nuclear matter can come in its normal hadronic form and at sufficiently high temperature and density, in the form of a deconfined state of quarks and gluons. The diagram shows how nuclear matter should behave as a function of density and temperature.

Statistical distributions plotted in pseudorapidity rather than rapidity undergo transformations that have to be estimated by using a kinematic model for the interaction.

2.7.3 The Quark–Gluon Plasma (QGP)

The primary motivation for studying ultrarelativistic heavy ion collisions is to gain an understanding of the equation of state of nuclear, hadronic, and partonic matter, commonly referred to as *nuclear matter*. Displayed in Figure 2.16 is a schematic phase diagram of nuclear matter. The behavior of nuclear matter as a function of temperature and density (or pressure), shown in Figure 2.16, is governed by its equation of state.

Conventional nuclear physics is concerned primarily with the lower left portion of the diagram at low temperatures and near normal nuclear matter density. Here normal nuclei exist and at low excitation, a liquid–gas phase transition is expected to occur. This is the focus of experimental studies using low-energy heavy ions. At somewhat higher excitation, nucleons are excited into baryonic resonance states, along with accompanying particle production and hadronic resonance formation. In relativistic heavy ion collisions, such excitation is expected to create hadronic resonance matter.

We now briefly discuss the QGP signatures in nucleus–nucleus collisions. For more details see, for example, Refs [24, 25]. One group of such signatures can be classified as *thermodynamic variables*. This class involves determination of the energy density ϵ, pressure P, and entropy density s of the interacting system as a function of the temperature T and the baryochemical potential μ_B. Experimental observables can be identified with these variables and thus their relative behavior can be determined. If a phase transition to QGP occurs, a rapid rise in the effective number of degrees of freedom, expressed by $/T^4$ or s/T^3, should be observed over a small range of T. The variables T, s, and can be identified with the average transverse momentum p_T, the hadron rapidity density $\mathrm{d}N/\mathrm{d}y$, and the transverse energy density $\mathrm{d}E_T/\mathrm{d}y$, respectively. The transverse energy produced in the interaction is $E_T = \sum_i E_i \sin\theta_i$, where E_i and θ_i are the kinetic energies of the ejectiles and the emission angles.

Electromagnetic (EM) probes, such as photons and leptons, provide information on the various stages of the interaction without modification by final state interactions. These probes may provide a measure of the thermal radiation from a QGP, if a region of photon energy, or equivalently lepton pair invariant mass, can be isolated for emission from a QGP relative to other processes. However, the yields for EM probes are small relative to background processes, which are primarily EM decays of hadrons and resonances. Lepton pairs from the QGP are expected to be identifiable in the 1–10 GeV invariant mass range. The widths and positions of the ρ, ω, and ϕ peaks in the lepton pair invariant mass spectrum are expected to be sensitive to medium-induced changes of the hadronic mass spectrum.

The production of J/Ψ particles in a quark–gluon plasma is predicted to be suppressed (Figure 2.17). This is a result of the *Debye screening* of a $c\bar{c}$ pair, initially formed in the QGP by fusion of two incident gluons. Less tightly bound excited states of the $c\bar{c}$ system, such as Ψ' and χ_c, are more easily dissociated and will be suppressed even more than the J/Ψ.

A long-standing prediction for a signature of QGP formation is the enhancement of strange hadrons. The production of strange hadrons relative to nonstrange hadrons is suppressed in hadronic reactions. This suppression increases with increasing strangeness content of the hadron. In a QGP, the strange quark content is rapidly saturated by $s\bar{s}$ pair production in gluon–gluon reactions, resulting in an enhancement in the production of strange hadrons. Thus, multistrange baryons and strange antibaryons are predicted to be strongly enhanced when a QGP is formed.

The connection between energy loss of a quark and the color-dielectric polarizability of the medium can be established in analogy with the theory of electromagnetic energy loss. Although radiation is a very efficient energy loss mechanism for relativistic particles, it is strongly suppressed in a dense medium by the *Landau–Pomeranchuk effect* [24]. Adding the two contributions, the stopping power of a quark–gluon plasma is predicted to be higher than that of hadronic matter. A quark or *gluon jet* propagating through a dense medium will not only lose energy but will also be deflected. This effect destroys the coplanarity of the two jets from a hard parton–parton scattering with the incident beam axis. The angular deflection of the jets also results in an azimuthal asymmetry. The presence of a quark–gluon plasma is

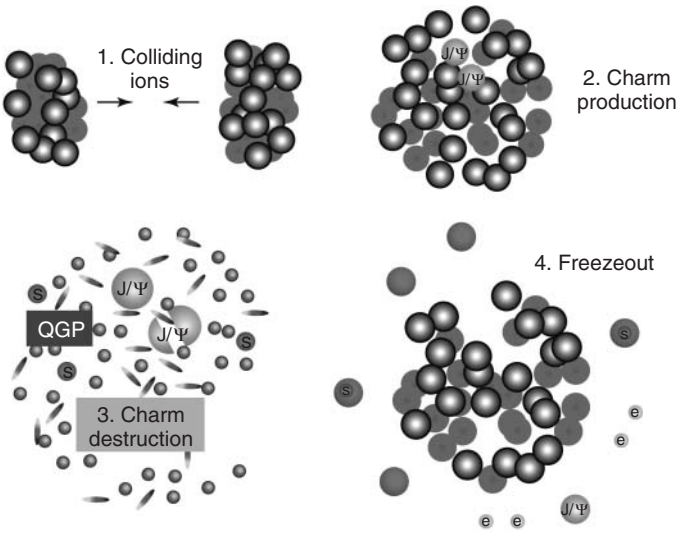


Figure 2.17 Formation and evolution of J/Ψ particles in relativistic heavy ion collisions. After the formation, the J/Ψs are dissociated in the plasma owing to color screening. The end effect is a smaller number J/Ψs than expected from pure hadron–hadron multiple collisions.

also predicted to attenuate the emission of jet pairs opposite to the trigger jet.

All the above quark–gluon plasma signatures have been studied to maturity at the CERN (European Organization for Nuclear Research) – Large Hadron Collider (LHC) and Super Proton Synchrotron (SPS) and the Brookhaven National Laboratory (BNL) – Relativistic Heavy Ion Collider (RHIC) heavy ion accelerator facilities. For more details, the following references discuss: (i) direct photons [26, 27], (ii) dileptons [28], (iii) J/Ψ production [29], and (iv) quark in-medium attenuation, also known as *highp_T quenching* or *jet attenuation* [30].

2.7.4 Ultraperipheral Relativistic Heavy Ion Collisions

Relativistic heavy ion collisions have also been used to study nuclear photoexcitation and photoproduction of hadrons. The strong electromagnetic fields from a heavy ion allow for the study of multiphoton excitation of nuclear targets, for example, the double-giant resonance [17, 21]. Reactions such as multiple vector meson production can be used for studies involving polarized photons [17]. The high photon energies can be used to study the gluon density in heavy nuclei [31] at low Feynman x.

Photon–photon (or two-photon) processes have long been studied at e^+e^- colliders. They are an excellent tool for many aspects of meson spectroscopy and tests of QED [32]. At hadron colliders, they are also used to study atomic physics processes, often involving electrodynamics in strong fields. One striking success was the production of antihydrogen atoms at CERN's LEAR2 [33] and at the Fermilab Tevatron [34]. At the highest energy colliders, reactions such as $\gamma\gamma \to X$ may be used to probe the quark content and spin

structure of meson resonances [17]. Production of meson or baryon pairs can also probe the internal structure of hadrons. At the LHC, electroweak processes such as $\gamma\gamma \to W^+W^-$ may be probed. The cross section for two-photon processes is, in analogy to Eq. (2.61), as follows [17]:

$$\sigma = \int \mathrm{d}E_1 \mathrm{d}E_2 \frac{n_1(E_1)}{E_2} \frac{n_1(E_2)}{E_2} \sigma_{\gamma\gamma\to X}(E_1, E_2) \tag{2.72}$$

where $n_i(E_i)$ is the number of virtual photons of energy E_i generated by the ion i, and $\sigma_{\gamma\gamma\to X}(E_1, E_2)$ is the cross sections for production of particle X by two real photons with energies E_1 and E_2.

2.8 Nuclear Reactions in Stars

2.8.1 Hydrogen and CNO Cycles

Energy production in stars is a well-known process. The initial energy that ignites the process arises from the gravitational contraction of a mass of gas. The contraction increases the pressure, temperature, and density, at the center of the star until the values are reached at which the thermonuclear reactions start, initiating the star's lifetime. The energy liberated in these reactions yield a pressure in the plasma, which opposes compression due to gravitation. Thus, an equilibrium is reached for the energy which is produced, the energy which is liberated by radiation, the temperature, and the pressure.

The Sun is a star in its initial phase of evolution. The temperature at its surface is 6000 °C, while in its interior the temperature reaches 1.5×10^7K, with a pressure given by 6×10^{11} atm and density 150g cm^{-3}. The present mass of the Sun is $M_\odot = 2 \times 10^{33}$g and its main composition is hydrogen (70%), helium (29%), and less than 1% of heavier elements, such as carbon, oxygen, and so on.

What are the nuclear processes that originate the huge thermonuclear energy of the Sun, and that have lasted 4.6×10^9 years (the assumed age of the Sun)? It cannot be the simple fusion of two protons, or of α-particles, or even the fusion of protons with α-particles, since neither ${}^2_2\mathrm{He}, {}^8_4\mathrm{Be}$, or ${}^5_3\mathrm{Li}$, are stable. The only possibility is the proton–proton fusion in the form

$$\mathrm{p} + \mathrm{p} \to \mathrm{d} + \mathrm{e}^+ + \nu_e \tag{2.73}$$

which occurs via β-decay, that is, due to the weak interaction. The cross section for this reaction for protons of energy around 1 MeV is very small, of the order of 10^{-23} b. The average lifetime of protons in the Sun due to the transformation to deuterons by means of Eq. (2.73) is about 10^{10}y. This explains why the energy radiated from the Sun is approximately constant in time, and not an explosive process.

The deuteron produced in the above reaction is consumed almost immediately in the process

$$\mathrm{d} + \mathrm{p} \to {}^3_2\mathrm{He} + \gamma \tag{2.74}$$

The resulting ${}^3_2\mathrm{He}$ reacts by means of

$${}^3_2\mathrm{He} + {}^3_2\mathrm{He} \to {}^4_2\mathrm{He} + 2\mathrm{p} \tag{2.75}$$

which produces the stable nucleus ${}^4_2\mathrm{He}$ with a great energy gain, or by means of the reaction

$${}^3_2\mathrm{He} + {}^4_2\mathrm{He} \to {}^7_4\mathrm{Be} + \gamma \tag{2.76}$$

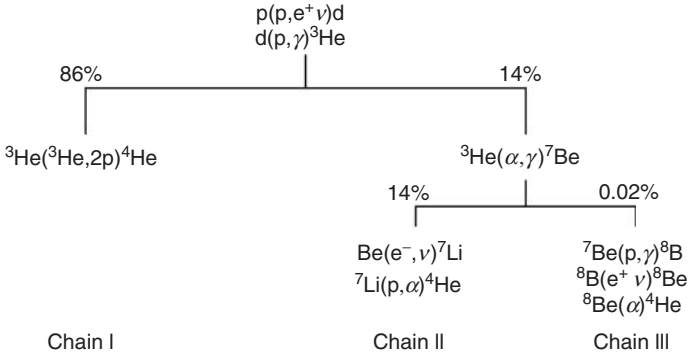


Figure 2.18 The p–p chain reaction (p–p cycle). The percentage for the several branches are calculated in the center of the Sun [35].

In the second case, a chain reaction follows as

$$^{7}_{4}\mathrm{Be} + \mathrm{e}^{-} \rightarrow {}^{7}_{3}\mathrm{Li} + \nu_e, \quad {}^{7}_{3}\mathrm{Li} + \mathrm{p} \rightarrow 2(^{4}_{2}\mathrm{He}) \tag{2.77}$$

or

$$\begin{aligned} {}^{7}_{4}\mathrm{Be} + \mathrm{p} &\rightarrow {}^{8}_{5}\mathrm{B} + \gamma, \\ {}^{8}_{5}\mathrm{B} &\rightarrow 2(^{4}_{2}\mathrm{He}) + \mathrm{e}^{+} + \nu_e \end{aligned} \tag{2.78}$$

The chain reaction (2.73)–(2.78) is called the *hydrogen cycle*. The result of this cycle is the transformation of four protons into an α-particle, with an energy gain of 26.7 MeV, about 20% of which are carried away by the neutrinos (Figure 2.18).

If the star contains heavier elements, another cycle can occur – the *carbon cycle* or *CNO cycle* [36]. In this cycle, the carbon, oxygen, and nitrogen nuclei are catalyzers of nuclear processes, with the end product also in the form $4\mathrm{p} \rightarrow {}^{4}_{2}\mathrm{He}$. Figure 2.19 describes the CNO cycle. Owing to the larger Coulomb repulsion between the carbon nuclei, it occurs at higher temperatures (larger relative energy between the participant nuclei), up to 1.4×10^{7}K. In the Sun, the hydrogen cycle prevails. However, in stars with larger temperatures, the CNO cycle is more effective.

After the protons are transformed into helium at the center of a star like our Sun, the fusion reactions start to consume protons at the surface of the star. At this stage, the star starts to become a *red giant*. The energy generated by fusion increases the temperature and expands the surface of the star. The star's luminosity increases. The red giant contracts again after the hydrogen fuel is burned.

Other thermonuclear processes start. The first is the helium burning when the temperature reaches 10^{8}K and the density becomes $10^{6}\mathrm{g\,cm^{-3}}$. Helium burning starts with the triple-capture reaction

$$3(^{4}_{2}\mathrm{He}) \rightarrow {}^{12}_{6}\mathrm{C} + 7.65\mathrm{MeV} \tag{2.79}$$

followed by the formation of oxygen via the reaction

$${}^{12}_{6}\mathrm{C} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{16}_{8}\mathrm{O} + \gamma \tag{2.80}$$

For a star with the mass of the Sun, helium burning occurs in about 10^{7}y. For a much heavier star, the temperature can reach 10^{9}K. The compression process followed by the burning of heavier elements

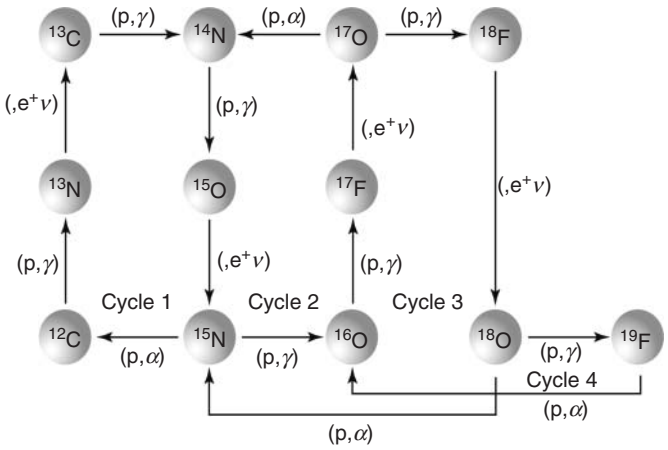


Figure 2.19 The CNO cycle. (Courtesy of Frank Timmes.)

can lead to the formation of iron and nickel. After that, the thermonuclear reactions are no more energetic and the star stops producing nuclear energy.

2.8.2 Synthesis of Heavier Elements

In Figure 2.20, we show the relative distribution of elements in our galaxy. It has two distinct regions: in the region $A < 100$ it decreases with A approximately like an exponential, whereas for $A > 100$ it is approximately constant, except for the peaks in the region of the magic numbers $Z = 50$ and $N = 50, 82, 126$.

The thermonuclear processes (2.73)–(2.80) can explain the relative abundance of $^4_2$He, $^{12}_6$C, and $^{16}_8$O. The processes occurring after $^4_2$He burning mainly form isotopes of $^{20}_{10}$Ne, $^{24}_{12}$Mg, and $^{28}_{14}$Si. We can understand the small abundance of the elements Li, Be, and B as due to the small velocity with which they are formed via the reaction (2.76) and the first equation of (2.77), while they are rapidly consumed by the second reaction in (2.77) and the first reaction in (2.78).

The synthesis of elements heavier than oxygen occur when, after helium burning, a new compression and heating of the star raises the temperature to values higher than 6×10^8K. This situation ignites an intense carbon burning:

$$^{12}_6\mathrm{C} + {}^{12}_6\mathrm{C} \rightarrow {}^{20}_{10}\mathrm{Ne} + {}^4_2\mathrm{He} \rightarrow {}^{23}_{11}\mathrm{Na} + \mathrm{p} \rightarrow {}^{23}_{12}\mathrm{Mg} + \mathrm{n} \rightarrow {}^{24}_{12}\mathrm{Mg} + \gamma \quad (2.81)$$

Carbon and oxygen can also burn simultaneously:

$$^{12}_6\mathrm{C} + {}^{16}_8\mathrm{O} \rightarrow {}^{24}_{12}\mathrm{Mg} + {}^4_2\mathrm{He}, \text{etc.},$$
$$^{16}_8\mathrm{O} + {}^{16}_8\mathrm{O} \rightarrow {}^{28}_{14}\mathrm{Si} + {}^4_2\mathrm{He}, \text{etc.} \quad (2.82)$$

For temperatures above 3×10^9K, more photonuclear processes appear. These yield more nuclei to be burned and heavier nuclei are produced:

$$\gamma + {}^{28}_{14}\mathrm{Si} \rightarrow {}^{24}_{12}\mathrm{Mg} + {}^4_2\mathrm{He},$$
$$^4_2\mathrm{He} + {}^{28}_{14}\mathrm{O} \rightarrow {}^{32}_{16}\mathrm{S} + \gamma, \text{etc.} \quad (2.83)$$

Owing to the large number of free neutrons, many (n, γ)-reactions (radiative neutron capture) elements in the mass range $A = 28, \ldots, 57$ are formed. This

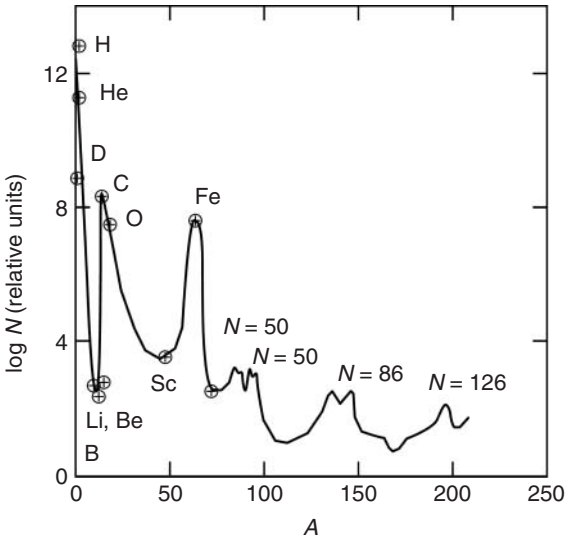


Figure 2.20 Relative distribution of elements in our galaxy.

leads to a large abundance of elements in the iron mass region, which have the largest binding energy per nucleon. For elements heavier than iron, the nuclear fusion processes do not generate energy.

For $A > 100$, the distribution of nuclei cannot be explained in terms of fusion reactions with charged particles. They are formed by the successive capture of slow neutrons and of β^--decay. The maxima of the element distribution in $N = 50, 82, 126$ are due to the small capture cross sections corresponding to the magic numbers. This yields a slew of isotopes at the observed element distribution.

2.8.3 Thermonuclear Cross Sections

The nuclear cross section for a reaction between target j and projectile k is defined by

$$\sigma = \frac{\text{number of reactions (target}^{-1}s^{-1})}{\text{flux of incoming projectiles}} = \frac{r/n_j}{n_k v} \tag{2.84}$$

where the target number density is given by n_j, the projectile number density is given by n_k, and v is the relative velocity between target and projectile nuclei. Then r, the number of reactions per cubic centimeter per second, can be expressed as $r = \sigma v n_j n_k$, or, more generally,

$$r_{j,k} = \int \sigma |v_j - v_k| \mathrm{d}^3 n_j \mathrm{d}^3 n_k \tag{2.85}$$

The evaluation of this integral depends on the type of particles and distributions that are involved. For nuclei j and k in an astrophysical plasma, obeying an MB distribution,

$$\mathrm{d}^3 n_j = n_j \left(\frac{m_j}{2\pi kT}\right)^{\frac{3}{2}} \exp\left(-\frac{m_j v_j^2}{2kT}\right) \mathrm{d}^3 v_j \tag{2.86}$$

Eq. (2.85) simplifies to $r_{j,k} = < \sigma v > n_j n_k$, where $< \sigma v >$ is the average of σv over the temperature distribution in Eq. (2.86).

In astrophysical plasmas with high densities and/or low temperatures, effects of electron screening become highly important. This means that the reacting nuclei,

due to the background of electrons and nuclei, feel a different Coulomb repulsion than in the case of bare nuclei. Under most conditions (with nonvanishing temperatures), the generalized reaction rate integral can be separated into the traditional expression without screening and a screening factor

$$< \sigma v >^{*}_{j,k} = f_{\mathrm{scr}}(Z_j, Z_k, \rho, T, Y_i) < \sigma v >_{j,k} \tag{2.87}$$

This screening factor is dependent on the charge of the involved particles, the density, temperature, and the composition of the plasma. Here Y_i denotes the abundance of nucleus i defined by $Y_i = n_i/(\rho N_A)$, where n_i is the number density of nuclei per unit volume and N_A Avogadro's number. At high densities and low temperatures, screening factors can enhance reactions by many orders of magnitude and lead to pycnonuclear ignition.

When in Eq. (2.85), particle k is a photon, the relative velocity is always c and quantities in the integral are not dependent on $\mathrm{d}^3 n_j$. Thus, it simplifies to $r_j = \lambda_{j,\gamma} n_j$ and $\lambda_{j,\gamma}$ results from an integration of the photodisintegration cross section over a Planck distribution for photons of temperature T

$$r_j = \frac{1}{\pi^2 (c\hbar)^3} \int \mathrm{d}^3 n_j \times \int_0^\infty \frac{c\sigma(E_\gamma) E_\gamma^2}{\exp(E_\gamma/kT) - 1} \mathrm{d}E_\gamma \tag{2.88}$$

There is, however, no direct need to evaluate photodisintegration cross sections, because, due to detailed balance, they can be expressed by the capture cross sections for the inverse reaction $l + m \rightarrow j + \gamma$.

A procedure similar to Eq. (2.88) is used for electron captures by nuclei. Because the electron is about 2000 times less massive than a nucleon, the velocity of the nucleus j is negligible in the CM system in comparison to the electron velocity ($|v_j - v_e| \approx |v_e|$). The electron capture cross section has to be integrated over a Boltzmann, partially degenerate, or Fermi distribution of electrons, depending on the astrophysical conditions. The electron capture rates are a function of T and $n_e = Y_e \rho N_A$, the electron number density [37]. In a neutral, completely ionized plasma, the electron abundance is equal to the total proton abundance in nuclei $Y_e = \sum_i Z_i Y_i$ and $r_j = \lambda_{j,e}(T, \rho Y_e) n_j$.

This treatment can be generalized for the capture of positrons, which are in a thermal equilibrium with photons, electrons, and nuclei. At high densities ($\rho > 10^{12}$ g cm^{-3}), the size of the neutrino scattering cross section on nuclei and electrons ensures that enough scattering events occur to thermalize a neutrino distribution. The inverse process to electron capture (neutrino capture) can occur even then and the neutrino capture rate can be expressed similarly to Eq. (2.88), integrating over the neutrino distribution. Also inelastic neutrino scattering on nuclei can be expressed in this form. Finally, for normal decays, like β or α decays with half-life $\tau_{1/2}$, we obtain an equation similar to Eq. (2.88) or r_j of the previous paragraph with a decay constant $\lambda_j = \ln 2/\tau_{1/2}$ and $r_j = \lambda_j n_j$.

2.9 Reactions with Radioactive Nuclear Beams

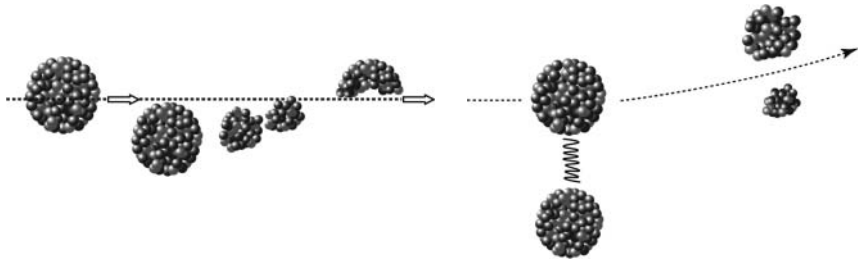
The basic research activity in nuclear physics, driven by the desire to understand the forces that dictate the properties of nuclei, has spawned a large number of beneficial applications. Among its many progeny, we can count reactor- and spallation-based neutron sources,

synchrotron radiation sources, particle physics, materials modification by implantation, carbon dating, and much more. It is an excellent example of the return to society of investment in basic research.

All these achievements have been realized by accelerating the 283 stable or long-lived nuclear species we find here on Earth. In recent years, however, it has become evident that it is now technically possible to create and accelerate unstable nuclei and that there are some 6000–7000 distinct nuclear species which live long enough to be candidates for acceleration. They are the nuclei within the so-called *drip-lines*, the point where the nucleus can no longer hold another particle. This has led to many new opportunities in industry, medicine, materials studies, and the environment.

Assume that a highly energetic uranium projectile ($N/Z \sim 1.6$) hits a target nucleus in an almost central collision, as shown in Figure 2.21. A part of the projectile (*participant*) is scrapped off and forms a highly excited mixture of nucleons with a part of the target. A piece of the projectile (*spectator*) flies away with nearly the same velocity as that of the beam. The neutron-to-proton ratio of the spectator is nearly equal to that of the projectile. Since the N/Z ratio of light nuclei (stable) is close to one, the fragment is far from the stability line. Statistically, a large number of fragments with different N/Z ratios are created and several new exotic nuclei have been discovered in this way. These nuclei can be collected in a secondary beam and further accelerated; they can induce reactions with a target nucleus. This method has become an important tool to study the properties of short-lived isotopes.

Experiments with secondary beams are limited by reaction cross section and *luminosity*. The luminosity L is defined as the product of beam intensity i and target thickness t, $L = i \times t$. The reaction rate N is the product of luminosity and reaction cross section σ_r, $N = \sigma_r \times L$. In most of the reactions, the usable target thickness is limited by the width of the excitation function (i.e., the cross section as a function of the excitation energy). Production reactions with a wide excitation function covering a broad energy range can profit in luminosity by the use of thick targets.



(a) (b)

Figure 2.21 (a) Schematic description of a nuclear fragmentation reaction producing rare isotopes. The lower fragments are called *participants*, while the upper one is called the *spectator*. Using uranium projectiles ($N/Z \sim 1.6$), one expects to produce (light) spectator nuclei of about the same N/Z ratio. (b) Coulomb fission of relativistic projectiles leading to the production of rare isotopes. For a heavy unstable projectile, an exchanged photon with the target can give it enough energy to fragment into several types of isotopes.

The condition for fragmentation of heavy ions is that the projectile should move faster than nucleons move inside the nucleus. The projectile energy should be sufficiently above the Fermi domain, for example, above 100 A MeV. The usable target thickness for these high energies is of the order of several grams per square centimeter, corresponding to 10^{23} atoms cm^{-2}. The excitation function for complete fusion of heavy ions, however, has a width of only 10 MeV. This corresponds to a usable target thickness of the order of 1 mg cm^{-2} or 10^{18}atoms cm^{-2}. Consequently, beam intensities for the investigation of complete fusion reactions must be by 4–5 orders higher to achieve the same luminosity as for fragmentation.

Figure 2.22 shows an example of the production cross sections for tin isotopes from complete fusion (dotted line), nuclear fragmentation (solid line), and Coulomb fission of ^{238}U (dashed line). The symbols represent experimental data. The fragmentation cross sections (solid line) have been calculated with a semiempirical code [Sue00].

The first experiments with unstable nuclear beams were designed to measure the nuclear sizes, namely, the matter distribution of protons and neutrons. For stable nuclei, such experiments are best accomplished with electron beams, which probe the nuclear charge (proton) distribution. Electron scattering experiments with unstable beams can only be performed in an electron–nucleus collider. However, the easiest solution is to measure the *reaction cross section* in collisions of unstable beams with a fixed target nucleus [38].

The reaction cross section in high-energy collisions is given by

$$\sigma_R = 2\pi \int [1 - T(b)] b db \qquad (2.89)$$

where

$$T(b) = \exp\left\{-\sigma_{\mathrm{NN}} \int_{-\infty}^{\infty} \mathrm{d}z \times \int \rho_P(\mathbf{r})\rho_T(\mathbf{R}+\mathbf{r})\,\mathrm{d}^3 r\right\} \qquad (2.90)$$

with $\mathbf{R} = (\mathbf{b}, z)$. σ_{NN} is the nucleon–nucleon cross section at the corresponding bombarding energy,

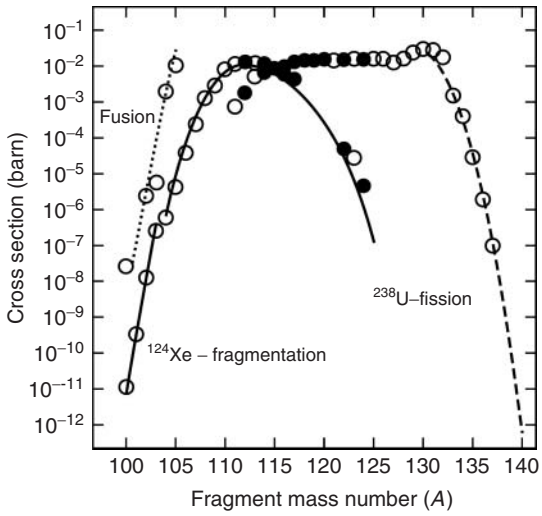


Figure 2.22 Production cross sections for the tin isotopes from complete fusion (dotted line), fragmentation (solid line), and projectile fission of ^{238}U (dashed line). The symbols represent experimental data. The fragmentation cross sections (solid line) have been calculated with a semiempirical code.

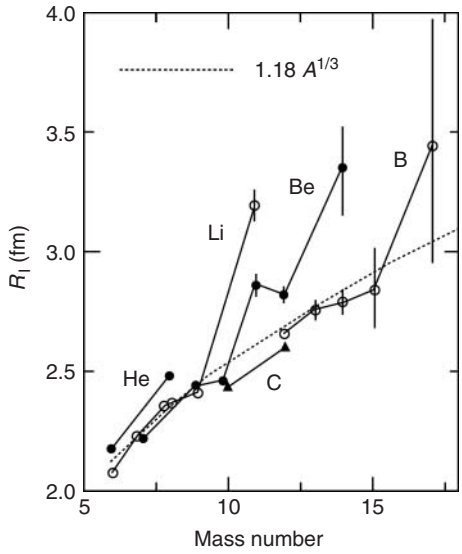


Figure 2.23 The matter density radii of several light nuclei compared to the trend $R \sim 1.18\ A^{1/3}$ fm (dashed line) for normal nuclei. The solid lines are guides to the eyes.

and $\rho_{P(T)}$ is the projectile (target) matter density distribution. $T(b)$ is known as the *transparency function* or *survival probability*. The factor $1 - T(b)$ is the probability that a reaction occurs for a collision with impact parameter b. The factor $\left[\sigma_{\mathrm{NN}} \int \rho_P \rho_T \mathrm{d}^3 r\right]^{-1}$ is interpreted as the mean free path for a nucleon–nucleon collision. The reaction cross sections can be calculated using the matter distributions of the target in Eq. (2.16).

Experimentally, a significant increase in the root mean square radii for the neutron-rich isotopes ^{6}He,^{8}He, and ^{11}Li (Figure 2.23) was observed. Thus, the addition of the neutrons to ^{4}He and ^{9}Li nuclei increase their radii considerably. This can be understood in terms of the binding energy of the outer nucleons. The large matter radii of these nuclei have lead the experimentalists to call them *''halo nuclei.''* The binding energy of the last two neutrons in ^{11}Li is about 300 keV. In ^{6}He, it is 0.97 MeV. These are very small values and should be compared with $S_n =$ 6–8 MeV, which is the average binding of nucleons in stable nuclei. Abnormally large radii were also found for other light neutron-rich nuclei.

The matter density radii of these nuclei do not follow the observed trend $R \sim 1.18\ A^{1/3}$ fm of normal nuclei. Thus, the halo seems to be a common feature of loosely bound neutron-rich nuclei.

Several other methods have been devised to probe the structure of nuclei far from the valley of stability. Among these are (i) Coulomb dissociation [39], (ii) Trojan horse method [40], (iii) asymptotic normalization coefficients [41], (iv) heavy ion charge exchange [42], (v) knockout [43], and (vi) fusion reactions [44]. These methods yield different insights into the structure of exotic nuclei and comprise most of the experiments in radioactive beam facilities [45].

Reactions producing rare nuclear isotopes have opened a new research front in nuclear physics with applications in many areas of science: the rapid production of short-lived nuclei, many of which are poorly known, is important for astrophysics (r-process) and cosmology; nuclear

medicine benefits from studies of new nuclear isotopes; and the list goes on.

However, the basic question still remains: what combinations of neutrons and protons can make up a nucleus? The experimental detection of new nuclear isotopes is an ongoing research that will ultimately lead to new insight and development of nuclear science, with enormous profit for mankind [46].

Acknowledgments

The author is grateful to Prof. Reinhard Stock and Prof. Konrad Gelbke for their support and encouragement. This work was supported in part by funds provided by the US Department of Energy (DOE) under contract No. DE-FG02-08ER41533.

References

1. Rutherford, E. (1911) *Philos. Mag.*, **21**, 669
2. Ridley, B.W. and Turner, J.F. (1964) *Nucl. Phys.*, **58**, 497.
3. Bohr, N. (1936) *Nature*, **137**, 344
4. Ghoshal, S.N. (1950) *Phys. Rev.*, **80**, 939
5. Facchini, U. and Saetta-Menichella, E. (1968) *Energ. Nucl.*, **15**, 54
6. Feshbach, H. (1992) *Theoretical Nuclear Physics - Nuclear Reactions*, John Wiley & Sons, Inc., New York.
7. Tveter, A. (1972) *Nucl. Phys.*, **A185**, 433
8. Marshak, H., Langsford, A., Wong, C.Y., and Tamura, T. (1968) *Phys. Rev. Lett.*, **20**, 554
9. Weisskopf, V.F. and Ewing, D.H. (1940) *Phys. Rev.*, **54**, 472
10. Hauser, W. and Feshbach, H. (1952) *Phys. Rev.*, **87**, 366
11. McVoy, K.W. (1967) *Ann. Phys.*, **43**, 91
12. Feshbach, H., Porter, C.E., and Weisskopf, V.F. (1953) *Phys. Rev.*, **90**, 166
13. Perey, F.G. (1963) *Phys. Rev.*, **131**, 745
14. Miura, K. *et al.* (1987) *Nucl Phys.*, **A467**, 79
15. McMillan, E. and Abelson, P.H. (1940) *Phys. Rev.*, **57**, 1185
16. Oganessian, Y.T. *et al.* (2006) *Phys. Rev.*, **C 74**, 044602
17. Bertulani, C.A. and Baur, G. (1986) *Phys. Lett.*, **B174**, 23; (1988) *Phys. Rep.*, **163**, 299
18. Leprêtre, A. *et al.* (1974) *Nucl. Phys.*, **A219**, 39
19. Axel, P. (1962) *Phys. Rev.*, **126**, 671
20. Mordechai, S. *et al.* (1988) *Phys. Rev. Lett.*, **61**, 531.
21. Schmidt, R. *et al.* (1993) *Phys. Rev. Lett.*, **70**, 1767.
22. Frankfurt, L., Strikman, M., and Weiss, C. (2005) *Ann. Rev. Nucl. Part. Sci.*, **55**, 403.
23. Danielewicz, P. (1995) *Phys. Rev.*, **C51**, 716.
24. Wong, C.Y. (1994) *Introduction to High Energy Heavy Ion Collisions*, World Scientific, Singapore.
25. Hwa, R.C. and Wang, X.N. (eds) (2004) *Quark-Gluon Plasma 3*, World Scientific, Singapore.
26. Aggarwal, M.M. *et al.* (2000) *Phys. Rev. Lett.*, **85**, 3595.
27. Bathe, S. *et al.* (2006) *Nucl. Phys.*, **774**, 103.
28. Damjanovic, S. *et al.* (2006) *Nucl. Phys.*, **A774**, 715.
29. Adare, A. *et al.* (2007) *Phys. Rev. Lett.*, **98**, 232302.
30. Buesching, H. *et al.* (2006) *Nucl. Phys.*, **774**, 103.
31. (a) Gonçalves, V.P. and Bertulani, C.A. (2002) *Phys. Rev. C*, **65**, 054905; (b) Adeluyi, A. and Bertulani, C.A. (2012) *Phys. Rev.*, **C 85**, 044904.
32. Bertulani, C.A., Klein, S., and Nystrand, J. (2005) *Ann. Rev. Nucl. Part. Sci.*, **55**, 271.
33. Baur, G. *et al.* (1996) *Phys. Lett.*, **B 368**, 251.
34. Blanford, G. *et al.* (1998) *Phys. Rev. Lett.*, **80**, 3040.
35. Bahcall, J.N. (1989) *Neutrino Astrophysics*, Cambridge University Press, Cambridge.
36. Bethe, H.A. (1939) *Phys. Rev.*, **55**, 434.
37. Fuller, G.M., Fowler, W.A., and Newman, M. (1985) *Ap. J.*, **293**, 1.
38. Tanihata, I. *et al.* (1985) *Phys. Rev. Lett.*, **55**, 2676.
39. Baur, G., Bertulani, C.A., and Rebel, H. (1986) *Nucl. Phys.*, **A458**, 188.
40. Baur, G. (1986) *Phys. Lett.*, **B178**, 135.
41. Xu, H.M., Gagliardi, C.A., Tribble, R.E., Mukhamedzhanov, A.M., and Timofeyuk, N.K. (1994) *Phys. Rev. Lett.*, **73**, 2027.
42. Steiner, M. *et al.* (1996) *Phys. Rev. Lett.*, **76**, 26.
43. Hansen, P.G. and Tostevin, J.A. (2003) *Annu. Rev. Nucl. Part. Sci.*, **53**, 219.
44. Canto, L.F., Gomes, P.R.S., Donangelo, R., and Hussein, M.S. (2006) *Phys. Rep.*, **424**, 1.

45. Geesaman, D.F., Gelbke, C.K., Sherrill, B.M., and Janssens, R.V.F. (2006) *Annu. Rev. Nucl. Part. Sci.*, **56**, 53.
46. Baumann, T. *et al.* (2007) *Nature*, **449**, 1022.

Further Readings

Bertulani, C.A. and Danielewicz, P. (2004) *Introduction to Nuclear Reactions*, IOP Publishing, London.
Bertulani, C.A. (2007) *Nuclear Physics in a Nutshell*, Princeton University Press, Princeton, NJ.
Danielewicz, P., Lacey, R., and Lynch, W.G. (2002) Determination of the equation of state of dense matter. *Science*, **298**, 1592.
Feshbach, H. (1993) *Theoretical Nuclear Physics, Nuclear Reactions*, Wiley-Interscience, New York.
Fröbrich, P. and Lipperheide, R. (1996) *Theory of Nuclear Reactions*, Oxford University Press, Oxford.
Iliadis, C. (2007) *Nuclear Physics of Stars*, John Wiley & Sons, Inc., New York.
Lilley, J.S. (2001) *Nuclear Physics: Principles and Applications*, John Wiley & Sons, Inc., New York.
Glendenning, N.K. (2004) *Direct Nuclear Reactions*, World Scientific, Singapore.
Wong, S.S.M. (2004) *Nuclear Physics*, John Wiley & Sons, Inc., New York.
Wong, C.Y. (1994) *Introduction to High Energy Heavy Ion Collisions*, World Scientific, Singapore.

3
Electrostatic Accelerators

David C. Weisser

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

3.1 Introduction

Electrostatic accelerators opened nuclear physics to extensive experimentation and, after more than half a century, they continue to play an important role. Basic research in nuclear and atomic physics was virtually the sole application of stand-alone electrostatic accelerators until the 1980s, when their uses for applied physics and as injectors to booster accelerators expanded rapidly. They became the powerful tools for the new techniques required for analysis and production in many emerging fields, especially in microelectronics and molecular biology. They were the natural choice for the injectors needed as heavy-ion research reached for higher energies and heavier mass beams. Advances in the designs and construction of electrostatic accelerators have given them the precision, reliability, simplicity, and versatility needed for these broad roles. These unique capabilities, developed during decades of use in basic research, armed them for these new tasks.

The widespread and varied applied uses of electrostatic accelerators include ion implantation in the manufacture of semiconductor devices, age determination of samples thousands and even millions of years old, and nondestructive analysis to characterize computer chips, analyze biomedical specimens, monitor pollution, and locate oil. High-power electron beams are used for enhancing insulation on wire, producing shrink tubing, and sterilizing disposable medical products, and for tunable, microwave lasers.

Nuclear physics using heavy ions blossomed through the exploitation of the wide variety of beams conveniently provided by tandem accelerators. It was natural that, as the research expanded to heavier beams, tandem accelerators were the tools-to-hand for use as injectors into booster accelerators. This has been a major role for the larger tandem accelerators from 1980 to the present. Although strides have been made to replace such accelerators with dedicated positive-ion sources, almost all tandem accelerator injects retain a continuing role. As research interest shifted toward radioactive ion beams, they were also on hand to accelerate rare and radioactive species produced by other accelerators.

An electrostatic accelerator consists of a high-voltage terminal supported by an insulating column, a means of generating the high voltage, a source of charged particles, and an evacuated acceleration tube through which the particles are accelerated. The potential drop from the high-voltage terminal to ground potential, accelerates the charged particles to an

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

energy equal to the charge on the particle times the terminal voltage ($E = qV$).

An electrostatic accelerator that operates above 10^6 V (1 MV) is usually enclosed in a steel pressure vessel containing electrically insulating, high-pressure gas. The most widely used of such gases is SF_6 whose dielectric strength is 2.7 times that of air at 1 atm (101 kPa). The dielectric strength rises approximately linearly with pressure to 7 atm, where SF_6 sustains $\sim$200 kV cm^{-1} fields.

For stable operation at high voltages, the insulators of the accelerator support structure and the acceleration tube must be subdivided by metal electrodes that are connected to a resistive voltage divider for voltage grading from the terminal to ground.

3.2 Types of Electrostatic Accelerators

Electrostatic accelerators are often identified by the type of high-voltage generator they employ. Those that generate high voltage by rectifying AC voltage are generally used for high-intensity beams at energy up to $\sim$5 MeV. For certain neutron-induced reactions, ion implantation to a few hundred kiloelectronvolts, and electron beam applications, the Cockcroft–Walton or symmetric cascade rectifier accelerator may be the most convenient. Very intense positive ion or electron beams of a few megaelectron volts, are most easily met by the insulating core transformer (ICT) and Dynamitron accelerators.

For analytical precision and for voltages above 5 MV, accelerators using electrostatic charging are far superior. These use a mechanical system to carry the charge to the terminal. Each type of accelerator described below has advantages for specific applications, and each is serving needs in the growing range of uses of accelerators.

3.2.1 Cockcroft–Walton Voltage Multiplier Accelerator

This accelerator was the first to be used successfully for nuclear transmutation and gained wide recognition when results were published in 1932. It employs a circuit developed by H. Greinacher in 1920, which drives AC voltage through two stacks of capacitors and rectifiers as illustrated in Figure 3.1.

One terminal of the left-hand capacitor stack is connected to a step-up transformer giving peak voltages of $\pm V$. Voltages at all points along this capacitor stack oscillate over a range of 2 V. Series-connected rectifiers link this capacitor stack to another one, connected at one end to the high-voltage terminal and at the other end to ground. As voltages on the transformer-driven stack oscillate, charge is transferred stepwise through the rectifiers from ground to the terminal. The terminal voltage is equal to 2 V N, where N is the number of stages, but has a voltage ripple superimposed, which increases with increasing current.

The charged particles from the ion (or electron) source are accelerated through the evacuated acceleration tube that is connected to the terminal. The acceleration tube sections may be connected to the power supply stages as in Figure 3.1 or may be equipped with resistors for potential grading from high voltage to ground.

Cockcroft–Walton accelerators, operating in open air near 1 MV, are large and require a very large room to avoid voltage flashover. Figure 3.2 shows an 850 keV accelerator of this type in a room that is many meters in all dimensions.

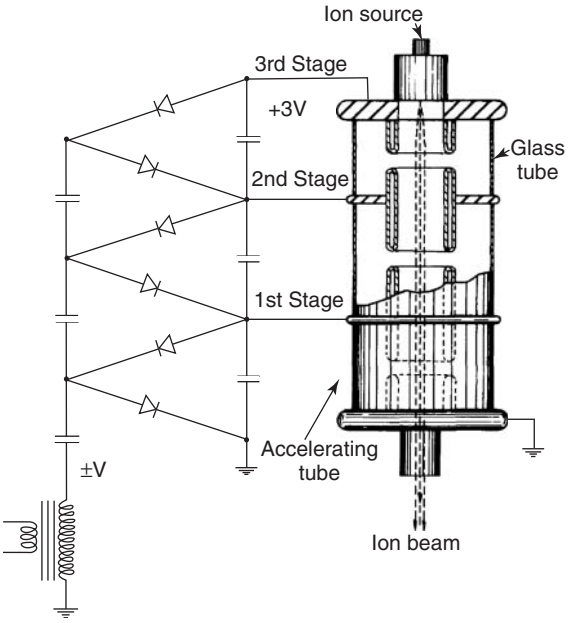


Figure 3.1 Schematic of a Cockcroft–Walton accelerator with three voltage multiplier stages.

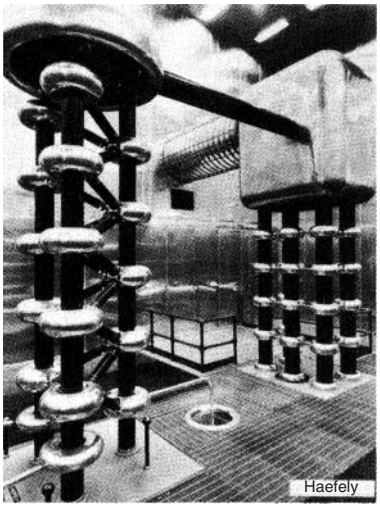


Figure 3.2 The 850 keV Haefely accelerator for injection of pulses of hydrogen ion into a linear accelerator at the Los Alamos National Laboratory. (Photo courtesy of Emile Haefely & Co., Ltd.)

It generates pulsed hydrogen ion beams for injection into the high-intensity linear accelerator at the Los Alamos National Laboratory for production of intense meson beams.

Cockcroft–Walton accelerators are in widespread use throughout the world and are especially advantageous for applications requiring beam currents of a few milliamperes at up to few hundred kilovolts. Many are used to accelerate deuterons to bombard targets containing deuterium or tritium to yield neutrons, for example, for neutron activation analysis.

3.2.2
Symmetric Cascaded Voltage Multiplier Accelerator

G. Reinhold of Haefely has shown that, up to a few hundred kilovolts, voltages developed by cascade voltage multipliers using the circuit of Figure 3.1 do not depart greatly from $2\,V\,N$. Above about 500 kV, however, voltages achieved fall substantially below this value. He developed a symmetrical cascade rectifier, which eliminates most shortcomings of the Cockcroft–Walton circuit. It employs two transformers and two AC capacitor stacks feeding one DC capacitor stack. Using this symmetric system, Haefely built an open-air voltage generator without an acceleration tube that operated up to 2.5 MV.

Such open-air accelerator systems with symmetric circuit charging are used for 500 kV ion implantation and ion beam analysis. Air-insulated accelerators ranging from 20 to over 500 kV are utilized as injectors for almost all megaelectron volt and gigaelectron volt accelerators. Size and space requirements for open-air systems increase rapidly with voltage and appear to set a practical upper limit of about 1.5 MV. Using high-pressure gas for insulation greatly reduces the size of higher voltage accelerators but at the cost of more difficult access to the components for maintenance.

3.2.3
Dynamitron® Accelerator

Cleland invented a cascaded rectifier system called the *Dynamitron*, in which the rectifiers are driven in parallel (Figure 3.3). Rectifiers, connected in series between ground and the high-voltage terminal, are in two rows on opposite sides of the acceleration tube. The high-voltage column is enclosed by half rings with smooth exterior surfaces to inhibit corona and spark discharge. Large, semicylindrical radio-frequency (RF) electrodes, mounted near the wall of the tank, surround the accelerator column. As indicated in Figure 3.3, the half rings on the column provide capacitive coupling between the rectifiers and the large RF electrodes. The RF electrodes form the tuning capacitance of an LC-resonant circuit driven by a separate high-frequency (~100 kHz) power supply.

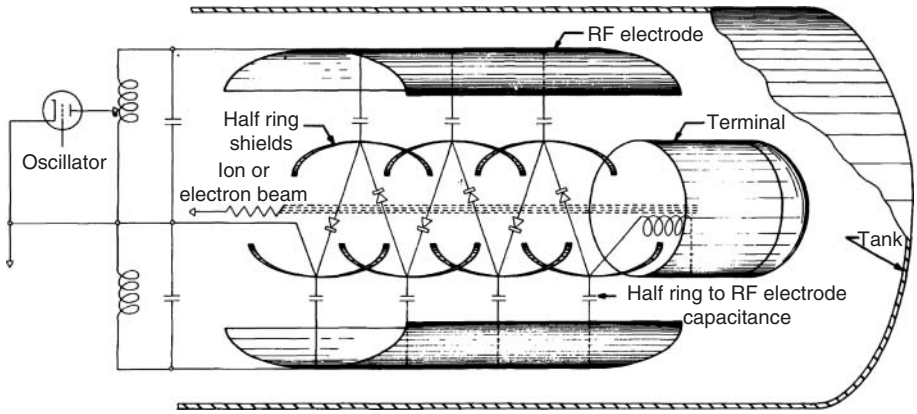


Figure 3.3 Schematic of a Dynamitron accelerator in a pressure tank. The half ring to RF electrode capacitors shown are not physical elements but only represent the effective capacitance of the pairs of electrodes.

Because RF power is fed in parallel to each of the rectifiers, there is no requirement for the large storage capacitors between successive stages as in other cascaded rectifier systems. This reduces stored energy to levels comparable with accelerators using electrostatic charging. Minimizing stored energy is important, especially in multi-megavolt accelerators, because large stored energy released in a discharge can damage capacitors, rectifiers, column components, and acceleration tubes. Single-stage and tandem Dynamitrons up to 4 MV are commercially available for research applications requiring high beam current. Most Dynamitrons are built to accelerate electrons for industrial polymerization and sterilization. These applications require electron beam power ranging from 50 kW at 0.5 MV to 200 kW at 5 MV.

3.2.4
Insulating Core Transformer (ICT) Accelerator

The High Voltage Engineering Corporation, HVEC, developed the ICT circuit for high-current electron beam applications. Power is coupled magnetically in the ICT in contrast to the capacitively coupled devices just described. Rectifiers are on modular, flat, circular decks that are stacked one above the other and are separated by insulation layers. Nominal 460 V, three-phase, 50 or 60 Hz line power drives the primary coil of the power supply transformer. Input power is magnetically coupled through the insulation layers, to secondary coils (three per deck) by a laminated three-phase iron core. The secondary coils on each deck energize the rectifier circuit on that deck. This minimizes electrical stresses within the stack. HVEC builds ICT electron accelerators for energies from 0.3 to 3 MeV and beam power up to 100 kW, primarily for electron beam processing systems.

3.2.5
Van de Graaff Accelerator

Accelerators that use an insulating belt to carry charge to the high-voltage terminal are named for Robert Van de Graaff, who first successfully demonstrated high-voltage generation with belt charging in 1929.

In this device, illustrated in Figure 3.4, electric charge is deposited on the insulating belt at the ground end. The motor-driven belt carries the charge to the spherical metal terminal.

To charge the belt, a corona discharge is maintained from sharp needles or from a fine wire held at high voltage to the grounded lower pulley on the opposite side of the belt. If the corona needles are at a positive voltage, the belt surface intercepts positive ions moving from them toward the grounded pulley. The belt carries the positive charge into the spherical terminal, where another corona discharge from another array of needles removes the charge. The charge then passes to the outer surface of the terminal. The positive voltage on the terminal rises until the charging current is matched by current paths to ground. These may be combination of a resistor string along the insulating support and ion current through an acceleration tube. A negative terminal voltage is produced by operating the charging needles at negative voltage for the acceleration of electrons or negative ions.

The first use of a Van de Graaff voltage generator to accelerate ions was in 1932 at the Department of Terrestrial Magnetism of the Carnegie Institution, Washington, DC. By 1934, this institution had built a second accelerator, which

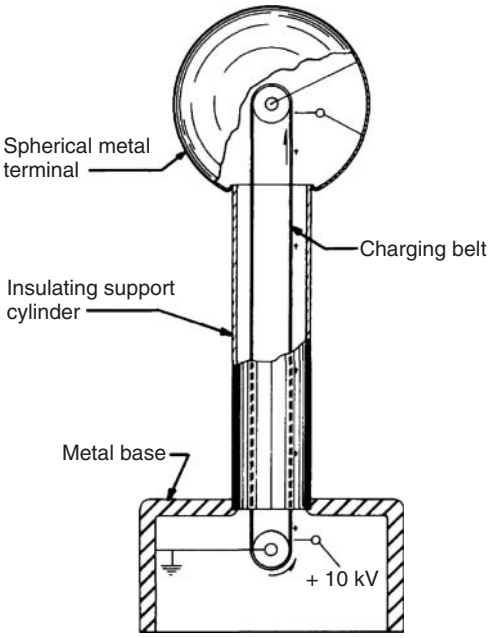


Figure 3.4 Schematic of a Van de Graaff generator for operation in open air.

was used extensively for nuclear physics research at energies up to 1.2 MeV and is now on display at the Smithsonian Institute in Washington, DC. Other belt-type accelerators built for 2–3 MV in open air were very large (up to over 12 m high).

Compact belt-charged accelerators, in steel tanks containing high-pressure insulating gas, were developed at the University of Wisconsin in the 1930s under the direction of Raymond Herb. Many basic design features of electrostatic accelerators, still relied on today, were developed in these early machines. In 1940, Wisconsin completed a 4.5 MeV accelerator, which was used for precise nuclear physics measurements. In 1947, HVEC began commercially manufacturing belt-charged accelerators insulated by high-pressure gas for electrons and positive ions up to 4 MeV. The belt-charging system was the workhorse of HVEC machines with energies up to 6 MV and still remains so for all their single-ended accelerators. However, genuine HVEC replacement belts are no longer available and the adaption of contemporary conveyor belts for the unusual role in Van de Graaff accelerators has been challenging with few successes. These machines do not have enough room for pellet chains or laddertrons.

3.2.6
Two-Stage (Tandem) Accelerators

In 1958, HVEC demonstrated an electrostatic accelerator that produced high-energy beams by accelerating ions twice. In this two-stage or ''tandem'' accelerator, illustrated in Figure 3.5, injected negative ions are accelerated from ground potential to a positively charged high-voltage terminal. In the terminal, a dilute gas-zone gas or thin-foil target strips electrons from the ions. The resulting positively charged ions are accelerated a second time, away from the terminal, back to ground potential.

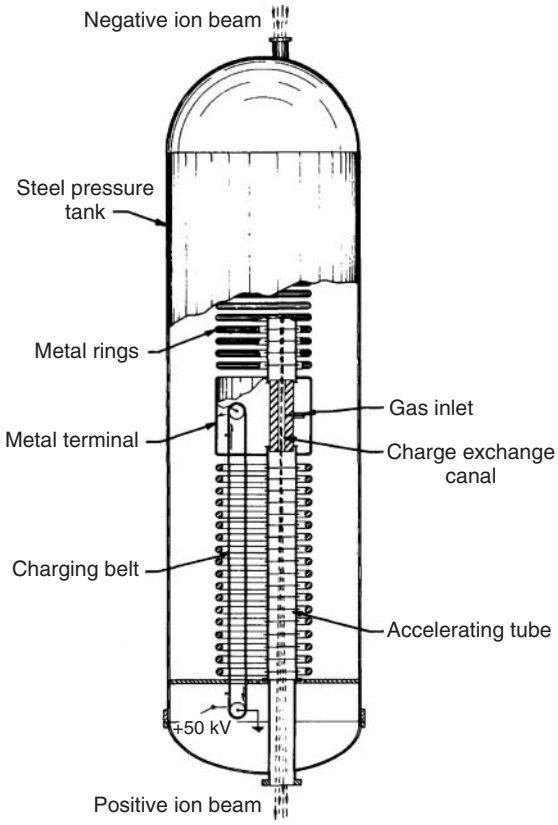


Figure 3.5 Schematic of a two-stage (tandem) electrostatic accelerator in a pressure tank.

Tandems accelerate protons to energy of 2 eV, where e is the electron charge and V is the terminal voltage. Heavier ions, which can be stripped to multiple positive charge states, are accelerated to even higher energies. For example, negative oxygen ions injected into a tandem operating at 10 MV are first accelerated to the positive terminal to 10 MeV. Some of the oxygen ions are stripped of all electrons and emerge from the terminal with a charge of +8e and gain an additional 80 MeV in the second acceleration for a total energy of 90 MeV.

The tandem principle was originally patented in 1935 by Bennett of Ohio State University. HVEC built many tandem accelerators, which are in operation in laboratories around the world. The largest reach terminal voltages of over 16 MV.

3.2.7 Tandetron® Accelerator

In 1980, Kenneth Purser designed the Tandetron line of compact tandem accelerators for materials analysis, accelerator mass spectrometry (AMS), megaelectronvolts ion implantation, and basic research. The high-voltage generator is a Dynamitron-type parallel-driven cascade rectifier circuit with silicon rectifiers and a 50 kHz driver. It delivers several hundred microamperes of beam current at up to 3 MV with high stability. For accessibility, the high-voltage stack is at right

angles to the accelerator column and SF_6, at a pressure of about 5 atm, insulates the accelerator and power supply.

3.2.8 Pelletron® Accelerator

Pelletron accelerators use a charging chain, rather than a belt or rectifier stack, to deliver charge to the terminal in order to generate high voltage [1]. The chain, shown in Figure 3.6, consists of steel cylindrical pellets joined by links of solid insulating material such as nylon. Pellet-charging current is adjustable and highly uniform, so that terminal voltages are easily maintained at very precise values with very little voltage ripple. This high charging stability is important and, in some cases, essential for applied uses.

The pellets are charged inductively so there is no contact with the charging electrodes called *inductors* and no electrical discharge between the pellets and inductors. The pellets carry the charge into the terminal, where the charge is removed at another pulley. The pellets leaving the terminal are then negatively charged so that the up-going and down-going runs of the chain contribute equally to charging. Laddertron charging systems, built by HVEC, are similar except that a pair of pellets is joined by a rung forming a ladder structure. Essentially, all large accelerators now use pellet chains or laddertrons.

The highest voltage electrostatic accelerator, the National Electrostatics Corporation (NEC) 25 MV tandem accelerator, shown in Figure 3.7, is at the Oak Ridge National Laboratory (ORNL) [2]. It is also one of the few "folded" tandems. In the folded design, both acceleration tubes are mounted in a single insulating support column. Ions are accelerated up to the terminal and stripped of the more loosely bound electrons. Then a 180° magnet reverses the beam direction for the second acceleration down to ground. The magnet completely filters out ions of unwanted charge states from the stripper before the second acceleration, so there are no extraneous beams loading the accelerating tube. The folded design, introduced by Henry Naylor [3] in New Zealand, conveniently allows the location of both ion source and experimental equipment at ground level near the system control room. For very large tandems (20 MV and higher), the folded design reduces the size and costs for the building, insulating gas, and pressure vessel.

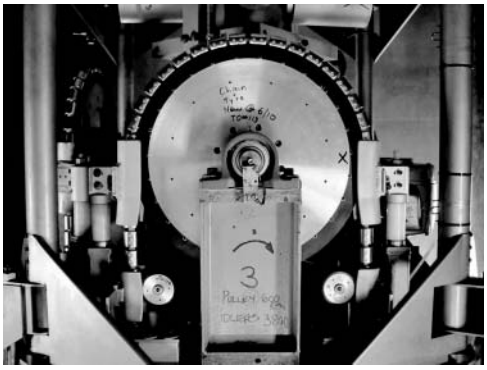


Figure 3.6 Charging chain and terminal pulley in the 14UD Pelletron accelerator with the terminal electrode slipped down for access.

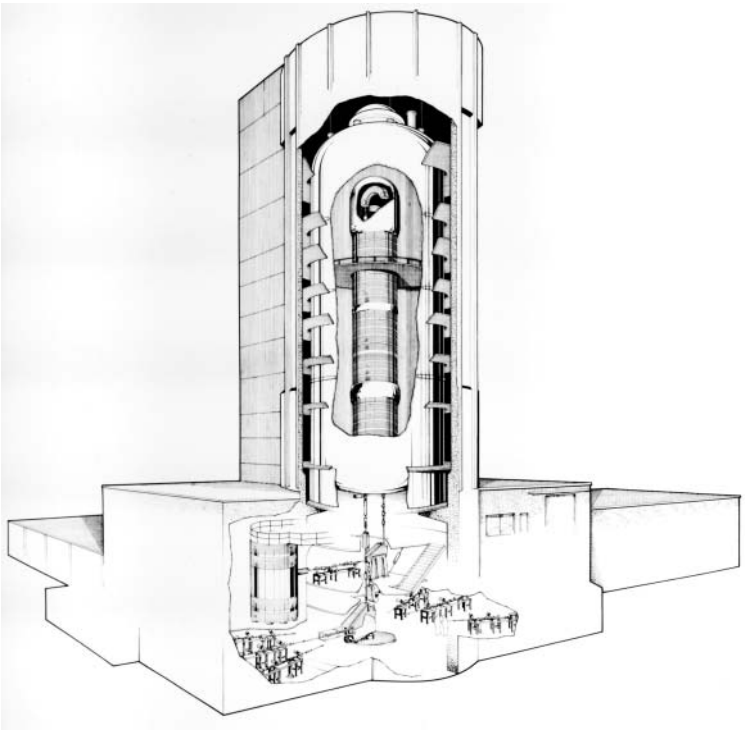


Figure 3.7 Drawing of the 25 MV folded tandem at the Oak Ridge National Laboratory, showing the 180° magnet in the high-voltage terminal. The cylinder (lower left) is a 500 keV ion injector. (Drawing courtesy of National Electrostatics Corporation.)

3.3 Applications of Low-Energy Electrostatic Accelerators

Basic research in nuclear and atomic physics drove the development of electrostatic accelerators from their beginnings in the early 1930s until well into the 1970s. In the 1980s, a broad range of new applications of accelerators rapidly expanded their use. By the end of the 1980s, nearly all new electrostatic accelerators had terminal voltages <5 MV and were designed for applied work. Many existing smaller tandem accelerators, with voltages up to 10 MV that were originally built for nuclear physics research, were converted for use for applied physics. Most of the larger ones, with voltages of 10 MV and above, expanded their role from stand-alone, heavy-ion accelerators to include injecting into other accelerators to provide heavy ion beams with higher energies and heavier masses.

There are three main areas of application of electrostatic accelerators with terminal voltages up to about 5 MV. For applications needing energies of up to several hundred kilovolts, insulated core transformers and open-air symmetric cascade generators are chosen. Higher voltages employ Dynamitrons, Van de Graaffs, and Pelletrons in competition with RF linear accelerators (LINAC).

3.3.1 Material Transformation

The first class of operation is the use of the beam particles themselves to alter the material on which they impinge transforming it in some way.

3.3.1.1 Ion Implantation 12

The major application of low-energy accelerators is the ion implantation into semiconductor material to produce microscopic electronic circuits. Megaelectron volt implanters produce ion beams at energies from a few hundred kiloelectron volts to over 1 MeV. At the higher energies, ions penetrate deeper into the semiconductor wafer. Deeper implants permit higher density circuits and can create conducting or insulating layers below surface circuitry to inhibit effects from stray charges, for example, in computer memory chips. Multiple implants at different energies up to over 1 MeV are used to produce deep complementary metal oxide semiconductor (CMOS) wells, with economic and performance advantages over wells produced by slower thermal diffusion. Megaelectron volt ion implantation in III–V materials such as gallium arsenide produces very high speed switching circuits and light-emitting and light-sensitive solid-state devices. Megaelectron volt implants can program logic circuits late in their production cycle and enhance circuit resistance to RF and radiation interference.

3.3.1.2 Surface Processing

Ion implantation is also used to improve the corrosion and wear resistance of metal surfaces. Implanting nitrogen into low-carbon steel, increases surface hardness for improved resistance to abrasive wear. Implanting titanium into steel reduces the coefficient of friction. Implants of elements such as Cr, Ta, Pd, P, and B into metals can reduce the chemical reactivity of the metals, for example, for longer component lifetime in chemical processing systems.

3.3.1.3 Electron-Beam-Induced Radiation Damage

The radiation damage produced by electron beams is used to increase the mechanical toughness of plastic insulation on electrical cable, to cure thin coatings, improve adhesives, for sterilization of single-item medical disposables, for crystal modification for permanent color in glass and gems, in the treatment of drinking water, disinfection of municipal sewage, and treatment of SO_2 and oxides of nitrogen in industrial smokestack exhausts.

3.3.2 Materials Analysis with Ion Beams

Materials analysis is a pervasive application of electrostatic accelerators in which the ions are used to interrogate the composition of the material rather than to transform it. These nondestructive techniques use megaelectron volt ion beams to identify elements and measure their concentrations and locations in a wide range of materials.

3.3.2.1 Particle-Induced X-Ray Emission (PIXE)

Particle-induced X-ray emission (PIXE) identifies elements and their relative abundances from the energies of characteristic X-rays emitted by atoms in the sample, which are excited by an incident megaelectron volt ion beam. It is used to measure trace elements in biomedical samples, ceramics, petroleum, polymers, and crystals, and on the surfaces of metals and semiconductor materials. PIXE analysis of air samples is used to monitor

aerosol pollution. With a scanning proton microprobe ion source, PIXE provides high-resolution lateral mapping of trace element distributions in organic tissues, minerals, and so on.

3.3.2.2 Rutherford Backscattering (RBS) Analysis

Rutherford backscattering (RBS) analysis measures depth and density profiles of elements at and near the surface of samples. Nuclei in the sample elastically scatter incident megaelectron volt helium ions to a backward angle where their energies are measured.

Rutherford scattering is simply elastic scattering due to the electrostatic repulsion of the positively charged incident and target nuclei. The energy of scattered particles and the counting rate can be calculated from first principles. For a given incident beam, the energy of the scattered ions increases with increasing target mass. The cross section, the probability of scattering, is proportional to the square of the target atomic number (Z) divided by the incident energy. From the calculated cross section and the measured incident beam current, RBS can measure elemental concentrations directly so calibration standards are not required. RBS is often used to measure concentrations of known target elements, for example, in surface layers created by vapor deposition or ion implantation.

The fast, simple, quantitative, and nondestructive nature of RBS analysis has opened many applications. RBS systems typically use small tandem accelerators, which produce microampere helium beams over a wide range of selectable energies. Lower energy beams increase depth resolution in thin films. Higher energies probe deeper and improve mass resolution for heavy elements.

3.3.2.3 Channeling

Channeling, an effect seen in RBS analysis of crystals, decreases backscatter yields by as much as 98% when the beam is aligned with a crystal axis. Ions incident parallel to the axis (within about 1°) are steered by correlated small-angle deflections along the axis, preventing their close approach to crystal lattice atoms. As close approach is required for large angle scattering only, lattice imperfections and impurities in the crystal will result in significant large angle yield.

Channeling is used, especially in the semiconductor industry, to measure crystal perfection, the amount and depth of lattice disorder, locations of substitutional and interstitial impurities, and lattice damage from ion implantation in silicon and subsequent crystal regrowth after annealing. Surface defects and amorphous surface layers increase the angular divergence of the beam, so that some of the beam cannot be channeled by the crystal below the surface. This effect is used to measure thickness and crystallinity in surface layers. Channeling improves RBS sensitivity to surface impurities by reducing backscatter from the substrate and permits deeper analysis with PIXE and nuclear resonance analysis in crystals.

3.3.2.4 Elastic Recoil Detection (ERD)

Elastic recoil detection (ERD) is similar to RBS analysis but measures the recoiling target ions at forward angles instead of the scattered beam particles. In applications on thick samples, the incident beam and detector are at grazing angles to the sample surface. Thin foil samples are exposed to the beam at normal incidence with the detector at 0°, where recoil cross sections are largest.

ERD is sensitive to light elements in heavy substrates and is used to profile

hydrogen in metals and semiconductors. Although nuclear reaction analysis (NRA) gives better depth resolution, ERD is faster and easy to analyze.

3.3.2.5 Resonance Scattering

Resonances increase the scattering cross section at particular energies and angles. Proton backscattering from carbon, resonant at 1.73 MeV, is used to measure photoresist thickness on wafers coated with a metal, such as gold. Resonance scattering is superior to helium RBS in which the high background of backscatter from the Au obscures the carbon signal. Elemental concentrations can be measured with resonance scattering if the detection system is calibrated using standard targets of known concentrations.

3.3.2.6 Nuclear Reaction Analysis (NRA)

As beam energy increases and the scattering no longer follows the Rutherford formula, the use of nuclear reactions becomes possible. The reaction $^{1}H(^{15}N, \alpha\ \gamma)$ is resonant at 6.4 MeV and provides good depth resolution for profiling hydrogen in metals, semiconductors, and glass. Unlike ions, γ-rays do not suffer energy loss in the material. Using channeled beams to induce γ-emitting reactions, NRA locates impurities deep within a crystal. The sensitivity of reactions to specific isotopes is useful for analyses in several fields, including mineralogy and biomedicine.

3.3.2.7 Activation Analysis

Particle-induced nuclear reactions are used to generate trace quantities (10^{-10}) of radioactive isotopes at the surface of objects such as gears or inside a pipe. Decreases in radioactivity in excess of natural decay measure the rate of surface wear or erosion.

Neutron activation analysis is used to locate underground oil deposits. Accelerators to generate the neutrons operate at about 100 kV and are small in diameter (<45 mm) to fit into drill holes. Neutrons are also used to detect explosives in airplane cargo.

3.3.2.8 Ion Beam Microprobe

Typical megaelectron volt ion beams are about 2–4 mm in diameter. While suitable for many applications, this is often too large for measuring lateral distributions of elements in important samples. The microprobe produces a beam spot of from 10 to 20 μm down to <0.3 μm diameter for measurements with spatial resolution of a few micrometers or better. It uses energy-analyzed beams from an accelerator and includes a beam collimator, strong focusing lenses, target chamber with sample holder, detectors, and a means of scanning the beam across the sample. Microprobes mainly use megaelectron volt proton beams and PIXE analysis, primarily for biomedical studies but also for archaeology, mechanical engineering, mineralogy, and solid-state studies.

3.3.3 Accelerator Mass Spectrometry (AMS)

AMS measures isotopic ratios down to 10^{-16} with high precision using only milligram and smaller samples. Naturally occurring radioisotope concentrations are measured to date art and archaeological objects, ground water, lunar samples, and meteors and to study glaciers, sediment formation, and ocean mixing. AMS of tracer isotopes is used in hydrological and atmospheric studies.

The samples of milligram size, analyzed by AMS are not in a target chamber, but in the ion source, which creates ion beams from the sample. Carbon dating,

the primary use of AMS, determines the age of samples from the concentration of ^{14}C, whose half-life is 5.73 ka (a = year). The beam from the ion source contains ^{12}C, ^{13}C, and ^{14}C ions in the same relative abundances as in the sample.

Tandem accelerators are best suited to most AMS applications, because they can transport all the isotopes without intermediate adjustments and accelerate them to high energy for high discrimination against contaminant beams. Molecular isobaric interferences, a common problem for other mass spectrometers, are eliminated in tandems, because the stripper dissociates molecules, and the fragments are easily resolved by the detector systems. The use of negative ions, required for tandems, fortuitously eliminates potential interference from ^{14}N (much more abundant than ^{14}C) in carbon dating, because negative N ions are not stable and thus cannot be created. For milligram samples, ages up to 5 ka require <1 h for 0.3% accuracy and ages to 50 ka can be measured with reasonable measurement times.

In the past decade, significant progress has been made in developing and improving AMS with small and compact tandem accelerators. In many cases, these dedicated and optimized systems outperform AMS facilities running at higher terminal voltages. The developments at ETH Zurich, Switzerland, showed that ^{14}C can be competently measured at tandem accelerators running as low as 0.2 MV. This is because high transmission (>40%) in charge state 1+ and adequate molecular dissociation can be obtained at achievable stripper gas pressures [4]. This allowed the building of a compact radiocarbon dating system (MICADAS, Figure 3.8) with measurement on a regular basis of precision in

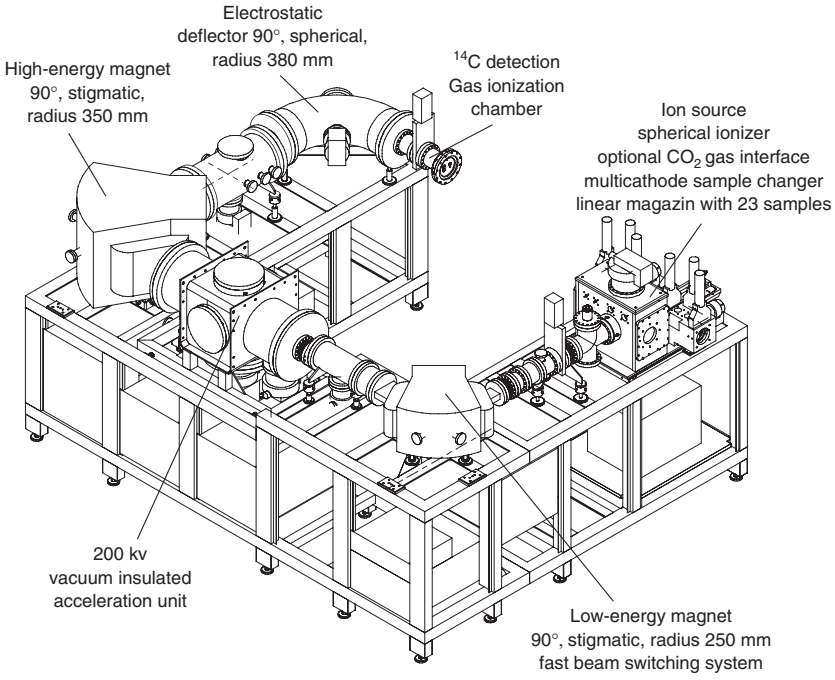


Figure 3.8 The ETH Zurich MICADAS system based on a 200 kV tandem accelerator.

the range of 0.1–0.2%. The voltage insulation of the tandem terminal is achieved by high vacuum instead of a high pressure gas. Thus, there is no need to separate the vacuum space, which the beam traverses, from the insulating region and so no need for an acceleration tube.

In addition to radiocarbon AMS, tandem accelerators with terminal voltages around 0.5 MV have become alternatives for AMS of many radionuclides where the separation from isobaric interference at the high-energy side of the facility is not required. The degrader method allows the separation of ^{10}Be–^{10}B, which even allows AMS of ^{10}Be at 0.5 MV, comparable to facilities running at higher terminal voltages [4]. With an adequate spectrometer at the high-energy end, the suppression of neighboring isotopes is adequate to perform AMS up to the actinides (e.g., ^{236}U, Pu isotopes). Furthermore, with the use of helium as a stripper gas, high transmission (>40%) through the tandem accelerator can be achieved [5]. This is a clear advantage over AMS at larger tandems, which requires the selection of a higher charge state than the most populated one because of the limitations set by the bending magnets and/or electrostatic analyzers (ESA)s, resulting in transmissions of well below 10%. Important AMS radionuclides such as ^{10}Be, ^{26}Al, ^{41}Ca, ^{129}I, ^{236}U, and Pu can be measured with less effort at these small and compact facilities, allowing studies where hundreds of samples (e.g., for biomedical or environmental studies) are involved.

As adequate molecular dissociation is practical at 0.2 MV, NEC has even eliminated the tandem accelerator itself. In their single-stage accelerator mass spectrometer introduced in 2005 (Figure 3.9), negative

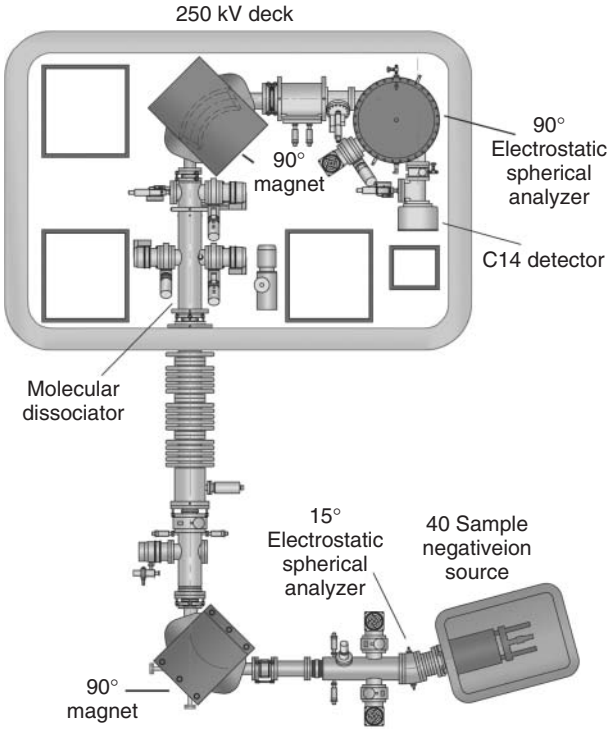


Figure 3.9 NEC single-stage accelerator mass spectrometer. (Courtesy of National Electrostatics Corporation.)

ions are produced in the usual way, from a multisample negative-ion source by sputtering with Cesium (Cs)-positive ions [1]. The negative ions emerge with 44 keV energy and then are energy analyzed in a 15° electrostatic deflector, followed by mass analysis using a 90° magnet. They are accelerated by an additional 250 kV for molecular dissociation in a gas stripper. The +1 ions are then again magnetically and electrostatically analyzed as in conventional AMS. Precision of better than 0.5% has been demonstrated. As of early 2012, about 20 such AMS facilities are in operation worldwide, running at 200–250 kV acceleration voltages.

With the advent of the kilovolt terminal machines, there has been a rekindling of interest in the stripping of negative to positive ions with a variety of gasses. The enhanced yield of low-charge-state heavy ions using helium stripper gas is promising. The stripping to high-charge states of heavy ions in higher-energy machines is benefiting as well [6, 7]. Further developments are in progress to extend radiocarbon measurements to only the source energy of 45 keV, eliminating the accelerator totally [8].

AMS employs several isotopes for exposure age dating, that is measuring the buildup of the radioactive isotope in a rock surface that has been bombarded by cosmic-rays. Isotopes commonly used are ^{10}Be (1.6 Ma half-life), ^{41}Ca (130 ka), and ^{36}C1 (310 ka). ^{36}Cl is also used for the dating of groundwater and ^{26}A1 (720 ka) for cosmology and medial physics. The isotope ^{129}I (16 Ma) is used to study petroleum deposit formation and tracing discharges from nuclear plants. The take-up of calcium in animals is studied using ^{41}Ca.

AMS also measures traces of stable isotopes and their distribution in geological samples. Concentrations below 10^{-10} are mapped by using a scanning microion beam sputter ion source to bombard the sample and a tandem AMS system to detect the emerging negative ions.

3.3.4 Simultaneous Bombardment by Multiple Accelerators

Materials used in nuclear reactors are exposed to intense irradiation by neutrons and α-particles. Historically, the norm has been the performance of sequential simulation of reactor exposure followed by materials analysis.. However, it has been recognized recently that relaxation processes between the irradiations may mask phenomena that could occur during simultaneous exposure. The JANNuS facility at Saclay, France [9], brings together nuclear recoil by neutrons simulated by heavy-ion bombardment and the implantation of helium ions. Figure 3.10 shows the layout of the multiple accelerator facility. The heavy ions are provided by an NEC single-ended 3 MV Pelletron equipped with an electron cyclotron resonance (ECR) ion source. The helium implantation is done at present with a 2.5 MV HVEC van de Graaff Yvette, which is to be replaced by an NEC 2.5 MV single-ended Pelletron in 2014. The consequences for material behavior are probed in the same experimental chamber. Ion beam analysis NRA, elastic recoil, and RBS are employed. Materials analysis beams are provided by beams from a 2 MV Pelletron.

3.4 High-Energy Electrostatic Accelerators

Electrostatic accelerators, with terminal voltage at and above 10 MV, were initially bought for nuclear physics research first with light ions such as the hydrogen and

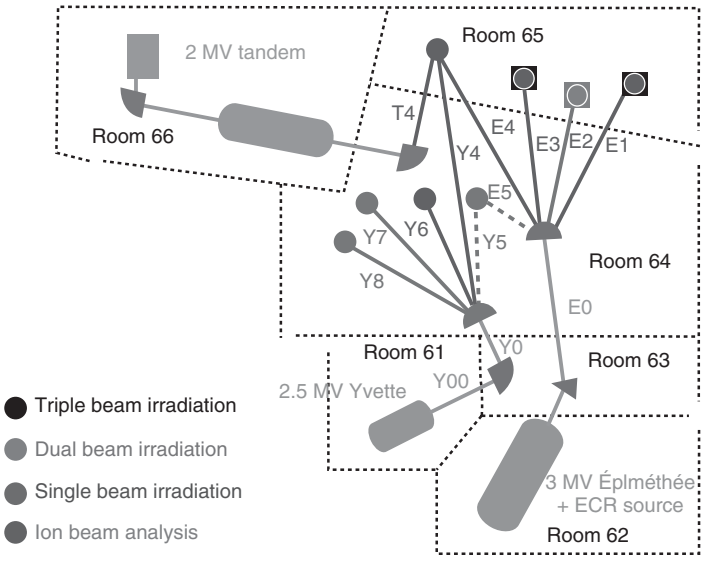


Figure 3.10 Layout of the multiaccelerator facility JANNuS at Saclay that brings implantation and analysis beams together in one experimental chamber in room 65.

helium isotopes, but later with heavier ions. The nuclear physics accessible with these machines was limited to ions lighter than about Ni. As this area of research was exploited, the push to heavier ions drove the expansion of the mass range of beams by the addition of booster accelerators. One might have thought that this relegated the tandems solely to the injector role. However, the tandems continued to dominate the research done even in laboratories with boosters because of their ease of operation compared to boosters and the ingenuity of the physics research staff. The productivity of these laboratories continues to earn a substantial place for them at the forefront of research in nuclear physics.

3.4.1 Injectors in Nuclear Physics Research

Electrostatic accelerators became the injectors of choice for facilities requiring higher-energy heavy-ion beams. The factors in their favor are they were already the heavy ion workhorses present in the research laboratories; they are easy to operate and reliable; they produce compact beams with the good energy resolution of $\sim$2 parts in 10^4; and changing beam species is rapid because the ion source is accessible at ground potential.

Electrostatic accelerators dominated over positive ion injectors in the injector role for heavy-ion research in spite of having much lower beam intensity and the nonexistence of negative noble gas ions other than helium. In laboratories with strong backgrounds in cyclotron technology such as those in Michigan State University (MSU) [10], South Africa [11], and Ganil [12], cyclotrons are the chosen injectors.

The DC beam from an electrostatic injector is not directly compatible with the booster's pulsed acceleration process. The two types of boosters, LINACs

and cyclotrons, both are driven with megahertz RF. Thus, they accelerate best over a limited number of degrees of an RF cycle – about 5° for a 150 MHz LINAC. Therefore, the DC beam must be bunched for injection – a process that reduces the available beam intensity.

There are several factors that limit the beam intensity from electrostatic tandem injectors compared to positive ion injectors: the negative ion current is one to two orders of magnitude smaller than that from a positive ion source; the fractionation of that beam, due to the stripping process from negative ions to positive ones in the terminal, costs a factor of about 4–7; and the loss due to bunching the DC beam into tight time bunches to match the booster RF phase acceptance. Bunching systems capture from 25 to 75% of the DC beam. Another significant drawback of tandem accelerators as injectors is the limited lifetime of the carbon foils used to strip electrons from the negative ions in the terminal. The downtime associated with opening the accelerator pressure vessel to replace the used foils, limits facility availability.

However, it is only from the mid-1990s that positive ion injectors have started to compete with tandems in existing laboratories. Argonne National Laboratory, ANL [13], has used positive ion injectors routinely. Even in 2012, the FN persists as a credible injector. Tandem injectors still coexist with positive ion sources at Legnaro [14]. The new-generation accelerators use LINAC first stages rather than electrostatic machines (e.g., those at TRIUMF [15] and MSU [10]).

3.4.2 Electrostatic Accelerator Injectors

The existing population of electrostatic accelerators determined the choice of injectors pressed into service for the expansion of nuclear physics research. Although many were manufactured by HVEC, they were modified to include Pelletron or laddertron charging systems and Dowlish accelerating tubes.

The 8UD Pelletron in Sao Paulo, Brazil [16], is the only injector with terminal voltage up to 8 MV that presently intends to service a LINAC. The next rank of injectors comprised the HVEC, 8–10 MV terminal FNs servicing superconducting LINACs at ANL [13] and at Florida State University FSU [17]. The LINACs have been shut down at FN laboratories at the State University of New York at Stonybrook [18], the University of Washington at Seattle [19], and Centre d'Etudes Nucleaires, Saclay, France [9].

The higher-energy electrostatic accelerators were generally used in larger laboratories and were adapted to inject a wider variety of boosters. Although the nuclear physics research of the combined accelerator pair was still targeted at beams at 5–10 MeV A^{-1}, where A is the projectile mass in atomic mass units, the range of masses was biased to heavier beams in these larger facilities.

The HVEC MPs and XTUs are capable of 12–16 MV on terminal and could deliver beams with β ranging from 0.06 to 0.08, where β is the ratio of the beam velocity to the velocity of light. The boosters would then further accelerate the beam to β of 0.1–0.15. This was accomplished with superconducting LINACs at Istituto Nazionale di Fisica Nucleare Laboratori Nazionali, Legnaro, LNL, Italy [14] and a superconducting cyclotron at Catania, Italy [20]. Chalk River National Laboratory, Canada [21] had an MP–superconducting cyclotron combination that has now been shut down. A room-temperature LINAC, injected by an MP at MPI, Heidelberg [22],

was adapted to inject a heavy ion storage ring.

3.4.2.1 LNS Catania Injector and Radioactive Ion Beam Accelerators

The Catania facility was reconfigured as shown in Figure 3.11. The cyclotron, using spallation reactions, produces radioactive species that are then accelerated in the electrostatic accelerator [20].

3.4.2.2 ORNL Injector and Radioactive Ion Beam Accelerator

The large NEC Pelletron accelerators were newer than the HVEC cohort and so more of them were adapted as injectors. The largest such facility, the 25URC at ORNL is capable of more than 25 MV on terminal [2]. Figure 3.7 shows the 25URC in stand-alone mode.

The 25URC was always intended to inject beam into the Oak Ridge Isochronous Cyclotron, ORIC, as shown in Figure 3.12. In this floor plan, the tandem is vertically above the partial circle at the lower left, and injects beam into ORIC on its axis. The first coupled operation beam was 400 MeV ^{16}O followed with 28 different ion species, ranging from Li to Pb. The portion of beam time in coupled operation ranged between 35 and 45%, went down to 25–30% during the years 1986–1989 then to 16.5% in 1990 when coupled operation ceased.

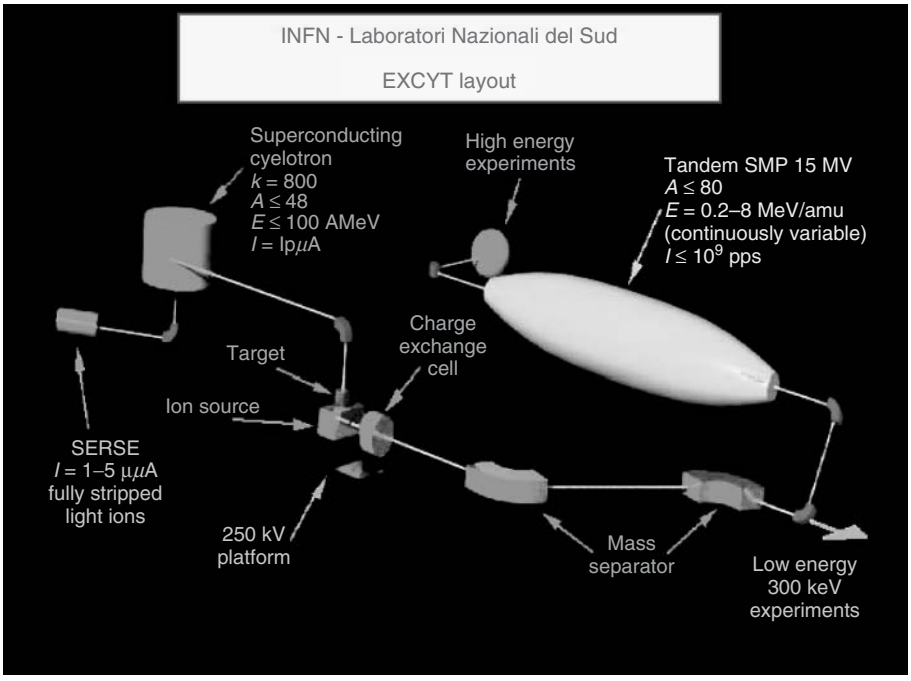


Figure 3.11 The radioactive ion beam facility at LNS, Catania. The cyclotron beam produces radionuclide in a target ion source where they are converted to negative ions for injection into the tandem for acceleration. (Courtesy of the INFN LNS.)

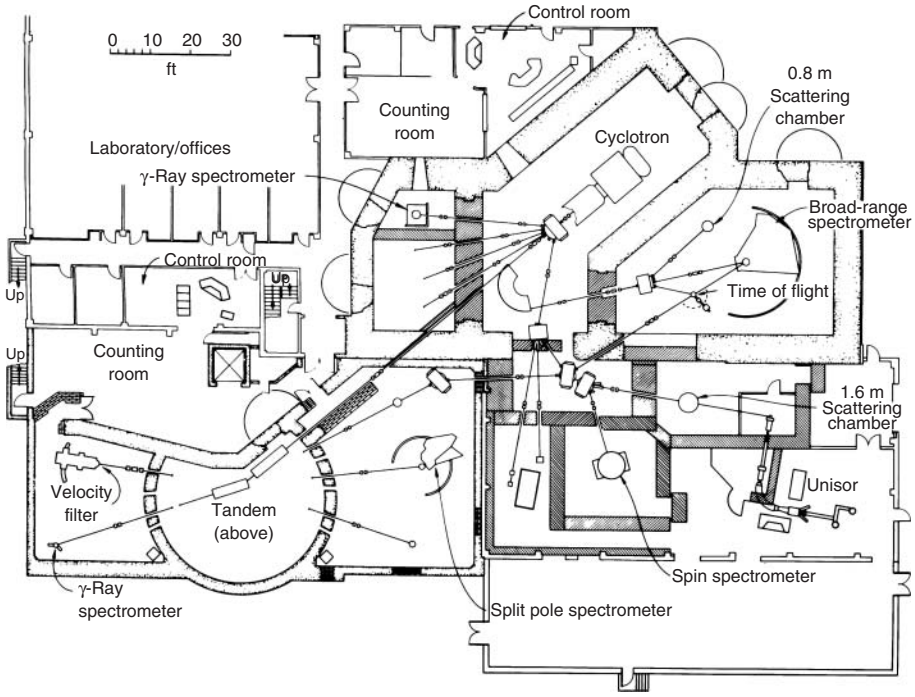


Figure 3.12 Layout of the ORNL 25URC injecting ORIC along its central axis. The tandem is vertically above the partial circle in the lower left. (Drawing courtesy of Oak Ridge National Laboratory.)

The fashion in physics moved from boosted beams of stable isotopes to radioactive beams and so the ORNL combination of accelerators was reversed as shown in Figure 3.13. Here again the tandem is above the partial circle on the left but now, ORIC delivers its proton beam to one of two target/ion sources in the heavily shielded rooms. There, the radioactive atoms are transformed to negative ions and are then mass analyzed for injection into the tandem for acceleration.

Although the Holifield radioactive facility provided a substantial contribution to radioactive ion beam research, the use of the cyclotron to create radioactive elements for further acceleration in the 25URC has been discontinued.

The accelerator combinations at Catania [20] and Oak Ridge are in the first generation of radioactive ion beam facilities defined as those that exploit existing infrastructure and provide beams at and below the Coulomb barrier for light and medium mass ions. Research over the full range of masses and energies will be the domain of purpose-designed facilities, which will not include electrostatic accelerators.

3.4.2.3 Brookhaven National Laboratory – Dual MP Injectors

One of the most successful examples of electrostatic accelerators performing the injector role is that of the MPs at Brookhaven National Laboratory, BNL [23]. It is commonplace in high-energy

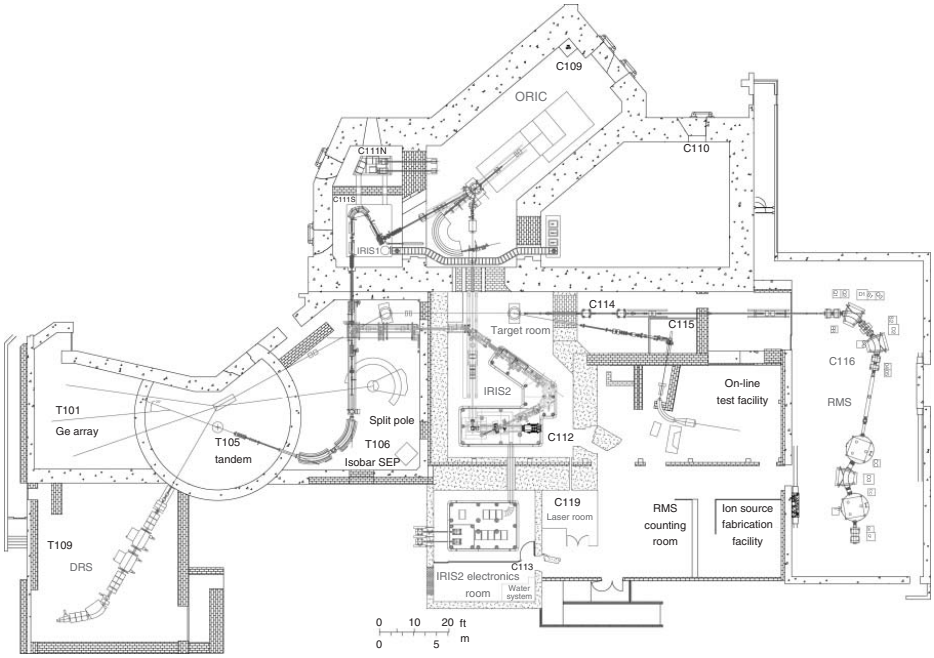


Figure 3.13 The ORNL radioactive ion accelerator system with ORIC bombarding a target in one of two well-shielded rooms each containing a negative ion source. The ions are then accelerated in the tandem. (Drawing courtesy of ORNL.)

laboratories, for a sequence of accelerators to be used to produce the final beam and for newer, larger accelerators to stand on the shoulders of their predecessors. This is also the path taken at BNL, in the journey toward the relativistic heavy ion collider (RHIC).

There already existed at BNL, a historic, high-energy physics program using the alternating gradient synchrotron (AGS), which provided multi-gigaelectronvolts proton beams. BNL also had a vigorous nuclear physics program based on two MP tandem accelerators. These accelerators were well equipped for the injector task since they had the newest style of acceleration tubes for stable operation at 14 MV, Pelletron charging systems for good energy stability, modern resistor grading, and high-intensity negative ion sources.

The AGS evolved into a heavy ion accelerator in a proof-of-principle exercise through being injected by one of the MP tandem accelerators for fixed target experiments. For head-on collisions envisioned for RHIC, the center of mass energy, that is the energy available for reactions, would be more than doubled. The higher energy is required for the shift to research on quark–gluon plasma. This physics demanded very energetic heavy ion beams that could be produced by using the AGS to populate counter rotating gigaelectronvolt heavy ion beams in a collider ring, RHIC. The layout of the BNL accelerator facility is shown in Figure 3.14.

Electron Stripping and Energy Gain The stripping of electrons from ion beams is an essential process in all heavy ion facilities and will be illustrated here using the BNL arrangement. The factor that determines the fate of a heavy ion, between the ion source and the target, is its interactions with matter. These collisions determine its charge state and its charge state determines how much energy it can gain in the electric fields of the accelerators.

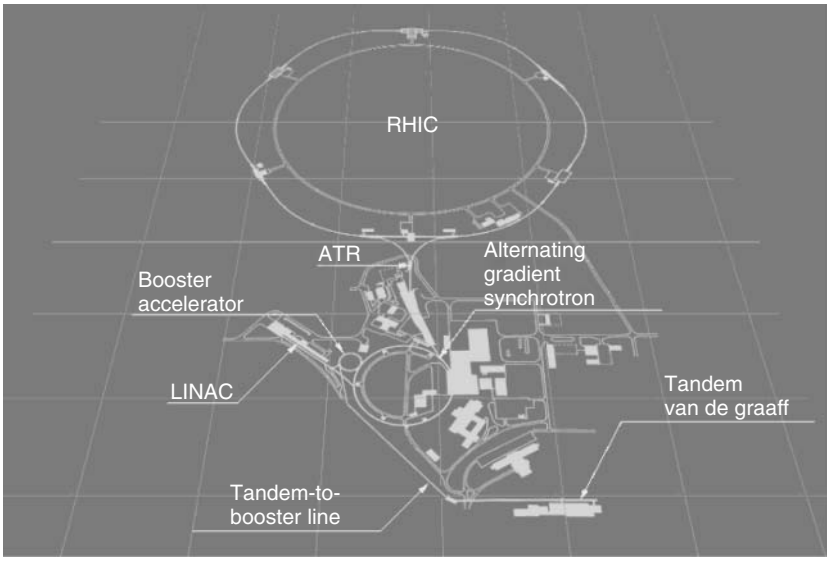


Figure 3.14 The RHIC complex showing the tandem Van de Graaff injection via the booster and the AGS. (Illustration courtesy of Brookhaven National Laboratory.)

At the ion source, an extra electron is attached to the neutral atom, often in the sputtering process using Cs ions, at energies of a few kiloelectronvolts. The velocity of the negative ion increases as it is accelerated on its way to the terminal of the tandem accelerator. If it strikes a residual gas molecule during this segment of its journey, it will be neutralized or made positive and so gain no further energy. The survivors reach the terminal with, for example, 14 MeV energy. Here an interaction with matter is imposed since it is necessary to change the charge state from negative to as high a positive charge state as possible. The higher the charge state the higher the energy the ion will gain being accelerated from the positive terminal to ground. The highest charge states are achieved by using a carbon stripper foil of a few micrograms per square centimeter thickness. This stripper foil will unfortunately also scatter the ions causing some of the beam not be transmitted through the rest of the machine. The thinner the foil, the less scattering, but thin foils are so fragile that they are easily broken even before being exposed to the beam. A 2–10 μg cm^{-2} typical carbon stripper foil is used in the tandem terminal. The production of higher charge states is enhanced by high ion velocity. That is why higher energy tandems predominate as injectors especially for heavier ions.

If the ion, in its travels from the tandem, encounters residual gas molecules, more electrons will be stripped from it and it will be eliminated from the beam in the deflecting magnets along the path to the booster. The second stripping is desirable because the booster will add energy in proportion to the charge state of the ions presented to it. Since the number of electrons stripped increases with the ion velocity, the higher energy ions emerging from the tandem, can be stripped to charge states much higher than that achieved in the tandem terminal.

In the case of Au ions from an MP at BNL, operating with 14 MV on the terminal, charge state +12 is chosen resulting in a final beam energy of $14(1 + 12) = 182$ MeV. In Figure 3.15, the stripper in the terminal of each MP is labeled S1. This beam, after tandem acceleration, is further stripped at S2 to +31 for its travel along the ~850 m long transfer line to the booster. If the +31 ion interacts with residual gas in this section, it is less likely to change charge state and be lost to the beam since its charge state is already typical for interactions at this velocity.

As seen in the figure, there are two additional foil strippers. The one after the booster, S3, takes the beam from +31 to +77 to maximize the energy gain and minimize the charge-changing loss in the AGS. For similar reasons, a fourth stripper, S4, is placed before RHIC to fully strip the Au to +79, precluding any further change of charge state.

All these stripping actions impose losses in intensity because only a fraction of the beam incident on a stripper foil will emerge in a particular charge state. In each of the first two stripping events, only about 15% of the incident beam goes to the desired charge states, +12 and +31. The stripping efficiency between the booster and AGS, S3 is much higher, 65%, because +77 Au is helium-like and retains only the two very tightly bound K-shell electrons. Thus $0.15 \times 0.15 \times 0.65 = 0.015$ of the negative ion beam will reach the fourth stripper only if there are no other losses due to apertures or scattering from residual gas in the vacuum system. The last striping, S4, is >99% efficient since, at this highest

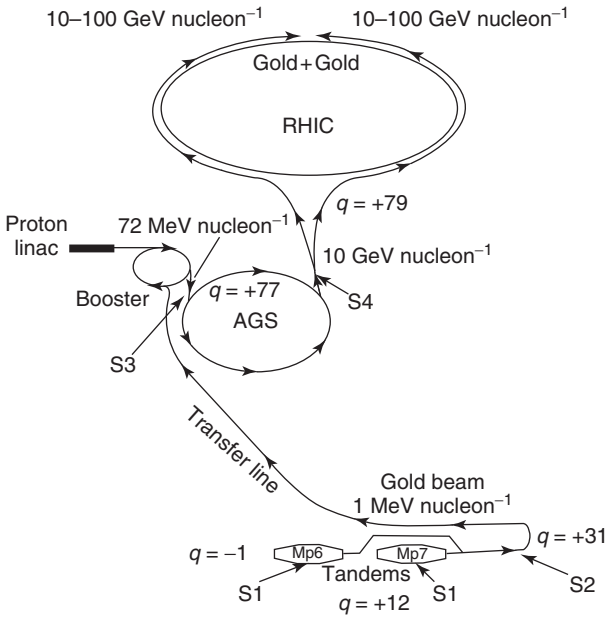


Figure 3.15 The RHIC combination of accelerators starting with the pair of HVEC MPs. The carbon strippers are shown as S1, in the terminals of the tandems; S2, at entry to the transfer line leading to the booster; S3, between the booster and the AGS; and finally S4, at the entrance to the RHIC transfer line. (Illustration courtesy of Brookhaven National Laboratory.)

energy, only the last two electrons need to be removed.

The BNL electrostatic accelerators have provided pulses of Au, Cu, and deuterons for RHIC. The Au ions from either tandem, travel through the tandem-to-booster transfer line to a booster synchrotron, which accelerates the beam to $\beta = 0.37$ for further injection into the AGS. There, the beam is stacked and accelerated to $\beta = 0.997$ in the AGS before being sent to RHIC to circulate in either the clockwise or counterclockwise direction. The countercirculating beams are made to collide at six interaction points where the experiments are mounted.

Having two MPs allows for either to operate while the other is being serviced. Using both machines, collisions can be arranged between different beam species, for example, deuterons from one MP and Au ions from the other. The MPs also provide heavy ion beams via the booster synchrotron to the NASA Space Radiation Laboratory as well as MP-only beams to a large user community.

Now that the ground work for RHIC has been done with the MPs, a purpose-designed positive ion injector has been commissioned. This comprises an electron beam ion source, an RF quadrupole accelerator, followed by a LINAC. This combination has the advantages over the MPs of higher intensity, noble gas beams, and shorter beam pulses to better match the time structure required by the booster synchrotron. If history is any guide, the MPs will continue to provide alternative injection service to RHIC and beam for the NASA Space Radiation Laboratory and stand-alone operation for other users.

3.4.2.4 Argonne National Laboratory – ATLAS

The Argonne tandem LINAC accelerator system (ATLAS) booster, a superconducting LINAC, was first injected by an HVEC

FN electrostatic accelerator. It has been upgraded with NEC accelerator tubes, a Pelletron charging system, and a modern resistor grading system. Here too, the FN was the convenient workhorse to get the heavy ion LINAC program off the ground and continues to inject beam even though two excellent positive ion injectors are available [13].

Figure 3.16 shows the three complimentary injectors at the ATLAS accelerator facility at ANL.

As is the case in BNL, the beam is stripped several times to optimize the energy gain provided by the LINACs. In addition to the stripping in the terminal of the FN from negative to multicharged positive, there is another stripper in front of the booster and another in front of the ATLAS section. The terminal voltage of the FN is only 8.5 MV, which results in lower charge states than is the case for the 14 MV BNL machines and thus limits the FN injection to beams of masses up to Ni.

The booster accelerating devices are superconducting split-ring resonators operating at 48 MHz. The time structure of the booster requires the tandem DC beam first to be bunched into 1 ns, wide pulses. This is accomplished with a four-harmonic gridded buncher, which captures 75% of the DC beam. The beam is then further compressed into 0.1 ns bunches using a superconducting resonator for acceleration in the LINAC. The ANL superconducting resonator technology and beam-pulsing techniques underpin booster projects in Sao Paulo, Brazil [13], in FSU [15], in Japan Atomic Energy Research Institute (JAERI), Tokai, Japan [24], and in the Inter-University Accelerator Centre, IUAC, New Delhi, India [25]. All these facilities use tandem accelerators as injectors.

3.4.2.5 NEC 14–16UD Injectors

NEC accelerators used as injectors are the 14–16 UDs at the Australian National University (ANU), Canberra, Australia, [26], at New Delhi, [25], and the Tata Institute for Fundamental Research, TIFR, Mumbai [27]. At these facilities, NEC tandems inject into superconducting LINACs for a growing part of their research programs. The LINAC that the 20URC at the Japan Atomic Energy Agency (JAEA) [24] used to inject has been shut down.

The large electrostatic tandem accelerator developed at the Nuclear Science Centre in Daresbury, England, mirrored some of the NEC design features. It was intended to have a 20 MV terminal and was adapted to inject a superconducting LINAC but funding was diverted just as the first boosted beams were demonstrated. The LINAC was resurrected at the ANU where it was developed further and the tandem in Daresbury scrapped.

Another family of accelerator facilities, in which electrostatic tandem accelerators are used as injectors, use variants of LINAC accelerating resonators based on a Caltech design of a split loop resonator and/or a quarter wave resonator developed by Ilan Ben Zvi at the Weitzman Institute, Rehovot, Israel [28], in conjunction with the University of Washington, Seattle [19] and the University of the State of New York, Stonybrook, USA [18]. Although a booster facility was not achieved at Rohovot, their contributions to the technology played an important role at several other laboratories including Legnaro [14], Stonybrook [18], Seattle [19], the ANU Canberra [26], and at TIFR [27].

Figure 3.17 shows the layout of the facility in Canberra. It is typical of such facilities that the arrangement of beam transport from the existing electrostatic

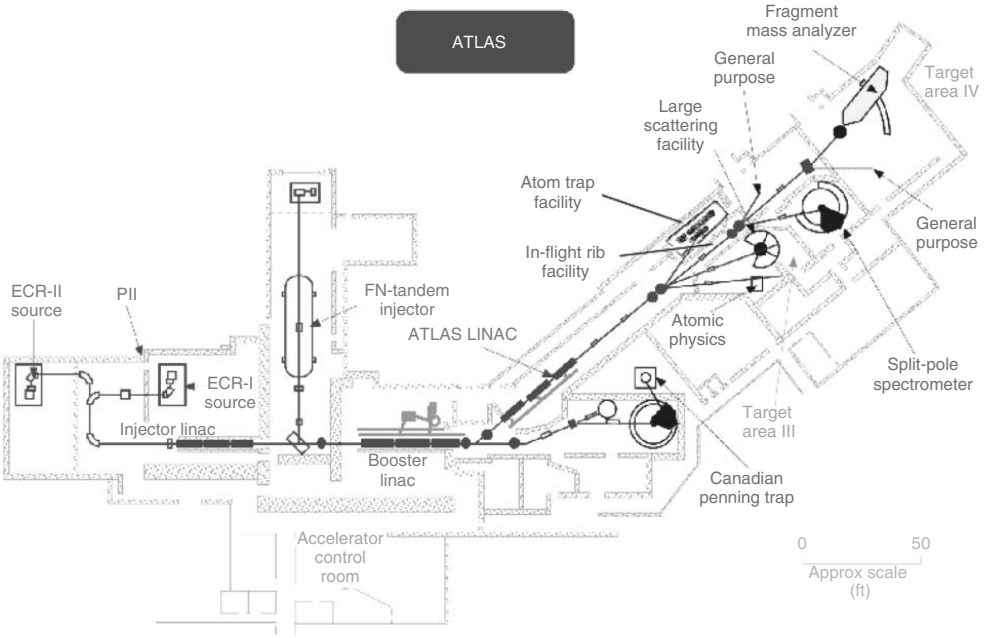


Figure 3.16 ATLAS layout showing the FN tandem injector and the two positive ion injectors. The negative ion beam is stripped in the terminal of the tandem, before injection into the booster LINAC and before entry to the ATLAS LINAC. (Illustration courtesy of ANL.)

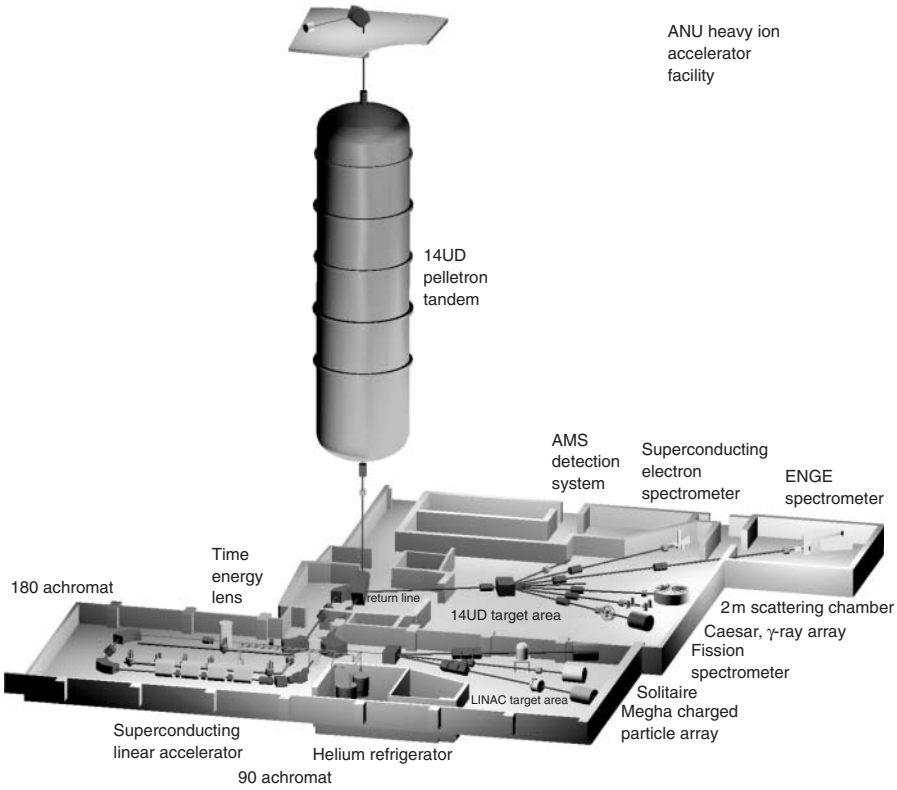


Figure 3.17 The 14UD injector at the ANU providing the beam to superconducting LINAC.

accelerator to the booster is complex and unique for each installation.

3.5 Summary

Electrostatic accelerators have provided the earliest and most productive tools for nuclear research. Their variety and robustness ensured their wide application in industry and diverse areas of science. As tandem accelerators became the preferred devices for heavy ion physics, their roles expanded to injectors into boosters and as accelerators of radioactive ion beams. Their place in the evolution of heavy ion physics was and is the link to future generations of accelerator systems. Because of their multiple strengths, they continue to make crucial contributions to research at the leading edge of heavy ion physics including the study of the quark–gluon plasma and the physics revealed with radioactive beams.

Acknowledgments

This is an update of ''Potential Drop Accelerators'' by R. G. Herb and G. M. Klody in the ''Encyclopedia of Applied High Energy and Particle Physics'' WILEY-VCH Verlag BmbH & Co KGaA, 1989. Much of the history and descriptions of applications is based upon that c. I would like to thank

all the colleagues and institutions that have generously provided illustrations and information. Greg Norton at NEC has been a crucial resource as have Keith Fifield at ANU, Dannie Steski at BNL, Richard Pardo at ANL, and Martha Meigs at ORNL. I especially thank Christ of Vockenhuber for most of the section on small AMS facilities.

Glossary

atm: Atmospheric pressure.

Dielectric strength: The ability of an insulating material to sustain electric fields without breaking down.

HVEC: High Voltage Engineering Corporation – Manufacturer of CN, EN, FN, and MP Van de Graaff and ICT accelerators. The company is no longer operating.

Ion implantation: The insertion of energetic ions into a material.

Lattice atoms: The structural atoms organized in a regular pattern that comprise a crystal.

LC resonant circuit: A combination of inductors and capacitors through which electric current will oscillate at a fixed stable frequency.

NEC: National Electrostatics Corporation – Manufacturer of Pelletron accelerators.

Nuclear transmutation: The conversion of a highly radioactive isotope into a less radioactive one using a nuclear reaction caused by ion beam bombardment.

RF electrode: One plate of a capacitor in a radio frequency circuit used to couple power.

Semiconductors: Materials such as silicon and germanium serving as the host for electrical devices.

SF_6: Sulfur hexafluoride gas used to insulate high electric fields.

Sputtering: The energetic liberation of atoms from a solid caused by the impingement of high-velocity atoms or ions.

Stripping: The process in which electrons are removed from an ion during its passage through either a low-pressure gas region or a thin carbon foil.

Tandem: An accelerator in which a negative ion is accelerated to a positive terminal where electrons are stripped from it to produce a positive ion that is then accelerated again to ground.

Terminal: The high-voltage terminal of an accelerator.

References

1. NEC *http://www.pelletron.com/* (accessed 12 February 2013).
2. ORNL *http://www.phy.ornl.gov/hribf/accelerators/* (accessed 12 February 2013).
3. Naylor, H. (1968) An electrostatic particle accelerator. Thesis (PhD--Physics). University of Auckland.
4. (a) Synal H.A., Schulze-König T., Seiler M., Suter M., and Wacker L. (2013) Mass spectrometric detection of radiocarbon for dating applications. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **294**, 349–352. (b) Synal, H., Stocker, M., and Suter, M. (2007) MICADAS: a new compact radiocarbon AMS system. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **259** (1), 7–13.
5. Vockenhuber C., Alfimov V., Christl M., Lachner J., Schulze-König T., Suter M., and Synal H.-A. (2013) The potential of he stripping in heavy ion AMS. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **294**, 382–386.
6. Wittkover, A.B., Betz, H., and Betz, D. (1973) *Phys. Rev. A*, **7**, 159–167.
7. Hotchkis, M.A.C., Child, D., Fink, D., Levchenko, V., and Wallner, A. (2012) Investigation of gas stripping at 4.1 MeV for high mass negative ions. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **294**, 387–391.

8. Müller, A.M., Lachner, J., Christl, M., Suter, M., and Synal, H.A. (2010) Competitive 10Be-measurements below 1 MeV with the upgraded TANDY AMS facility. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **268** (17–18), 2801–2807.
9. Saclay, B., Cauvin, J.G., and Ramstein, G. (1990) *Nucl. Instrum. Methods Phys. Res., Sect. A*, **287**, 263–267 and IBA05-5.
10. MSU *http://www.nscl.msu.edu/ria/* (accessed 12 February 2013).
11. South Africa *http://www.tlabs.ac.za/Research.htm*.
12. Ganil *http://www.cnrs.fr/Cnrspresse/en370a1.htm* (accessed 12 February 2013).
13. ANL *http://www.phy.anl.gov/atlas/index.html* (accessed 12 February 2013).
14. INFN – LNL *http://www.lnl.infn.it/accelerators/accelerators.html* (accessed 12 February 2013).
15. TRIUMF *http://www.triumf.info/public/* (accessed 12 February 2013).
16. Linac *http://www.dfn.if.usp.br/pesq/linac/* (accessed 12 February 2013).
17. FSU *http://www.physics.fsu.edu/nuclear/Brochures/SuperconductingLinearAccelerator Laboratory/default.htm*.
18. *http://www.stonybrook.edu/SUNY http://www.sunysb.edu/nsl/laboratory.html* (accessed 12 February 2013).
19. Center for Experimenal Nuclear Physics and Astrophysics (CENPA) *http://mist.npl.washington.edu/overview.html* (accessed 12 February 2013).
20. INFNLNS The EXCYT radioactive beam facility, *http://www.lns.infn.it/excyt/index.html* (accessed 12 February 2013).
21. CRNL, Hoffmann, C.R., Bingham, C.B., Heighway, E.A., Hulbert, J.A., Omrod, J.H., and Schneider, H.R. (1975) The Chalk River Heavy Ion Superconducting Cyclotron. *IEEE Transactions on Nuclear Science*, Vol. NS-22, No. 3, 1647–1650.
22. Heidelberg – mpi *http://www.mpi-hd.mpg.de/blaum/accelerators/post-accelerator/index.en.html*.
23. RHIC *http://www.bnl.gov/RHIC/RHIC_complex.htm* (accessed 12 February 2013).
24. JAEA *http://www.lnl.infn.it/∼HIAT09/papers/oral/MO3ST.pdf*.
25. Prakash, P.N., Datta, T.S., Ajith, B.P., Kumar Antony, J., Barua, P., Chacko, J., Choudhury, A., Chandari, G.K., Ghosh, S., Kar, S., Krishnan, S.A., Manoj, K., Rajesh, K., Mandal, A., Mathuria, D.S., Meena, R.S., Metha, R., Mistri, K.K., Pandy, A., Suresh Babu, M.V., Sahu, B.K., Sarkar, A., Sonti, S.S.K., Rai, A., Venkartramanan, S., Zacharias, J., Bhowmik, R.K., and Roy, A. (2002) Superconducting linear Accelerator system for NSC. *Pramana – J. Phy.* **59**, 849–858.
26. ANU *http://wwwrsphysse.anu.edu.au/nuclear/accelerators.php* (accessed 12 February 2013).
27. BARC-TIFR *http://www.tifr.res.in/∼pell/* (accessed 12 February 2013).
28. The Weizmann Institute Accelerator Laboratory *http://www.weizmann.ac.il/physics/oldfiles/accel1.html* (accessed 12 February 2013).

Further Readings

Hellborg, R. (ed.) (2005) *Electrostatic Accelerators, Fundamentals and Applications*, Springer-Verlag, Berlin, Heidelberg.

Large Electrostatic Accelerators, (1974) *Nuclear Instruments and Methods*, in D. A. Bromley (ed.), Vol. 122, North Holland, Amsterdam.

4

Linear Accelerators

Robert Jameson, Joseph Bisognano, and Pierre Lapostolle†

† Deceased

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

4.1
Introduction

Particle accelerators were originally invented as enabling tools for physics research, for example, nuclear and elementary particle physics. Many, if not most, of the important concepts were invented for almost all accelerators and, in particular, linear accelerators (LINACs), around 1930. However, technological capabilities at that time did not permit anything but a proof of principle of LINAC operation. By using the term *LINAC*, we mean devices where, typically, the particle passes only once, on a straight trajectory through a possibly long structure to gain energy, and contrast this to ''circular'' devices such as synchrotrons and cyclotrons, where the particle passes a great number of times through the same accelerating structure. Accelerators such as microtrons and recirculating LINACs, where particles travel only a ''few'' times through the accelerating structure, show the limitations of such distinctions. With one or only a few passes of acceleration, LINACs typically provide the highest throughput of particles at a given energy.

After the development of high-power short-wave and microwave radiofrequency (RF) sources for radar, it was possible to undertake the construction of useful LINACs. Successive improvements pushed by continuous requests for still more energetic and powerful beams of better and better quality led from the initial crude machines to the very elaborate LINACs of today. In the past 20 years, the use of superconducting materials such as niobium has been a most significant innovation.

Besides being used in nuclear and particle physics research, LINACs are now also used as a source of energetic particles in equipment for medical, biological, and industrial applications, including industrial irradiation, radiotherapy, medical tool sterilization, food conservation, and baggage inspection, and provide intense beams of neutrons or photons for use as probes for condensed matter, materials, and in biological sciences. In the future, LINACs will enable methods to deal with radioactive waste and other societally important applications.

In this chapter, Section 4.2 explains the general principles of linear acceleration. A historical approach has been taken. Both the evolving goals and the difficulties encountered that have led to the sophisticated and complex designs of today's electron and proton LINACs are explored. Electron LINACs are described with some details of their components, limitations, application areas, and advanced research for the future. Proton LINACs are

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

described next, explaining in more detail the problems encountered at both low and high energies. Again, intensity and beam quality limitations are considered. Present and future applications are presented.

Heavy-ion LINACs, which have some special considerations and unique realizations, represent the final general category of LINACs.

The induction LINAC and alternating-phase-focused (APF) LINACs are covered briefly. The work being done on the search for new methods of acceleration aimed at reaching still higher energies in shorter distances for future generations of colliding beam machines for high energy physics and for tabletop devices for small laboratory research is described.

The final section, on engineering topics, covers the various aspects of LINAC construction: accelerating structures, magnetics, RF engineering, wideband electronics, instrumentation, automatic controls, and computers.

4.2 General Principles and Brief History

4.2.1 The Original LINACs

In the first part of the twentieth century, the development of nuclear physics and, later on, the discovery of induced radioactivity prompted the search for energetic projectiles to produce new nuclear reactions. The first particle accelerators were soon made from DC high-voltage generators, but source voltage and, therefore, particle energy were limited by electrical breakdown. To obtain more energetic projectiles, new methods had to be invented.

The proposal by Ising [1] that led to the development of LINACs called for an array of metallic tubes with an applied potential creating an electric field between tubes, but not inside them. Charged particles launched along the axis of the tubes were to be accelerated while traversing the gap between two adjacent tubes but drifted at constant velocity when inside a tube. Ising suggested that pulsed potentials be applied across the drift tubes at the time the charged particle bunches would enter the gap, and be switched off once they entered the tubes. Such a method, difficult to apply accurately enough then, is today the principle of the induction LINAC, mentioned in Section 4.6.1. Instead of using very accurately synchronized high-voltage pulses, Wideroe [2] proposed more simply to apply RF voltages on a succession of drift tubes, as shown in Figure 4.1. He tested a small device with two gaps and a single drift tube, and, with an RF peak voltage of 25 kV, accelerated Na^+ and K^+ ions produced by a heated filament to an energy of 50 keV.

Sloan and Lawrence [3] at Berkeley used 30 tubes of progressively increased length and a gap voltage of 42 kV at an RF frequency of 10 MHz to reach an energy of 1.25 MeV for Hg ions.

4.2.2 Phase Stability

To produce not only a few energetic particles but also a good useful beam, the following problems have to be resolved:

- how to obtain a large enough number of particles (*beam intensity*)?
- how to concentrate the particles – or focus them – into a thin enough beam (*beam quality*)?

Let us look further into this succession of accelerations between drift tubes at two

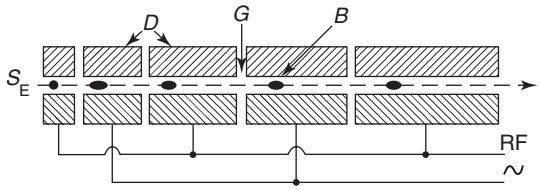


Figure 4.1 Principle of the Wideroe method of RF linear acceleration used in 1931 on the first machine of Sloan and Lawrence. *S* is a source of positive ions. *D* represents drift tubes alternatively connected to the two terminals of an RF voltage. The ion bunches *B* are accelerated in the gaps *G*. Because the velocity increases as the bunches are accelerated, the length of the drift tubes must increase along the machine.

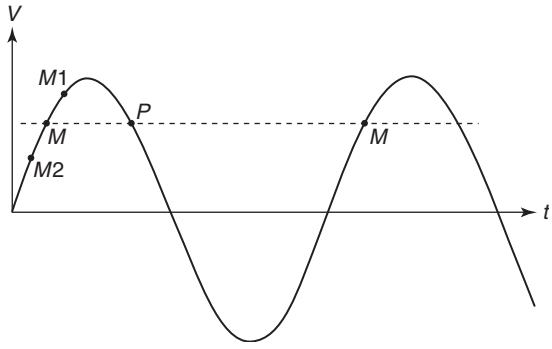


Figure 4.2 The voltage *V* in the gap *i* between the drift tubes has a sinusoidal variation with time *t*. To reach gap $i+1$ at the right time, the proper acceleration is necessary; the required amplitude occurs at two phases per period: *M* and *P*. While points *P* are unstable, points *M* are stable and particles form bunches around such phases.

consecutive gaps numbered i and $i+1$, where a sinusoidal RF voltage is applied (Figure 4.2). Suppose a particle arrives at gap i with the correct energy. The length of the drift tube between the gaps i and $i+1$ has been chosen for a given energy increase through gap i. Such a gain is produced according to Figure 4.2 at only two phases on the RF cycle: phases *M* and *P*. If particles crossing the gaps at these exact phases were the only ones that could be accelerated properly, the beam intensity would be extremely small. Let us examine, however, what happens around phase *M*, say at *M*1 and *M*2. A late particle at *M*1 is accelerated more and catches up with *M*. An early one at *M*2, on the contrary, is slower and also meets with *M*. It is said that there is phase stability around phase *M*. It can be seen that around *P* there would be instability. Thus, particles are caused to bunch around phases such as *M*, where the voltage is rising, and a more intense beam can be accelerated.

4.2.3 Focusing

Let us look now at focusing. Figure 4.3 sketches an electric field distribution in an accelerating gap; off-axis particles are focused at the gap input and defocused at the exit. For phase stability, however, at phase *M* (Figure 4.2), the field must be rising with time; defocusing is then larger than focusing. The overall effect of accelerating gaps is defocusing.

Additional focusing has to be provided, particularly for low-velocity particles [the velocity is small compared to the velocity of light: $v = \beta c$ (where $\beta \ll 1$)] that are nonrelativistic and are easily deflected by transverse electric fields. At the highest energies, focusing is introduced with

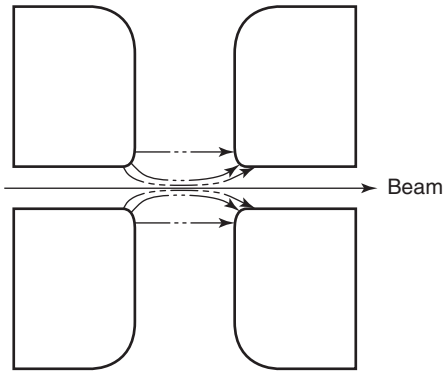


Figure 4.3 The electric field lines in the gaps terminate normally on the metallic surfaces of the drift tube holes. There is a focusing action at gap entrance and defocusing at gap exit if the gap is accelerating. If the voltage is increasing while particles cross a gap, the net effect is defocusing.

quadrupole magnets, typically in an alternate gradient configuration.

4.2.4 Electron Accelerators

In electron LINACs, the velocity rapidly approaches the velocity of light. The scheme of Wideroe is not easy to use; extremely high frequencies are needed, and the drift tubes become resonant antennas supporting very high currents with high losses (Section 4.4.3). Another method had to be invented. Using high-power microwave radar sources available in 1947, Ginzton *et al*. [4] designed an accelerating structure consisting of a stack of pillbox resonators with a central hole, called the *iris-loaded waveguide* (Figure 4.4). The holes provide a passage for the beam and also allow the electromagnetic energy to travel from resonator to resonator. The cavity array behaves as same as an RF filter of relatively narrow bandwidth; the phase relation between two successive cavities depends on the frequency. With a proper choice of the dimensions, at a given frequency, the field in successive cavities can be such that a traveling wave is generated with phase velocity equal to that of the electron beam: this synchronization ensures acceleration in every cell. The waveguide is interrupted at intervals to replenish the structure with energy and to install quadrupole magnets that provide the extra focusing needed for transverse stability.

Slater [5] at MIT, Walkinshaw in England, and others derived the theory of the iris-loaded waveguide and studied many other types of structures. Further developments at Stanford led to the longest electron LINAC to date; the 3 km long Stanford Linear Accelerator Center (SLAC) LINAC has accelerated electrons and positrons to 50 GeV.

In the 1980s, decades of work on superconducting accelerating structures came to fruition, with application to both circular [the CERN LEP (large electron–positron) collider] and the Continuous Electron Beam Accelerator Facility (CEBAF, now Jefferson Laboratory) recirculated LINAC [6]. For continually operating [continuous wave (c.w.)] or long-pulse LINACs of high Q, standing wave, multicell resonators rather than traveling wave structures are used. With the decision to use superconducting technology on the International Linear Collider (ILC) [7], free electron lasers (FELs) as optimal drivers [8], and energy recovery light sources [9], superconducting standing wave LINACs will likely become

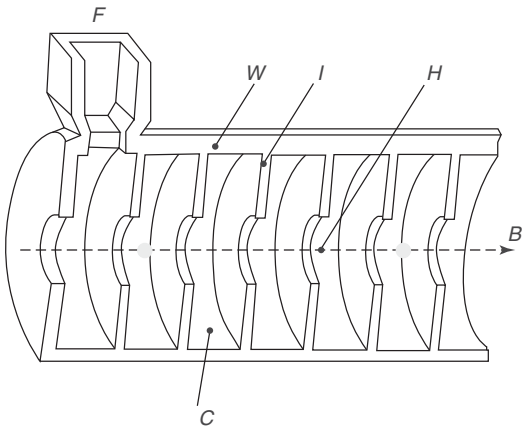


Figure 4.4 The cylindrical waveguide *W* is loaded by irises *I*; the holes *H* provide coupling between the cells *C* for electromagnetic energy transfer and beam passage along the axis *B*. The structure can also be considered as a stack of pillbox resonators *C* coupled through the axial holes *H*. Electromagnetic energy is fed at one end from a feeder *F*, through a ceramic window (not shown) to keep the interior of the structure under high vacuum.

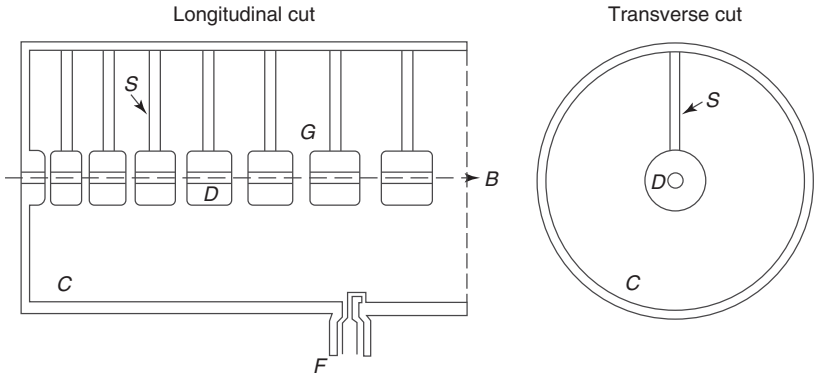


Figure 4.5 Alvarez structure. Drift tubes *D* supported by stems *S* are put along the axis of a cylindrical cavity *C*. The cavity is supplied with RF power from a coaxial line *F* (through a window not shown) by a coupling loop, and excited in a longitudinal mode such that the electric component of the electromagnetic field is directed along the axis and concentrated in the gaps *G*. The field varies sinusoidally with time but the drift tube lengths are chosen so that the beam is cumulatively accelerated along the axis *B*.

the dominant technology for large electron accelerators at the highest energies. Room-temperature standing wave structures (often of a side-coupled configuration) also have applications, especially as part of circular accelerators.

4.2.5 Proton Accelerators

For protons of low velocity, in 1947, Alvarez *et al.* [10] designed a 200 MHz structure at Berkeley, using surplus radar components. In a 12 m long, 1 m diameter resonant cavity, excited in a mode with uniform electric field on the axis, a succession of metallic drift tubes with axial holes provided cumulative acceleration of the beam (Figure 4.5). At first, additional focusing was provided with the help of flat foils or grids located at the exit of each gap (Figure 4.6). These devices suppressed the defocusing effect described in Figure 4.3, but stopped a nonnegligible fraction of

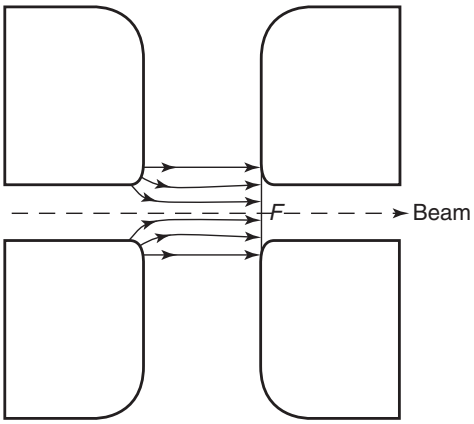


Figure 4.6 Foil (or grid) focusing. By closing the hole at the exit of the gap, the foil or grid suppresses the defocusing effect (Figure 4.3). The gaps are now focusing, but the beam is scattered and some particles are lost.

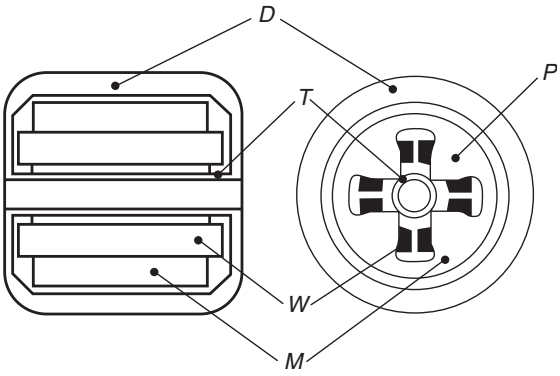


Figure 4.7 Magnetic quadrupole lens inside a drift tube. The windings *W* create a quadrupolar field between the four poles *P*. The magnetic flux returns through the magnet yoke *M*. The whole lens is encapsulated inside the drift tube *D* closed on the axis by a thin tube *T*.

the particles and deflected others in an irregular way, reducing the intensity and the beam quality.

It was then discovered by Courant *et al.* [11] that effective transverse focusing could be provided by a succession of focusing and defocusing quadrupole magnets, and Blewett [12] suggested installing such magnets inside the drift tubes. A four-pole magnet (Figure 4.7) constitutes a focusing lens in one transverse direction and a defocusing lens of equal strength in the orthogonal one. A succession of focusing and defocusing lenses, properly spaced, is globally focusing; with quadrupole magnets, one thus gets focusing in the two transverse directions [13]. This method is called *alternating-gradient* (AG) focusing.

The Alvarez structure, with various improvements, is used now for ion beam β values in the range 0.035–0.5; other types have been developed for lower and higher energies (Section 4.4.4).

As with electron accelerators, superconducting radiofrequency (SRF) LINACs have found important applications for protons and ions, with the Spallation Neutron Source (SNS) being the largest application to date [14]. A proposed rare isotope accelerator (RIA) using SRF is also under consideration to generate short-lived ion species.

4.3
Electron Linear Accelerators

4.3.1
Electron Source and Preacceleration (LINAC Injector)

Electrons are very light particles. According to Einstein's energy relation, one has $E = E_o + W = mc^2 = m_o c^2/\sqrt{1-\beta^2} = E_o/\sqrt{1-\beta^2}$, where the total energy E is the sum of the rest energy E_o (equal to $m_o c^2$) and kinetic energy W, where m is the total mass and m_o the rest mass; $\beta = v/c$ is the ratio of the velocity of the particle to the velocity of light. For electrons, $E_o = m_o c^2 = 511\,\text{keV}$. Table 4.1 shows how β changes with kinetic energy W.

As can be seen, with an energy gain in typical modern electron LINACs of $10\,\text{MeV m}^{-1}$ or more, the velocity is almost everywhere very close to the velocity of light, apart from a first section called *the injector*. A high-energy electron LINAC is then an injector followed by a number of identical sections of an accelerating structure, with the addition of some focusing. The simplest form of electron injector consists of a thermionic cathode emitting electrons into a DC voltage of 20–150 kV. A grid pulser modulates the beam to form "bunches" of adjustable length and amplitude in the range of microamperes to kiloamperes for certain applications. The conventional buncher section of an RF LINAC is an accelerating structure (Figure 4.8) with variable β bringing the energy to the order of 10 MeV and compressing the output bunches to 10° or less. Focusing is generally provided by magnetic solenoids.

Over the past few decades, electron injector technology has been revolutionized by the introduction of the laser-driven

Table 4.1 β as a function of W for electrons.

W (MeV)	0.020	0.05	0.10	0.20	0.50	1	2	5	10
β	0.272	0.413	0.548	0.695	0.863	0.941	0.979	0.996	0.999

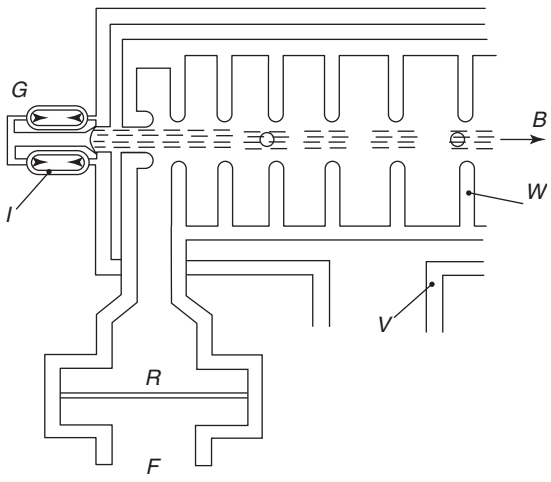


Figure 4.8 In this very schematic drawing of the bunching section of an electron LINAC, the gun *G*, at high voltage, is supported by an insulator *I*. The iris-loaded waveguide *W* is put inside a vacuum vessel *V* evacuated to high vacuum. A feeder *F* brings microwave power to the accelerating structure through the window *R*. The beam *B* is accelerated in the form of bunches. The details of the gun, the grid, and the chopper or prebuncher, as well as all the focusing elements, have been left out of this drawing.

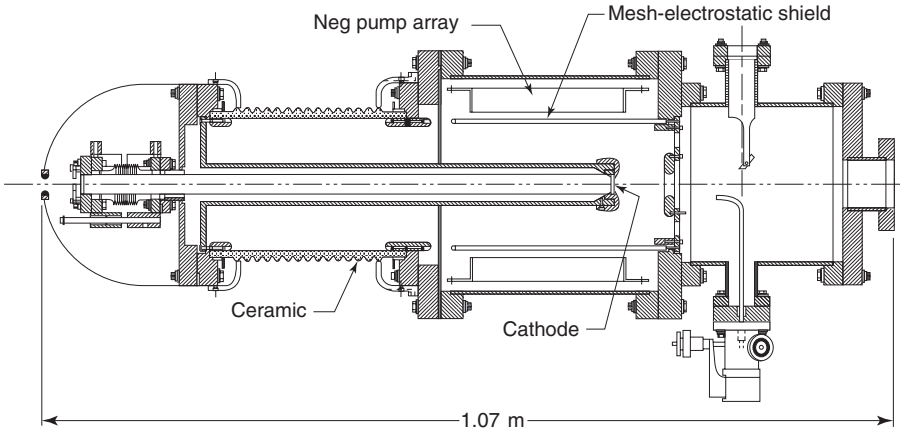


Figure 4.9 Schematic of Jefferson Laboratory polarized electron source. (Courtesy: Jefferson Laboratory.)

photocathode. This advancement has allowed the production of electron beams of significantly higher phase space density (smaller beam emittance) and shorter pulse lengths. For RF LINACs, the laser can be modulated to produce short electron bunches occupying only a narrow range about the proper RF phase. The photocathode is located at the entrance of an RF accelerating cell gap or a DC field (Figure 4.9). With high gradients, the electrons are accelerated very rapidly, helping to alleviate the effects of space–charge and preserve beam brightness. Bunch charges ranging from pico- to nanocoulombs can be obtained with normalized emittances below 10 μm at peak currents of tens of amperes. With the introduction of DC voltages approaching 1 MV or RF gradients of several tens of megavolts per meter, R&D is progressing toward submicrometer normalized emittances. The success of this effort will strongly impact the future of light sources such as FELs and energy recovery linear accelerators (ERLs).

Moreover, by using a photocathode and a laser, it is possible to produce longitudinally polarized electrons, which are of great interest in certain physics experiments. For example, circularly polarized near infrared (IR) light from a Ti:sapphire laser is projected onto a gallium arsenide (GaAs) cathode in the CEBAF polarized electron source [15].

4.3.2 Electron LINAC Characteristics

An RF electron LINAC consists of one or several sections of an accelerating structure, each generally several meters in length. Accelerating structures for short-pulse LINACs are of the iris-loaded waveguide, traveling wave type. They are usually operated in the microwave band around 10 cm wavelength (3 GHz). The shape of the cavities has been optimized to reduce (for a given accelerating gradient) both the peak electric field on the surfaces (to avoid the risk of electrical breakdown) and the RF power losses in the walls resulting from the finite conductivity of the material (usually copper). The SLAC cavities are pulsed for a few microseconds in a traveling wave mode, the first microsecond being used to fill the structure

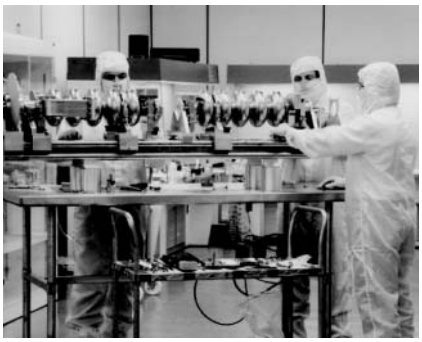


Figure 4.10 Assembly of CEBAF cavity in a clean room. (Courtesy: Jefferson Laboratory.)

from input to output (3 m). Only 1/10th or at most a few tenths of the power remains at the output, the rest being absorbed by losses in the walls or transferred to the beam. Electrons at relativistic velocity ride on the RF wave crest. As the RF cavity fields are raised by increasing the RF source power, the energy gain per unit length is raised accordingly.

For longer pulse and up to fully c.w. operation, RF LINACs are made with standing wave cavities. For optimal coupling to the electron beam, the side-coupled structure (Section 4.4.4) is typically chosen for normal conducting applications. The introduction of superconducting materials such as niobium has had a dramatic impact on standing wave LINAC technology. At RF, superconductors do still exhibit small losses, but resonators can offer Qs that approach 10^{10} at gradients of order $20\,\mathrm{MV\,m^{-1}}$. This is to be compared to a typical room-temperature copper structure where Qs of 10^4 are more typical. These low losses have enabled superconducting LINACs to be operated in a c.w. mode at these high gradients, which would be entirely impractical for room-temperature structures. True c.w. operation (and also the inherently high Q) provides at least an order of magnitude improvement in beam stability, which in turn has enabled applications that require smaller energy spread and beam temporal jitter. In fabricating superconducting cavities, much care must be taken to ensure the quality of the superconductor with minimization of imperfections in the material, meticulous assembly in clean-room environments, conditioning, and cool-down protocols (Figure 4.10).

Energy in big LINACs is mainly limited by cost considerations involving the length of the structure and the amount of RF power. In short-pulse machines such as SLAC, the energy can be increased economically by using a device called the *Stanford Linear Accelerator Center Energy Doubler* (SLED) [16, 17]. It makes use of high-Q resonant cavities that are filled with electromagnetic energy during the first part of the RF pulse. In the second part, the RF phase of the klystron is reversed and the cavities are automatically emptied, effectively resulting in pulse compression and a momentary reflected power increase by a factor of almost 3. With the help of 3 dB couplers, this power can be fed into the accelerating structures, and, over a very short pulse, can result in about 80% more acceleration.

4.3.3
Electron Beam Intensity Limitations

In applications such as high-energy LINACs for physics experiments or an

RF LINAC providing the driving beam for an FEL, beams of extremely good quality are required: very small transverse beam dimensions (measured in a tenth or hundredth of a millimeter or even in micrometers) and very small energy spread (0.1% or less). In these and other applications, beam intensity must also be as high as possible. Energy spread from the electron bunch phase width along the RF wave can be compensated using very short bunches and by ensuring that the bunches, on average, ride on the crest. However, to achieve the shortest bunch lengths, designs for FELs and colliders use bunch compression, which requires an energy chirp of the bunch that has passed through dispersive optics. The chirp is typically produced by off crest acceleration.

One limitation in short-pulse electron LINACs comes from beam loading. To be accelerated, the beam has to extract energy from the electromagnetic field, thus reducing the field strength in proportion to the beam intensity. As a result, under short-pulse operation, it is not easy to obtain a flat energy gain over the pulse. For long pulses or c.w., automatic control of the fields can be applied and then heavy beam loading (i.e., a large ratio of power supplied to the beam to that supplied to the accelerator structure losses) is used to raise the overall efficiency. In superconducting LINACs, for example, the structure losses are miniscule in comparison with the power transferred to the beam.

With the ever-decreasing bunch length and emittance and ever-increasing peak currents, wakefield phenomena and coherent synchrotron radiation (CSR) processes often limit the overall beam performance. As a generalization of beam loading, whenever a beam pulse passes through variations in its vacuum environment (e.g., the RF structure, bellows, and beam position monitors), the image charges change direction and radiate. The electromagnetic energy comes from the beam's kinetic and field energy. With a frequency content determined by the details of the environmental discontinuities, the electron bunch may find its momentum (both longitudinal and transverse) perturbed with a time variation. This is the basic notion of wakefield: the field generated by the beam, which is left behind and can act back on succeeding beam particles. A related but somewhat different notion is that of CSR. A single electron will radiate when it is accelerated, for example, on being steered along a curved path by a bending dipole magnet. If the electron is relativistic, this radiation has a very high frequency content (γ^3 of the typical revolution frequency), and the characteristic wavelength is typically much shorter than current modulations on the bunch or the bunch length itself. However, in the short bunch regimes typical of cutting-edge LINACs, the bunch length, in fact, is shorter than the characteristic wavelength of the radiation. The bunch can radiate as a single charge, with the radiation power enhanced by a factor corresponding to the number of electrons radiating in concert. Again, if the electron bunch is emitting energy, the beam energy must be diminished and will have a temporal variation. This is CSR.

The wakefields and CSR cause momentum changes that vary along the length of the electron bunch. This effectively increases the bunch energy spread and transverse emittance, which can destroy necessary resolution for spectroscopy and degrade performance of FELs. Moreover, the momentum variations can be exchanged for position variations in magnetic transport systems, yielding pulse lengthening, and spot size

increases. Finally, beam instabilities can be generated, which can grossly degrade beam properties.

Another serious limitation comes from a class of phenomena called *beam breakup* (BBU), which results in beam deflection and eventual particle loss in the accelerator walls. The LINAC structure can be excited not only on the fundamental accelerating mode but, just the same way as an ordinary waveguide, in other modes, the most dangerous of which are transverse hybrid modes that deflect the beam. These modes can be excited by particles traveling off the cavity axis, or as a result of cavity asymmetries. Along a single accelerator section, the interaction of the beam with such a transverse mode can lead to uncontrollable beam deflection when the current is above a certain threshold. The deflected beam can couple its energy into the transverse mode, which has a longitudinal electric field component off axis. A feedback loop is generated, which can go unstable at high currents (effectively, high gain). This is called *regenerative BBU*. In a long accelerator, each cavity, although below the threshold of this unstable oscillation, can nevertheless behave as same as an amplifier. A long chain of such amplifiers, even if each gain is small, can eventually deflect the beam beyond the edge of the irises; this is called *cumulative BBU*. The wakefield effect discussed earlier leads to excitation of higher order deflecting modes, which can affect succeeding beam bunches. These wakefields can build as the chain of bunches get deflected and more strongly excite downstream cavity modes. Their integrated effect on particles passing through them can cause effective emittance growth that at least dilutes brightness, or may even cause beam loss. The transverse effects are combated by eliminating boundary discontinuities wherever possible, and by making geometrical changes in the accelerating cavities to prevent the beam from being synchronized with the offending modes. The effects in RF structures become more serious as the operating frequency is raised and the iris diameter becomes smaller, and this becomes a limitation in high brightness designs. In the design of c.w. LINACs using the latest generation of superconducting RF cavities, either cost optimization or energy recovery for high average currents favors configurations where the beam passes several times through the same accelerating structure. It was discovered during the commissioning of the Stanford Recyclotron that a new kind of instability, a hybrid between regenerative and cumulative BBU, can limit beam average currents. As in cumulative BBU, a series of marginally offset beam bunches can have their displacement amplified by a higher order mode (HOM) when passing through a long series of accelerating cavities. In addition, with recirculation, these offset bunches are allowed to pass through a given cavity for a second time, where the offset beam can excite the same HOM as happens in regenerative BBU. A closed feedback loop is formed, which can go unstable if the time delay is unfortunate and the beam current (which is equivalent to loop gain) is high enough. This "multipass BBU" is the principal limitation to average beam current in recirculated LINACs. Since superconducting cavity HOMs have relatively high Qs, prevention of BBU requires careful damping of the strongest modes with mode couplers. For applications in the milliampere range, this has proved to be sufficient to provide beam stability, for example, in the Jefferson Laboratory nuclear physics machine CEBAF. For ERLs to drive high-average-power FELs (e.g., the Jefferson Laboratory IR source)

or for ERLs to generate incoherent synchrotron radiation (e.g., Cornell ERL) [9], average currents of order 100 mA are envisioned, and BBU and HOM damping must be monitored closely in cavity design and construction to ensure success.

4.3.4 Uses of Electron LINACs

Most of the pioneering electron LINACs in the world have been built for physics experiments. Table 4.2 lists four of the largest.

An important feature of large electron machines is the possibility of producing and accelerating positrons. Such particles are produced from the impact of high-energy (33 GeV at SLAC) electrons on a converter from which $e + e^-$ pairs emerge. Production efficiency is proportional to incident electron energy and is usually in the range of a small percent per incident gigaelectronvolt. Since the mid-1960s, energetic electrons and positrons from LINACs have been frequently used to fill collider rings (where counter-circulating beams collide, maximizing available energy) for physics experiments and to fill storage rings for the generation of synchrotron light.

In the early 1980s, the requirements for performing electron–nucleus coincidence experiments to look into the quark nature of the nuclear force and for very high resolution spectroscopy of nuclear states pointed out the need for high-duty factor electron accelerators. Although first conceived as a pulsed LINAC filling a pulse stretcher storage ring, it soon became clear that c.w. operation of a superconducting LINAC would offer the best, most upgradable solution. The CEBAF was designed as a multipass recirculated LINAC, with ultimately five beams of distinct energy being accelerated simultaneously through superconducting cavities originally designed for the Cornell electron positron storage ring (CESR) (Figure 4.11). Initially, the gradient in c.w. operation was specified at $5\,\mathrm{MV\,m^{-1}}$. The electron beam (now from a photocathode polarized source) is sent up to 5 times through two antiparallel 400 MeV superconducting LINAC segments connected by beam lines to deliver a 4 GeV, 200 μA average current continuous beam as designed. The LINAC is constructed using 1.5 GHz, five-cell niobium cavities operating at a temperature of 2 K. Advances in SRF technology led almost immediately to operation at 6 GeV, and an upgrade to 12 GeV is planned for the near future, with average gradients nearing $20\,\mathrm{MV\,m^{-1}}$ for supplemental seven-cell cavity cryomodules. In addition, the technology was applied to smaller recirculated LINACs at Jefferson Laboratory, configured in an energy recovery mode, where the beam passes a second time through the accelerating structure 180°

Table 4.2 The four largest electron LINACs.

Laboratory	Energy (GeV)	Length (m)	Date of completion
Stanford (California)	50.0	3050	1966
Tsukuba (Japan)	2.5	320	1981
Orsay (France)	2.3	230	1959
Kharkov (USSR)	2.0	240	1964

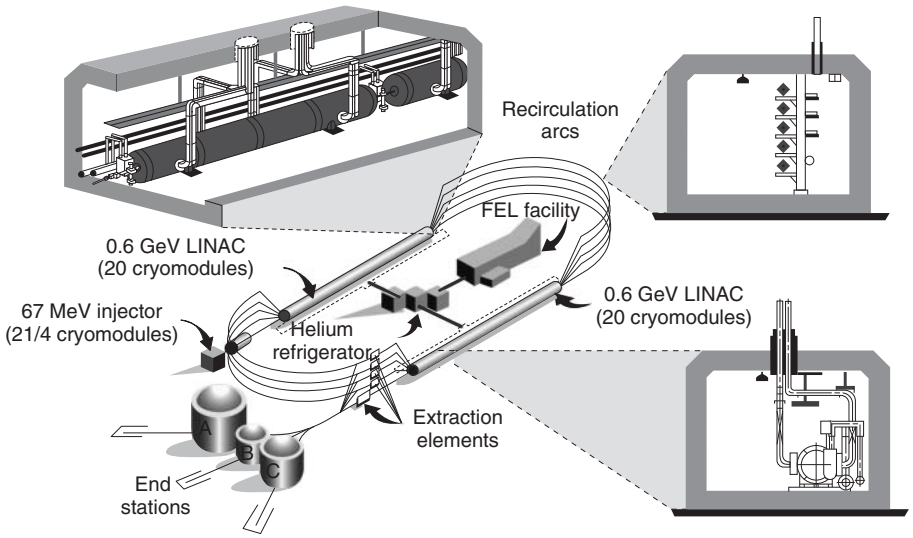


Figure 4.11 Configuration of the CEBAF recirculated LINAC at Jefferson Laboratory. (Courtesy: Jefferson Laboratory.)

out of phase with the field to deposit its kinetic energy for reuse. In this way, very high currents (approaching 10 mA) can be sustained for high-average-power FEL operation without making absurd demands on RF power. This notion of energy recovery is also of interest in the design of LINAC-based synchrotron light sources where fresh, low emittance electron bunches are sent through undulators, periodic arrays of magnets that coherently enhance the synchrotron radiation process.

After the initial findings of the large hadron collider (LHC) in the decade that followed, exploration of the high-energy frontier in detail will require a next generation of electron–positron colliders. Since 1980, high-energy physicists, seeking methods to collide electrons and positrons in the center-of-mass range of 100 GeV and higher, have reached the conclusion that above a certain energy level, linear colliders (two opposing LINACs whose beams collide at a target point) have an economic advantage over colliding rings. This advantage arises from the fact that a LINAC beam only loses a negligible amount of energy through synchrotron radiation because the particle trajectories are not bent in circles. Much R&D has gone into finding the best technology for linear colliders, and in the 1980s–1990s, work progressed on room-temperature short-pulsed, higher frequency copper LINACs, superconducting long-pulsed LINACs, and two-beam acceleration as potential approaches. In 2004, an international committee determined that long-pulsed superconducting RF technology was the most attractive candidate and an international collaboration has been created to bring this exciting idea to fruition.

The ILC is conceived as a positron–electron collider offering center-of-mass energies of 200–500 GeV, and upgradable to 1 TeV. Polarized electron beams are envisioned. In pulsed mode, the superconducting structures are expected to achieve

gradients in excess of 35 MV m^{-1}, with a theoretical limit of about 50 MV m^{-1}.

Apart from physics experiments, many electron LINACs of lower energy are also now made for a host of applications. Medical or industrial irradiation, radiotherapy, industrial radiography, chemical reactions, or hardening processes of various materials induced by irradiation, medical sterilization, and conservation of agricultural products through ionization are some of the various applications that have been developed and are more and more extensively in use. Almost all hospital X-ray machines now in use contain small standing wave RF electron LINACs. For such applications, high power (several tens of kilowatts of beam power) and low energy (in most cases $\leq$10 MeV) are needed; beam quality is not critical, and the beam is scanned or widened to cover the needed irradiation area. The crucial parameter is average intensity, and costs are optimized in this respect. Sources based on laser Compton backscattering are also under development.

A special type of application of LINACs is flash or stroboscopic radiography; there, extremely intense but very short pulses (100 ns, for instance) are needed. Induction LINACs (Section 4.6.1) are used for this application, and RF LINACs can also be used. The Dual Axis Radiographic Hydro-Test (DARHT) Facility project at Los Alamos National Laboratory (LANL) represents the latest advance, with two induction LINACs producing 20 MeV electrons at 4 kA.

4.4 Proton Linear Accelerators

4.4.1 Injection into a Proton LINAC

Unlike electrons, protons need to have a very large kinetic energy to reach a velocity close to that of light. With a rest energy $E_o = 938$ MeV, the relation between β and W is shown in Table 4.3.

Even at the highest energy reached in proton LINACs today, the final velocity is <90% of the velocity of light. At injection, it is only a few hundredths of c, increasing along the machine at a typical accelerating gradient of around 2 MeV m^{-1} according to usual practice.

As explained in Section 4.2, protons, especially at low energy, are subject to a strong defocusing effect of the electric field, not at all compensated by the effect of the magnetic component. Furthermore, the magnetic field strength of static focusing lenses must increase as particle velocity decreases, and there is insufficient room in the drift tube to provide it. This led to the choice of at least 500 kV for the injection of protons in LINACs built before ~1980.

The source of protons, from a DC or RF discharge in hydrogen, was traditionally of the duoplasmatron type; that is, one in which the axial magnetic field is made to vanish at the extraction hole from an expansion cup. Such a source can produce a beam of very high quality with intensities of several hundred milliamperes of protons either pulsed or c.w. This source

Table 4.3 β as a function of W for protons.

W (MeV)	0.1	0.2	0.5	1	2	5	10	20	50	100	200	500	1000
β	0.015	0.021	0.033	0.046	0.065	0.103	0.145	0.203	0.314	0.428	0.566	0.758	0.875

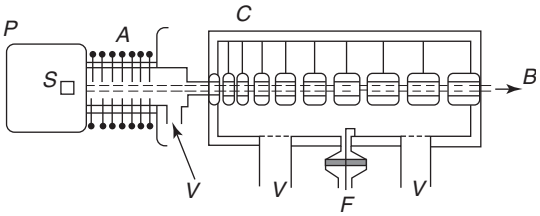


Figure 4.12 Schematic drawing of a proton LINAC with a high-voltage injector. The source *S* is placed on a high-voltage platform *P* and the beam is injected into the Alvarez structure (at ground potential) through a DC accelerating column *A*. A high vacuum is maintained all along the machine by vacuum pumps *V*. A coaxial line brings in *F* the RF power into the Alvarez cavity *C* by a coupling loop; vacuum tightness is provided by a window. Prebuncher and focusing elements are not shown in the drawing.

was traditionally put on a high-voltage platform at ~750 kV DC, and the beam injected from it into a drift-tube accelerator through an evacuated accelerating column (Figure 4.12).

Modern ion sources of the electron cyclotron resonance (ECR), the electron beam ion source (EBIS), and other types are used to produce intense beams of H^+, H^-, D^+, and other ion species; research continues – for example, using laser-driven sources – on producing higher currents of high-charge-state beams.

4.4.2 RFQ Accelerators

The first practical device using the idea that the RF field could be used to produce both an accelerating field and the necessary transverse focusing field, avoiding the need for magnetic elements, was invented by Teplyakov [18] and tested in the USSR by 1968. Radiofrequency quadrupole (RFQ) accelerators have replaced high-voltage DC injectors and moved injection energy into the drift-tube-type accelerating structure up to a few megaelectronvolts. The RFQ can provide strong AG focusing at low energy [19]. With it, the space–charge limit (Section 4.4.5) is sufficiently increased so that protons or other ions can be injected at energies from around 20–100 keV depending on the intensity and accelerated to around 2 MeV for injection into a drift-tube-type LINAC. Very high percentages (95–97%) of the input beam can be captured and accelerated with minimal growth in emittance.

Using this principle, the first device built outside the USSR was tested in 1980 [20, 21, 17]. Figure 4.13 shows the sketch of a four-vane RFQ; when excited on the quadrupolar RF mode, the field on the axis is mainly quadrupolar in configuration but also contains, owing to the modulation of the pole tips, a longitudinal component that is accelerating. The fields in the RFQ are continuous, rather than intermittent as in the drift tube approach. The period of the modulation is synchronized with the ion velocity. The four-vane structure is currently used at or above about 200 MHz (Figure 4.13). A lower frequency variant is the four-rod RFQ structure (Figure 4.14). There the tip profile is materialized by minivanes properly supported to achieve the chosen resonant frequency and the whole system is put into a cavity; such a structure can be operated at a lower frequency without becoming too large in size [22].

The efficiency of RFQ structures for focusing at low velocity comes from

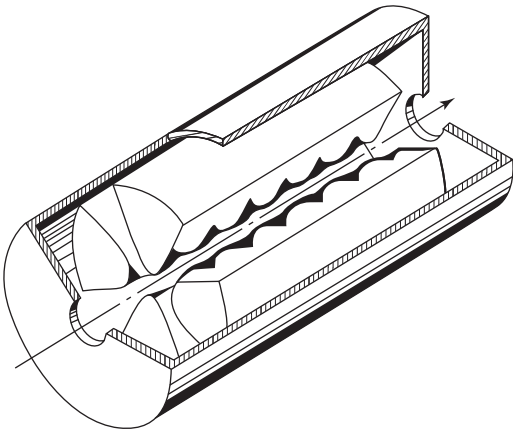


Figure 4.13 RFQ four-vane structure. The four vanes, in a quadrupolar arrangement, have their pole tips modulated; this produces a longitudinal RF field that allows the bunching and the acceleration of the beam.

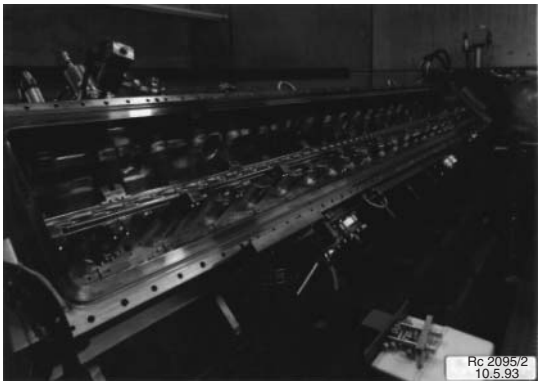


Figure 4.14 A four-rod-type RFQ built for GSI by Prof Dr A. Schempp, Institute for Applied Physics, Goethe University, Frankfurt.

the use of RF electric fields higher in strength than could be achieved with electrostatic fields and having a much larger focusing action than magnetic fields. RFQs are used to focus the beam emerging from the ion source to form bunches and accelerate them to an energy where adequate magnetic quadrupole focusing in the drift-tube-type structure can easily be provided (usually around 2 MeV). Because the RFQ has many cells in a reasonable length, the bunching and the acceleration characteristics can be slowly and carefully tailored. This nearly adiabatic process preserves the beam quality remarkably well and is a fundamental advantage of the RFQ with far-reaching consequences in modern high brightness proton and ion LINAC design.

4.4.3 Drift-Tube-Type LINACs

The RFQ loses efficiency as its cells become longer, necessitating a change in technique to accelerate to higher energies. The drift-tube principle is used.

The Alvarez type, with a focusing magnet located in every drift tube, provides the strongest and most uniform transverse focusing. The RF frequency of most of the large drift tube proton LINACs built in the 1970s is around 200 MHz, at which high-power triodes or tetrodes are available

for long-pulse operation. New machines tend to use higher frequencies to achieve higher brightness. In a 200 MHz Alvarez structure, the length of a few tens of drift tubes varies from about 5 to 15 cm as the particles accelerate from 0.5 to 5 MeV.

Magnetic quadrupoles are enclosed in each drift tube, which, supported by stems, must be carefully aligned. Drift tubes are installed in long cylindrical cavities of typically a few meters; the cavity length is sized to match the power delivery capability of the amplifier units attached either singly or through multiple-feed ports.

The Alvarez structure is efficient for medium velocities: $0.05 < \beta < 0.3$. The large diameter drift tubes necessary to house the quadrupole focusing magnets require more RF power. The structure can be modified to save RF power. Groups of a few cells having small diameter drift tubes without magnets provide acceleration without focusing, interspersed with a group with larger drift tubes containing magnets. Thus this type of LINAC can be termed a *separated-function* LINAC; compared to the drift tube LINAC (DTL), where transverse focusing is provided in every cell, the separated-function LINAC uses a series of cells designed for high acceleration efficiency and low RF power consumption, with transverse focusing inserted periodically as needed. There are now many variations of this idea, including superconducting ones. The overall transverse focusing is weaker but often adequate, and the savings in RF power can be substantial.

4.4.4 LINACs for Higher Velocity Ions

Above ~100 MeV, for protons, the efficiency of these LINAC structures again becomes too low, and another structure is needed. The Los Alamos Meson Physics Facility (LAMPF) LINAC was the first high-intensity LINAC to push beyond ~100 MeV. The ''separated-function'' principle is used, and the side-coupled cavity structure [23] (Figure 4.15) was developed using the standing wave mode at an RF frequency that is an integer multiple of

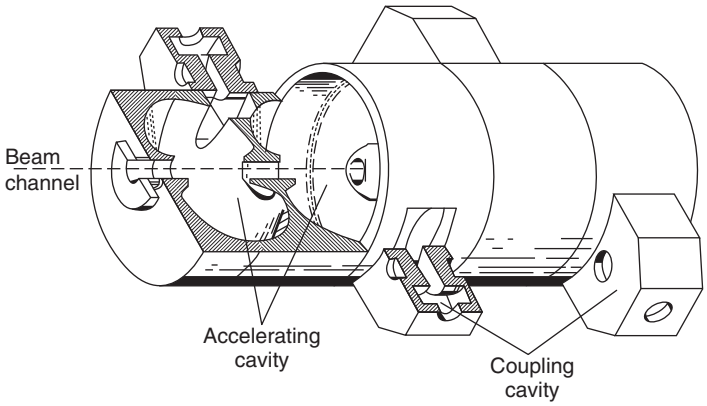


Figure 4.15 Side-coupled cavity structure. The on-axis accelerating cavities are designed with noses on the beam holes that prevent direct coupling and concentrate the accelerating field over a shorter length and time. The coupling between cavities is then made through lateral holes *H* via the side-coupling cavities.

2–4 times the frequency of the Alvarez section. In modern practice, this "room-temperature" structure type is used as a bridge between the structures of Section 4.4.3 to an energy at which a superconducting structure type becomes practical.

In Figure 4.15, one sees that the accelerating cells have noses on the beam holes in this structure. Such noses behave as same as drift tubes and concentrate the accelerating field over a shorter length, reducing the loss in acceleration resulting from the transit time through the gap. They also reduce the coupling between cells. The coupling necessary for the wave propagation is obtained through lateral cavities alternatively on one side and the other. There are two cells and two side-coupled cavities per wavelength. In a wavelength, there are two peaks and two nodes. The mode excited is such that the accelerating cells are at the peaks, that is, the field is high in them; the side cavities are on the nodes; so the RF losses are negligible in them. In this mode, the group velocity for energy propagation in the structure is high, and the cell fields are tightly synchronized – these features are very valuable in practice because the effect of errors is thereby reduced. The shaping of the cells optimizes the shunt impedance in room-temperature structures. In superconducting LINACs, the RF losses are low; so the cavity shape can be optimized on other criteria, and such structures can be used from below 100 MeV up to the velocity of light. The cell sequence is broken when it is necessary to place magnetic quadrupole lenses to provide focusing; special bridge coupler cells may couple RF power around the quadrupole.

Automatic control is applied to maintain the cavity field amplitude and phase constant by supplying extra drive when the beam is on. Because the required energy gain is fixed by the geometry, if beam loading were allowed to reduce the field, the trapping area in longitudinal phase space (called *acceptance*) would be smaller.

This side-coupled structure, with $\beta = 1$ for electrons, is presently used in most dental X-ray machines.

4.4.5 Beam Dynamics in Nonrelativistic LINACs

Almost all modern applications have to achieve an intensity as high as possible at the desired energy, along with a very good beam quality in terms of the beam confinement, aiming, or focusing. One figure of merit used is the beam brightness, defined as the *beam power* (or current, when the energy is fixed) divided by the phase space appropriate to the problem at hand. Phase space for the beam as a whole is six-dimensional, describing the physical size of the beam in x, y, and z, and the change in size with time or distance x', y', and z'; the area projected on one plane is called *emittance*.

Another figure of merit is to minimize the beam loss along the LINAC, and to concentrate any unavoidable beam loss to the lowest possible energies, as lost beam causes induced radioactivity in proportion to the lost particle energy.

The achievable beam intensity and quality are nonlinearly related, and raising the intensity tends to spoil the quality. Particle losses along the LINAC are to be strictly limited, as they would cause radioactivity buildup and complicate maintenance. The rule is that maintenance should be possible without remote manipulators. Achieving high brightness accelerated particle beams involves work on both the numerator and the denominator of the brightness

equation – beam intensity divided by beam emittance.

The numerator can be raised by brute force, but the large power requirements and engineering problems that result can be formidable, and better system efficiency becomes a key issue. A given accelerator channel can accelerate only a certain amount of current, depending on the imposed constraints. If more current is desired, the constraints or basic parameters might be changed, or several accelerator channels or modules could be used. The latter approach would multiply the system power requirement directly; therefore, one would prefer to find a more efficient way to raise the current, if possible.

Decreasing the denominator of the brightness factor involves building low emittance (but also intense and, therefore, bright) particle sources and then preserving the source brightness through all the subsequent steps of beam transport and acceleration to the final energy and the beam target. Owing to basic beam dynamics, the minimum phase space volume anywhere in an accelerator is no smaller than that at the particle source.

The main problem at nonrelativistic velocities is space–charge effects. The beam particles must be charged to be accelerated, and all have the same sign; so they tend to repel each other. When the beam intensity is increased, the charge in the bunches becomes such that defocusing occurs. In a bunched beam of ellipsoidal shape, defocusing takes place in both transverse directions against the external focusing, and also in the longitudinal directions acting against the phase stability mechanism. Such effects mainly take place at low velocity when the particles are very sensitive to electric fields; the situation improves at higher velocity.

Simple but very effective design equations for nonrelativistic transported or accelerated beams can be obtained by writing general root-mean-square (RMS) envelope equations for the beam as it is constrained in the accelerator channel. The external channel-focusing forces are offset by the internal space–charge forces in the beam. For a beam that is matched to the machine and does not oscillate in shape, one finds then the following relations:

$$\varepsilon_t = \frac{\sigma^t a^2}{N\beta\lambda} \quad \text{and} \quad \varepsilon_l = \frac{\sigma^l b^2}{N\beta\lambda} \qquad (4.1)$$

in terms of the transverse and longitudinal planes, where ε_t and ε_l are the transverse and longitudinal RMS emittance, a is the average transverse RMS beam radius, b is the physical RMS bunch length, and σ^t and σ^l are the phase advances of the oscillatory motion in phase space over the focusing period $N\beta\lambda$. Each of the σ terms can be expanded into two terms, one containing the machine parameters and the other containing the beam current and also the beam sizes a and b, explicitly. It is necessary to solve both parts of Eq. (4.1) simultaneously; the result is a beam that is ''matched'' to the RMS shape factors of the machine acceptance phase space. We know that RMS matching minimizes the growth of the effective emittances. It is important to note that these equations apply locally; thus, they should be applied not only as initial conditions at the entrance point of various sections of a machine but also with care to ensure that they remain true at every point.

In space-charge-dominated beam/accelerator systems, effective emittance growth could occur if the average energy in each of the coupled degrees of freedom is unequal [24, 25]. Requiring equality

produces another set of equations:

$$\frac{\varepsilon_l \sigma^l}{\varepsilon_t \sigma^t} = 1, \quad \text{which also requires}$$
$$\frac{\varepsilon_l}{\varepsilon_t} = \frac{\sigma^t}{\sigma^l} = b/a \tag{4.2}$$

If Eqs. (4.1) and (4.2) is not satisfied, the effective emittance will shift between the degrees of freedom (with conservation in six dimensions). The physical mechanisms for the redistribution have been clarified as a charge density distribution occurring at an abrupt parameter change within about one-quarter of the plasma oscillation period of the system, and in a smooth channel, a slower kinetic exchange toward the equipartitioned state. Systems satisfying Eqs. (4.1) and (4.2) are both matched and ''equipartitioned.'' Again, the condition can be applied locally and may be applied at injection and all along the machine to ensure minimum emittance growth.

Another problem can be the formation of a diffuse halo of particles around the beam bunch. Halos are to be avoided, as these particles may strike the walls of the LINAC. A halo forms when single particles interact with resonant driving forces in the channel [26]. If Eq. (4.1) is not satisfied, the beam is ''mismatched'' and this is a prime source of a halo.

A primary design requirement is to maintain matching and to avoid resonances between the beam and the accelerator structure parameters that can cause emittance growth and particle loss. The equilibrium, equipartitioned beam has no free energy to drive a resonance; so, even if a resonance is traversed, the growth rate will be zero. Although the equilibrium condition is desirable, as with all LINAC parameters, it may be relaxed or abandoned in favor of a particular objective; in this case, if a resonance area is encountered in the LINAC, it should be crossed quickly to avoid excessive emittance growth.

From these equations, many of the tradeoffs and scalings in machine design can be explored to first order. For example, an important result is that if high brightness is the goal, the maximum current achievable for a fixed transverse emittance strongly favors higher frequency LINACs. As a practical example, the low-β section of a modern 100 mA average current proton LINAC would operate at a frequency around 400 MHz. Fortunately, there are also other factors that are advantageous at higher frequencies, such as smaller size, ability to withstand higher electric fields without breakdown, and reduction in the amount of energy that must be stored in a resonant structure. Engineering factors such as heat removal in high duty factor or c.w. applications, on the other hand, argue for lower frequency; so a balance must be struck. Optimization for low and minimum energy beam loss is much more difficult, still not expressible in compact form, but considerably aided using the equipartitioned equilibrium and resonance avoidance.

Thus, in the final design for an application, many factors constrain the brightness that can be achieved. However, in summary, a few key requirements and procedures have been postulated that have major effects on achieving lower beam loss and/or higher brightness. These include using higher frequencies for ion LINACs, maintaining external focusing fields as high as possible, maintaining local matching per Eq. (4.1), the use of beam equilibrium and resonance avoidance locally per the simultaneous solution of Eqs. (4.1) and (4.2), and the important techniques for initial acceleration from the source as embodied in the RFQ LINAC for ions. Further research is required to fully understand and exploit

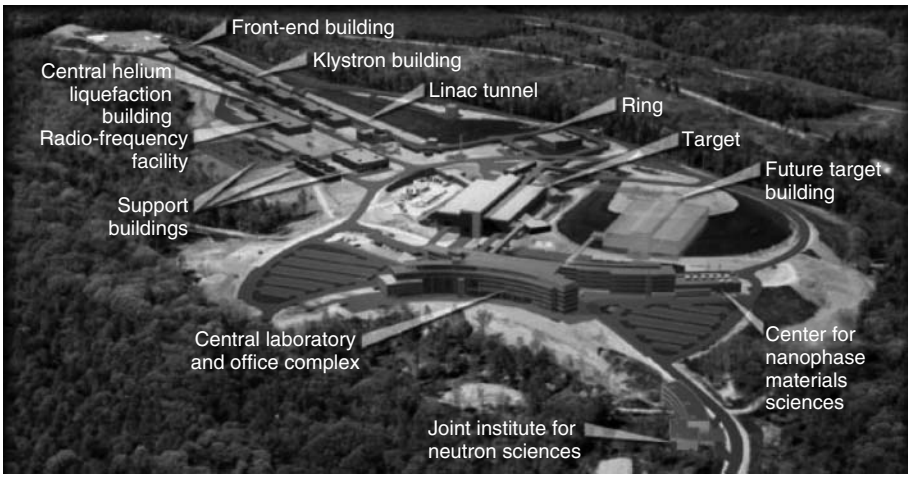


Figure 4.16 View of the ORNL SNS facility, indicating how the LINAC injects into the storage ring, from which the beam is extracted to the target.

these constraints, especially for controlling particles that tend to form a low-density ''halo'' around the beam and are likely to be scraped off, causing radioactivity problems for machine maintenance if their energy is above a few megaelectronvolts and to incorporate optimization algorithms.

4.4.6 Proton LINAC Applications

The LAMPF in New Mexico, producing an energy of 800 MeV with an average current of 1 mA of protons for medium-energy physics experiments, was completed in 1972; its energy is approximately 10 times and beam intensity 1000 times higher than any previous operational LINAC. Its length is close to a kilometer. It contains both 201.25 MHz Alvarez and 805 MHz side-coupled cavity structures. The SNS LINAC at Oak Ridge National Laboratory (ORNL), put into operation in 2005, has similar charge per bunch, and includes 402.5 MHz RFQ and DTL, and 805 MHz coupled cavity LINAC (CCL) and superconducting LINAC structures (SNSs, 2006) (Figure 4.16).

Ion LINACs are frequently used as injectors for storage rings or circular accelerators of higher energy. Frequently, H^- ions are accelerated for injection into a circular machine (e.g., SNS); a thin foil strips the two electrons, leaving the protons to circulate. An H^- LINAC just differs from a proton LINAC by the source and the sign of its extraction voltage.

Proton LINAC applications include drivers for intense neutron sources for physics and materials research such as the SNS, and the planned International Fusion Materials Irradiation Facility (IFMIF) for fusion-related materials development. Smaller proton and heavy-ion (mainly carbon) LINACs are being used for cancer therapy machines. Major new ideas were proposed at Los Alamos in 1990 [27] for closing of the fission power cycle by treatment of radioactive wastes, and for practical implementation of the thorium fuel cycle, using a high-intensity proton LINAC to drive a

subcritical nuclear reactor. These proposals were extended to the international community [28], and an active program EUROTRANS [29] is in progress to construct an R&D facility. The initial accelerator would be able to accelerate up to ~20 mA protons to 600–800 MeV; superconducting structures are ideal for a major part of the accelerator because a continuous beam is required. For such machines, the major design concerns are the minimization of beam losses at energies above a few megaelectronvolts, and strict prevention of sudden beam turn-off (discussed later).

More recently, advances in superconducting RF technology have been applied to the construction of proton LINACs. The main application is the SNS at the US ORNL [14]. The SNS is a second-generation pulsed neutron source, with a LINAC delivering 1 GeV protons at 1.4 MW average power to a liquid mercury target for neutron spallation with applications to the study of the structure of materials. The bulk of the LINAC is superconducting RF (0.812 GeV out of 1 GeV). In addition to greater efficiency and higher gradients for shorter lengths, the larger apertures allowable with SRF reduce beam loss, which is a critical concern because of activation. The superconducting LINAC utilizes two cavity types, for β of 0.61 and 0.81.

4.5 Heavy-Ion Linear Accelerators

4.5.1 Special Difficulties Relative to Ions

Under the term heavy ions, it is common practice to include all ionized atoms heavier than deuterons; accelerated ions usually range from lithium or carbon up to uranium.

Much of heavy-ion physics has been done in the past with the help of circular machines, usually cyclotrons, but there have also been several LINAC facilities, for instance, the heavy-ion linear accelerator (HILAC) at Berkeley – a big low-frequency Alvarez LINAC, later converted as an injector for the Lawrence Berkeley Laboratory Bevatron/Bevalac Facility (BEVALAC), a large heavy-ion synchrotron.

A center for heavy-ion physics was created at the Gesellschaft für Schwerionenforschung (GSI) laboratory in Darmstadt, Germany, and a heavy-ion LINAC complex called the universal linear accelerator (UNILAC) [30] was built. It is made of a Wideroe section, an Alvarez section, and an independent cavity section; more developments, such as special RFQs and H-type interdigital structure sections, have been added later, and the facility remains very active.

From this short description, it appears that a heavy-ion LINAC may be rather complex and differ appreciably from the proton machines.

A fundamental difference comes from the experimental requirements for a wide variety of ions with different masses and charge states. The charge-to-mass ratio (q/m) then differs from one ion to another and is always smaller than 1/2 (compared to the proton case) and often much smaller. In a given LINAC, where the geometry is fixed and only certain adjustments of the phase and amplitude of the accelerating fields are possible, only some combinations of q/m can be accelerated. For this reason, short, separately adjustable resonators are preferred to afford retuning for each separate ion beam.

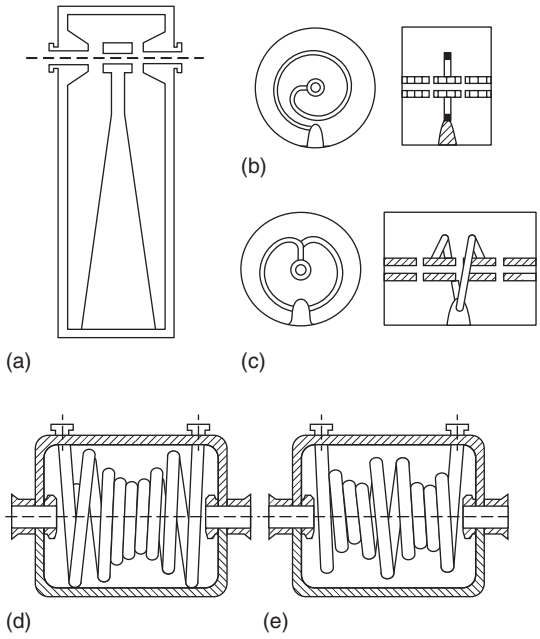


Figure 4.17 Superconducting cavities for heavy ions. (a) Quarter wave cavity: in this sort of coaxial resonator, the open end supports a drift tube and acceleration takes place across the gap. (b) Spiral ring resonator: to reduce the size, the resonant quarter wave line has been curved into a spiral. (c) Split ring resonator: to increase the energy gain, two spirals are put together; the gain should be doubled (in practice less), but the band width in velocity is reduced. (d) Half wave helix: the construction is easy, but there is a risk of mechanical vibrations and the field axis is not well defined. (e) Full wave helix: the rigidity is better, but the velocity bandwidth is smaller. Such cavities currently deliver >0.5 MV to the beam.

4.5.2
Accelerating Structures – Normal and Superconducting

Lower frequencies (20–100 MHz) offer more space for focusing; so structures such as the Wideroe structure (Section 4.2), or a variety of so-called H-type interdigital structures are used, some of which are represented in Figure 4.17. Independent short superconducting cavities are ideally suited, as they can be phased independently, affording complete flexibility in the velocity program. Fifty or one hundred cavities designed to have an optimum acceleration in several velocity steps (but with usable acceleration over a velocity range about a factor of 2 around the optimum) are arranged in succession. Their amplitude and phase are adjusted and controlled for the best possible acceleration of a chosen ion. The first machine of this type was made at Argonne near Chicago and is called *ATLAS* (Argonne tandem LINAC accelerator system). Its length has been progressively increased. Initially, as with the other machines of this type, it was used as a booster after an electrostatic tandem. Now, with proper low-β cavities ($\beta < 0.01$), beams from ECR sources at a relatively low voltage can be accepted directly [31].

4.5.3
Application

Heavy-ion accelerators have been mainly constructed for nuclear physics experiments; one of the largest facilities is the GSI in Darmstadt, Germany. More and more nonnuclear applications are appearing (solid-state physics, materials research, etc.), however, as also more and more industrial uses (implantation, surface treatments, ultrafiltration, etc.). An ambitious project that would reach the limits of the LINAC technology is a heavy-ion LINAC for inertial fusion. Fusion reaction could be induced in a small deuterium pellet

if bombarded by a short and extremely intense bunch of energetic heavy ions (10 GeV, several kiloamperes with a 10 ns pulse length). It is clear that a special type of accelerator would be needed to reach that goal (Section 4.6).

4.6 Other Types of Linear Accelerators and New Methods of Acceleration

4.6.1 Induction LINAC

The induction LINAC is a candidate for short-pulse applications. Long induction LINACs are built up from a series of modules of the type shown in Figure 4.18. They behave as same as pulse transformers with the beam as one winding, and are accurately phased to produce acceleration (and possibly bunching) exactly at the time a bunch of particles crosses the gaps. The method is, in fact, the modern version of the original Ising proposal (Section 4.2). Induction LINACs are inherently pulsed devices of pulse length ~50 ns. Longer pulses are limited by the energy that can be stored in the pulser cores. Repetition rates of up to a few kilohertz are being developed using special magnetic modulators. This device can accelerate extremely high intensities. The largest electron induction LINAC built is the Advanced Test Accelerator at Livermore, a 50 MeV machine that has accelerated up to 10 kA in 50 ns pulses [32]. The induction LINAC is also well suited for the heavy-ion fusion project [33] mentioned in the previous section (its length would be several kilometers).

4.6.2 Alternating-Phase-Focused LINACs

As noted in Section 4.4.2, the RF field can be used for both an accelerating and the transverse focusing field, avoiding the need for expensive magnetic elements. Except for the RFQ, it has been very difficult in general to design such LINACs, but recently a new practical method has been found [34].

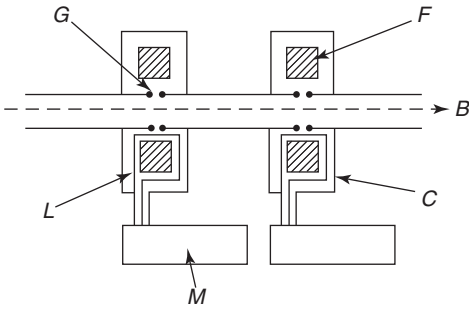


Figure 4.18 Induction LINAC modules. Along the beam path *B*, a succession of cylindrical cavities *C* are installed with an accelerating gap *G*. Inside these cavities, toroidal ferrite cores *F* are excited via a coupling loop *L* by the pulse-forming networks of the modulators *M*. The cavities behave as same as a transformer and the pulse voltage applied on the loop appears on the accelerating gap. Timing of the modulators must be accurately controlled so that the pulses are applied very correctly at the time the bunch crosses the gaps.

4.6.3 New Methods of Acceleration

Even in pulsed mode, microwave structures are limited to gradients of the order of 100 MV m^{-1}, and as the need for higher energy beams and compact sources has grown, strong interest has developed in using the intense fields generated by lasers or plasma waves to accelerate beams. For example, plasma perturbations can yield field gradients of order 100 GV m^{-1}. A device based on this technology could shrink the length of LINACs by three orders of magnitude! To generate these plasma waves, either a high-powered laser or a particle beam would be used to set up the plasma oscillation (laser wakefield or plasma wakefield concept). An alternative approach would use gratings, ministructures such as photonic bandgaps, or exotic materials (e.g., active or negative index of refraction) to couple a laser beam to a particle beam. Also, a particle beam passing through the alternating fields of an undulator can couple to a copropagating laser beam for acceleration in a device called an *inverse FEL*. Limitations of these ideas include structure damage, plasma instabilities, accelerating high currents, and beam quality maintenance. However, there have been significant experimental advances that hold the promise for future applications. See [35] for more details and references on these new methods of acceleration.

4.7 Engineering Aspects

The full spectrum of engineering disciplines is applied to accelerator technology; a close, multidisciplinary approach with the physical, mathematical, material, and computer sciences and technologies is essential. Deciding on trade-offs between the possibilities and constraints of the many aspects constitute a large part of the design, construction, and operation of accelerators. Broad engineering categories applied to most accelerator systems are described in the following sections.

4.7.1 Accelerator Structure Design

The structures in which beams are accelerated must efficiently provide precisely the correct fields and are mechanically complicated vacuum vessels that must be thermally stabilized, structurally rigid, extremely clean for sustaining high electromagnetic fields, precisely aligned, and maintainable. They must be structurally stiff to hold their shape under vacuum and component loading. RF accelerating structures are copper clad or copper plated, or of special material such as niobium if superconducting. Structures and beamlines must be constructed of materials that are hard-vacuum compatible because they must maintain a vacuum quality typically 10^{-6} to 10^{-9} Torr. Fabrication techniques used in accelerators are diverse. High-precision machining is done on numerically controlled milling machines and by other specialized fabrication techniques such as electron discharge machining (EDM). Specialized joining and forming techniques include brazing, electron beam and laser welding, explosive bonding of dissimilar metals, and explosive hydroforming. Very good mechanical tolerances and cooling systems are needed. Complex brazing is used in the latest vane-type RFQs (Figure 4.19) and side-coupled LINACs (Figure 4.20). The highest power RFQ built to date, the

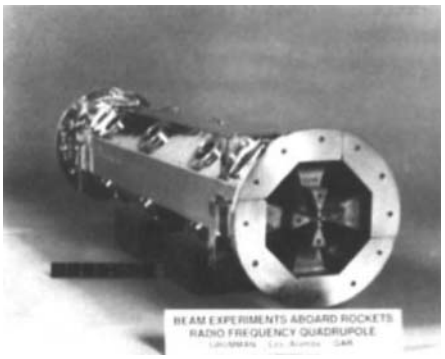
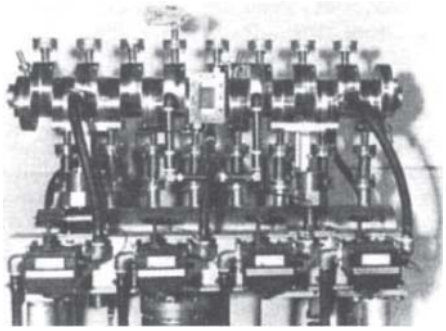
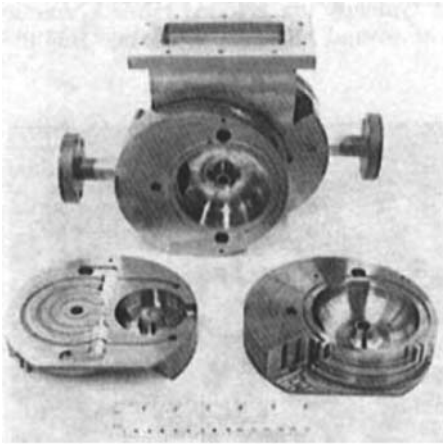


Figure 4.19 An RFQ manufactured in four segments.



(a)



(b)

Figure 4.20 (a) A side-coupled LINAC cell before brazing, and (b) a section of a 2400 MHz c.w. side-coupled LINAC.

Los Alamos Low-Energy Demonstration Accelerator (LEDA) c.w. 100 mA, 6.7 MeV proton LINAC (Figure 4.21), almost disappears in its support structure, which contains RF power, water-cooling, and vacuum connections.

Rapid advances in the performance of superconducting cavities have made RF superconductivity a key technology for accelerators. They are especially suitable for applications that require c.w. or long-pulse accelerating fields of many megavolts per meter. The RF surface resistance of an SRF cavity is many orders of magnitude less than copper, and even allowing for the inefficiencies of heat removal at cryogenic temperatures, a net gain in efficiency of 100 can be easily obtained. (Since superconducting Cooper pairs exhibit inertial effects in AC fields, they do not completely cancel the electric fields in a superconducting material. Nonsuperconducting electrons from ''ionized'' cooper pairs or impurities above absolute zero can respond to these residual fields and dissipate small amounts of energy.) Gradients of up to $20\,\mathrm{MV\,m^{-1}}$ and Qs approaching 10^{10} are practical today, and further improvement can be expected. SRF structures are

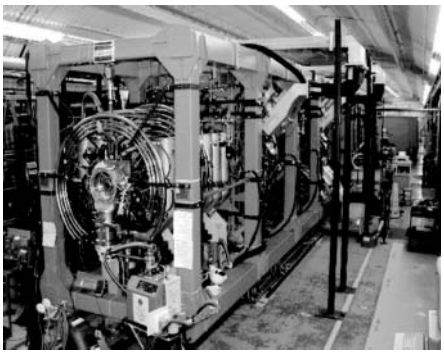


Figure 4.21 Low-Energy Demonstration Accelerator (LEDA) RFQ at Los Alamos.

shaped to control resonant electron effects (multipactoring). Control of contamination by macroparticles by assembly in clean rooms and careful surface preparation has eased the limits of field emission and thermal breakdown.

4.7.2 Magnetics

Complex magnet systems provide magnetic steering and focusing elements that control the position, size, and transport characteristics of particle beams. Modern transport systems use and control magnetic properties as high as fifth order in the optics. Many of the magnets used in accelerator systems are physically short because of packaging constraints or must deviate from ideal geometry because of the field coils. The fields must be measured, often to an accuracy of better than 1 part in 10^4, using magnetic mappers supported by sophisticated instrumentation and computers. Radiation-hardened magnets are also required.

Permanent magnets are quite applicable to high-frequency DTLs and considerably simplify the drift-tube design because they require no power or cooling. Small, accurate quadrupoles (with 1.2 cm bores operating at 18.33 kg cm^{-1} gradients) have been built and tested. They too can be built to a size appropriate for a beam-transport line. New design techniques can be used for adjustable-field multipole magnets and for magnets in which permanent magnets and normal-magnet steels are mixed. Recent experimental results on rare-earth cobalt materials show, however, that serious demagnetization occurs in some designs at 10^9 rad dose levels. Thus, although permanent-magnet technology grows rapidly in its own range of application, use in high radiation environments is precluded at present.

4.7.3 Radiofrequency Engineering

The complex electromagnetic requirements of the accelerator structures must be met, and high-power RF driver systems provided for them. Radiofrequency power systems typically use gridded tubes to frequencies of around 200 MHz, with solid-state amplifiers, klystrons, and inductive output tubes (IOTs) at higher frequencies. The power requirements can be very large; for example, a 40 MeV, 125 mA IFMIF RF LINAC requires 5 MW just for the beam. LINACs are divided into sections sized to the power sources; this ranges from single cells to tanks containing many cells. To maintain synchronism, all the sections must be tightly controlled to

within typically 1% in amplitude and synchronized to $\pm 1^{\circ}$ in phase, by the automatic control of a reference oscillator. Bunch compression scenarios can demand more stringent RF control by up to a factor of 100. Development of efficient RF power sources is an urgent requirement for future applications.

4.7.4 Availability and Maintainability

LINACs for physics research machines have almost always been cutting-edge technology, always breaking new ground, and on a very tight budget. Usually a loss of beam for some period, from missing pulses to some hours, could be tolerated without any problem by the experiments. Thus, accelerator facilities to date have considered availability only as "best-practice" and not as a top-level specification. The users of research facilities such as SNS may come for short periods of a few days; so their productivity must be ensured; SNS has availability goals built into its development schedule over the next decade. For IFMIF, the question of beam trips resulted in the choice of building at least two parallel LINACs, in modules of 125 mA each, greatly increasing the probability that at least one would always be operating. Also, the material samples would not be damaged because the thermal cooling could be adjusted quickly. For transmutation, however, a beam trip would induce thermal stress on the subcritical core, and the number of unwanted "beam trips" should not exceed a few per year, a requirement that is several orders of magnitude above usual performance. For these and other future facilities, availability becomes a top-level specification and a major design challenge, as accelerators contain thousands of components, and often the failure of one will require the beam to be shut off.

4.7.5 Computers

Accelerators, in general, have long provided a major impetus for the development of advanced computer systems. Design teams and facilities use the full panorama of computers, from supercomputers to simulate particle motion and electromagnetic fields in the accelerator and magnet systems and on-line real-time computer control systems with distributed processing, to personal computer tools. Modern, fully integrated computer control using instrumentation and automatic control bring accelerator commissioning and operation under strict and very effective management, as witnessed by the recent commissioning of the SNS.

4.7.6 Wide-Bandwidth Electronics and Instrumentation

The pulsed nature of particle beams requires wide-bandwidth electronics, some at higher power, and various diagnostic instrumentation to monitor machine- and beam-operating parameters and control them. Most of these tasks can now be handled digitally.

An example of high-power, high-bandwidth electronics is the pulsed beam chopper shown in Figure 4.22. This device deflects part or all of a 750 keV H^- beam in the LAMPF injector transport line. A pulse rise time of <5 ns and pulse lengths from 15 ns to 750 μs are required at repetition rates up to 4.7 MHz. Chopping is accomplished by launching along this structure a pulse that propagates at the ion velocity of the beam, $1.2\,\mathrm{cm\,ns^{-1}}$. The very

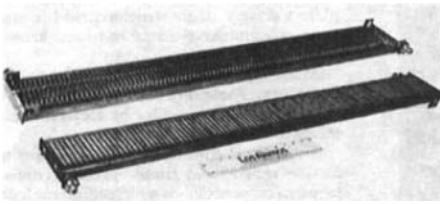


Figure 4.22 Deflector plate assemblies for a beam chopper.

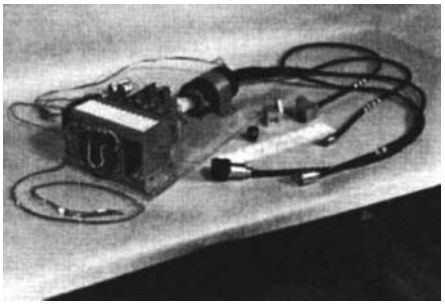


Figure 4.23 Four-armed fiber optics beam-sensing system.

fast high-power electronics uses planar triodes and vertical metal-oxide semiconductor fields-effect transistors (VMOS FETS) in a parallel cascade arrangement to deliver both positive and negative 600 V pulses to the opposing coaxial plate structures.

Diagnostic instruments must often measure beam parameters without intercepting the beam. This is essential with high-intensity beams, which would destroy intercepting materials. Beam parameters of interest include current, energy, density distribution, emittance, and centroid position. Beam density distribution is sometimes measured by optical sensors that view through optical ports the light generated by beam scattering from residual background gas. The optical system uses four-armed fiber-optic image scopes (Figure 4.23) to transmit four-beam profiles from a given beam tube/optical-port array to a microchannel light intensifier and then onto a pair of linear photodiode arrays, from which the light intensity is digitized. The use of a number of these ports distributed along the beam tube allows beam size and trajectory information to be reconstructed into beam emittance. The emittance reconstruction codes are executed in a microcomputer along with beam-transport dynamics codes and several display routines. Longitudinal beam phasing and shape information may be sensed using a variety of high bandwidth capacitive pickups, such as shown in Figure 4.24.

4.7.7 Automatic Controls

Automatic controls are used widely in accelerator technology. Examples include the following:

- control of parameters in complex subsystems such as the injector;
- cooling-system temperature control;
- RF amplitude and phase control;
- beam trajectory control.

Control of water temperature is fundamental to high-power accelerator tasks because operating temperature affords a vernier adjustment to the structure's RF.

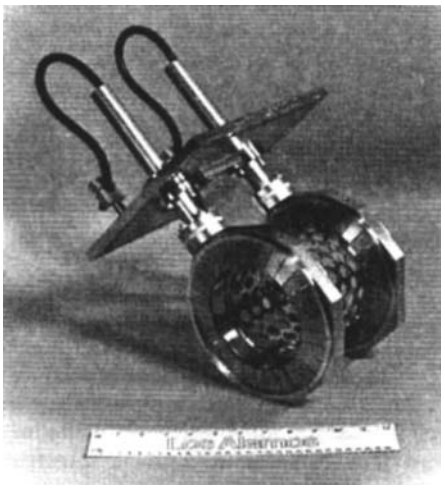


Figure 4.24 Capacitive pickup probes for time-of-flight measurement.

The proper RF field phase and amplitude are crucial to operation and must be controlled to within $\pm 1^{\circ}$ and $\pm 1\%$, typically, under all possible transient and steady-state conditions. In some cases, multiple amplifiers drive a single structure; field control of these complex systems has been accomplished.

Many types of errors can affect the beam trajectory. These must be found and corrected or compensated for by correction elements located along the beamline.

A variety of controllers are used, covering a wide range from simple process control loops, to complex loops with feed forward, domain switching depending on system state, and adaptive and nonlinear systems; modern implementation is digital with computer processing. Work is currently being done using expert system and neural network techniques.

4.7.8 Civil Engineering

The accelerating structures and magnets defining the beam trajectory must be aligned relative to each other to exquisite tolerances (a few mils; down to micrometers for linear colliders) and also absolutely over the whole facility, again to mil-level tolerances. The alignment must be maintained over the lifetime of the facility. Civil engineering procedures must therefore take many factors into account, such as thermal cycling on stresses in the buildings, placement and rearrangement of heavy shielding, vibrations from utilities, and replacement procedures when accelerator parts are maintained. The surveying techniques incorporate the latest laser- and computer-supported techniques.

4.8 Concluding Remarks

Details on current developments can be found in the proceedings of the biannual US Particle Accelerator Conferences (PACs) published for many years as part of the IEEE Transactions on Nuclear Science, and recently as IEEE Conference Proceedings. Also biannually, LINAC Conferences have been held with proceedings published by a host laboratory. A third biannual conference series is ''Application of Accelerators in Research and Industry.'' A

European Particle Accelerator Conference (EPAC) series started in 1988, and is now a triannual feature, rotating with the Asian Particle Accelerator Conference (APAC). Many of these proceedings are available at [36]. Russia also has a long-standing biannual particle accelerator conference series. These conference proceedings constitute the bulk of the accelerator technology literature, along with many laboratory internal reports and workshop proceedings (a number of these have been published as AIP Conference Proceedings, e.g., 156, 177, 184, 193, and 480). There is an active USPAC particle accelerator school and the CERN Accelerator School series. Since the previous (2007) edition, when conferences were almost the only source of accelerator literature, the American Physics Society has established its Division of Physics of Beams, and the Physical Review has a special topic component on Accelerators and Beams – PRSTAB.

Glossary

Acceptance: The maximum beam emittance that can be accepted without loss, through a transport-focusing channel or an accelerating system, determined by machine parameters, independent of the actual beam.

AG focusing (alternating-gradient focusing): A beam-focusing system made of a succession of focusing and defocusing lenses, normally of quadrupole type. Such a system provides net focusing.

Alvarez structure: An accelerating structure for protons, of the drift-tube type, invented by Luis Alvarez.

BBU (beam breakup): Intensity-dependent transverse phenomena in relativistic linear accelerators that limit the current in intense machines (Section 4.3.3).

Beam loading: Power absorption in RF accelerating cavities by the accelerated beam compensated by the application of RF power and control of the accelerator fields.

Bunch: The group of particles accelerated per RF period.

Disk-loaded structure: See Iris-loaded structure.

Drift tube: A tube connected to an RF voltage, through which the beam can drift. When a beam bunch is inside a drift tube, it drifts without being submitted to any RF field.

DTL (drift-tube LINAC): LINAC structure based on the use of drift tubes, for example, the Alvarez type.

Emittance: The positions of individual particles in an accelerated bunch can be expressed by three physical coordinates and their conjugate momenta. The areas occupied by the projections in the three-phase planes (see **Phase space**) are called the *beam emittances*; there are two transverse emittances (often equal in LINACs and expressed in terms of divergence angle rather than in terms of transverse momentum), and a longitudinal emittance (often expressed in terms of energy spread and time duration).

Equipartition: A balanced division of field energy between the three emittances of the beam that guarantees the best stability; the term is used similarly in the case of thermal equilibrium in a gas.

Focusing: The action of concentrating particles to form a beam, similar to focusing of light in optics.

GSI: Gesellschaft für Schwerionenforschung, Darmstadt, Germany.

Gap, accelerating gap: Interval between drift tubes where a strong RF field is concentrated, accelerating the particles of a bunch at the time they cross the gap.

IFMIF: International Fusion Materials Irradiation Facility, a complex of two 125 mA, 40 MeV, c.w. deuteron LINACs providing beams to a molten Li target, generating neutrons with a fusionlike spectrum for development of neutron-resistant materials for fusion reactors.

Iris-loaded structure (or disk-loaded): An accelerating structure for relativistic particles made of a circular waveguide loaded by irises to slow down the phase velocity of the RF wave, making it synchronous with the beam.

LAMPF: Los Alamos Meson Physics Facility, commissioned 1972, with 800 MeV, 1 mA average current proton LINAC, later known as LANSCE, Los Alamos Neutron Science Center.

LINAC: Linear accelerator.

Matching: The machine parameters define a preferred region in phase space (acceptance) for transport or acceleration. If the beam particle phase space distribution has a congruent orientation, called *matched*, beam passage will be smooth, without oscillations. A *mismatched* beam will oscillate in position or shape, which may cause loss of quality or actual particle loss.

Phase space: The six-dimensional space of physical coordinates and conjugate momenta.

Photocathode: A photocathode is a negatively charged electrode, often with a coating to improve emission, that generates electrons through photoemission when struck by a laser beam.

Quadrupole lens: Lenses, usually magnetic, with four poles around the beam: two of N polarity diametrically, and two S in between. Such a lens focuses in one transverse direction and defocuses in the other.

RFQ (radiofrequency quadrupole): An accelerating structure for low-velocity particles where the electrodes produce an RF field that is the sum of a transverse, quadrupolar, AG focusing field and an accelerating field.

SCL (side-coupled LINAC): A normal-conducting, standing wave LINAC structure used for accelerating electrons and protons when the RF pulse length is longer than a few microseconds, or for c.w. operation. The coupling between successive cells is arranged to take place through side cavities alternately on one side and then the other. The accelerating and side cavities resonate at the same frequency, resulting in a $\pi/2$ mode operation having maximum group velocity and field stability.

SNS: Spallation Neutron Source, Oak Ridge National Laboratory, USA, commissioned in 2006 [37].

Space–charge effects: Nonrelativistic particles in a bunch, all charged with the same sign, repel each other, tending to force the dimensions to grow, and requiring more focusing. This also leads to possible instabilities and couplings, and to the formation of a halo of particles around the bunches (Section 4.4.5).

Standing wave cavity: A hollow structure bounded by a conductive wall that supports a resonant electromagnetic mode when excited by an external RF source. The mode pattern is typically designed to accelerate a charged particle passing along the axis, although deflecting modes are also used in some accelerator applications.

Unit of energy: The energy of particles is expressed in electronvolts (eV), energy

given to a unit charge (charge of the electron) through a potential difference of 1 V. Higher energies are expressed in kiloelectronvolts, megaelectronvolts, gigaelectronvolts, and teraelectronvolts.

Wakefield: The electromagnetic field left by a charged particle when it passes a discontinuity in its environment. For example, such discontinuities may be resonators or gaps, bellows, or steps in a conductive wall boundary. Resistive wall boundaries can also generate wakefields, which describe the loss of beam energy as image currents are driven in the wall.

References

1. Ising, G. (1924) Prinzip Einer Methode Zur Herstellung Von Kanalstrahlen Hoher Voltzahl. *Ark. Mat. Astron. Fyzik*, **18**, 1.
2. Wideroe, R. (1928) Über ein neues Prinzip zur Herstellung hoher Spannungen. *Arch. Electotech.*, **21**, 387.
3. Sloan, D.H. and Lawrence, E.O. (1931) *Phys. Rev.*, **38**, 2021–32.
4. Ginzton, E.L. *et al.* (1948) A linear electron accelerator. *Rev. Sci. Instrum.*, **19**, 89.
5. Slater, J.C. (1948) The design of linear accelerators. *Rev. Mod. Phys.*, **20**, 473.
6. Leemann, C., Douglas, D., and Krafft, G. (2001) The continuous electron beam accelerator facility. *Annu. Rev. Nucl. Part. Sci.*, **51**, 413.
7. Barish, B. (2006) The global design effort for an international linear collider. Proceedings of the 2006 European Particle Accelerator Conference, Edinburgh, Scotland.
8. Free Electron Laser (2007) The World Wide Web Virtual Library: Free Electron Laser Research and Applications, *sbfel3.ucsb.edu/www/vl fel.html* (accessed 28 February 2013).
9. Merminga, L., Douglas, D., and Krafft, G. (2003) High current energy recovery linacs. *Annu. Rev. Nucl. Part. Sci.*, **53**, 387.
10. Alvarez, L.W. *et al.* (1955) Berkeley proton linear accelerator. *Rev. Sci. Instrum.*, **26**, 111–133210–219.
11. Courant, E.D., Livingston, M.S., and Snyder, H.S. (1952) The strong-focusing synchrotron – a new high energy accelerator. *Phys. Rev.*, **88**, 1190.
12. Blewett, J.P. (1952) Radial focussing in the linear accelerator. *Phys. Rev.*, **88**, 1197–1199.
13. Hereward, H.G. *et al.* (1956) Proceedings of the Conference on High Energy Accelerators, CERN, Genevap, p. 167.
14. Holtkamp, N. (2006) Status of the SNS linac: an overview. Proceedings of the 2006 European Particle Accelerator Conference, Edinburgh, Scotland.
15. Sinclair, C.K. *et al.* (2007) Development of a high average current polarized electron source with long cathode operational lifetime. *Phys. Rev. ST Accel. Beams*, **10** (2), 023501.
16. Farkas, Z.D. *et al.* (1975) Recent progress on SLED, the SLAC energy doubler. *IEEE Trans. Nucl. Sci.*, **NS-22** (3), 1299–1302.
17. Ahearne, J.T. (1980) RFQ is Alive and Well . . . Atom (Los Alamos Scientific Laboratory (LASL)) (Jul./Aug.), Vol. 17, No. (4).
18. Teplyakov, V.A. (2006) The first CW accelerator in USSR and a birth of accelerating field focusing. EPAC2006 Accelerator Prize, Edinburgh, Scotland, July 26–30, 2006.
19. Kapchinskij, I.M. (1970) *IEEE Trans. Nucl. Sci.*, **26**, 3462.
20. Stokes, R.H., Crandall, K.R., Stovall, J.E., and Swenson, D.A. (1979) RF quadruple beam dynamics. *IEEE Trans. Nucl. Sci.*, **NS-26**, 3462.
21. Potter, J.M., Williams, S.W., Humphry, F.J., and Rodenz, G.W. (1979) Radio frequency quadrupole accelerating structure research at Los Alamos. *IEEE Trans. Nucl. Sci.*, **NS-26**, 3745.
22. Schempp, A. (2002) Developer of more than 40 different RFQs. New development in high-duty cycle high-current RFQs. Proceedings of LINAC2002, Gyeongju, Korea, p. 539.
23. Knapp, E.A. *et al.* (1968) *Rev. Sci. Instrum.*, **39**, 979–91.
24. Jameson, R.A. (1981) Beam intensity limitations in linear accelerators. *IEEE Trans. Nucl. Sci.*, **28** (3), 2408.
25. Jameson, R.A. (2007) Equipartitioning in linear accelerators. Proceedings of the 1981 Linear Accelerator Conference, Report LA-9234-C, Los Alamos National Laboratory, February 1982, ORNL/TM-2007/001, p. 125.

26. Jameson, R.A. (1994) *AIP Proceedings of the 1994 Joint US-CERN-Japan International School on Frontiers of Accelerator Technology*, World Scientific, ISBN: 981-02-2537-7, pp. 530–560.

27. Bowman, C.D. *et al.* (1992) Nuclear energy generation and waste transmutation using an accelerator-driven intense thermal neutron source. *NIM Phys. Res.*, **AA320**, 336–367.

28. Jameson, R.A. Compiler, (1991) Specialist Meeting on Accelerator Driven Transmutation Technology for Radwaste and other Applications, June, 24–28, 1991, Saltsjöbaden, Stockholm, Sweden, spons. by LANL and Swedish National Board for Spent Nuclear Fuel (SKN), LA-12005-c, SKN Rpt. No. 54.

29. Knebel, J.U. (2006) EUROTRANS: European Research Programme for the Transmutation of High-Level NuclearWaste in an Accelerator Driven System. FI6 W-CT-2005-516520: Integrated Project on European Transmutation (EUROTRANS), Sixth Framework Programme.

30. Bohne, D. (1976) The UNILAC, development and present status. Proceedings of the Linear Accelerator Conference, Chalk River, Vol. AECL-5677, p. 2.

31. Bollinger, L.M. 1986 Linear Accelerator Conference, Stanford, California. CEBAF Report 89–001, p. 237.

32. Briggs, R.J. (1986) Linear induction accelerators. Proceedings of the 1986 Linear Accelerator Conference, SLAC-303, CONF-860629, UC-28, Stanford, California.

33. Fessenden, T. 1982 Linear Accelerator Conference, GSI 84-11, Darmstadt, Germany, p. 485.

34. Jameson, R.A. (2012) *Practical Design of Alternating-Phase-Focused (APF) Linacs* submitted for publication to NIM.

35. Blumenfeld, I. *et al.* (2007) Energy doubling of 42 GeV electrons in a metre-scale plasma wakefield accelerator. *Nature*, **445**, 741–44.

36. JACoW (2006) Joint Accelerator Conference Website, *www.JACoW.org* (accessed 28 February 2013)

37. Neutron Sciences (2006) The SNS Facility–Neutron Sciences, *neutrons.ornl.gov* and *www.sns.gov* (accessed 28 February 2013).

Further Readings

Accelerators and Beams (2012) *Tools of Discovery and Innovation*, Division of Physics of Beams of the American Physical Society, November 2012, *www.aps.org/units/dpb* (accessed 28 February 2013).

Bollinger, L.M. *et al.* (1984) Proceedings Linear Accelerator Conference, Darmstadt, Germany, Report GSI 84–11, p. 217.

Grunder, H. (1988) The continuous electron beam accelerator facility. Proceedings 1988 Linear Accelerator Conference, Williamsburg, Virginia, CEBAF Report 89–001, p. 3.

Hanna, S. (2012) *RF Linear Accelerators for Medical and Industrial Applications*, Artech House, Boston, MA March 31, 2012.

Lapostolle, P.M. *et al.* 1963 International Conference on High Energy Accelerators, Dubna, Atomizdat, Moscow, p. 513 (U.S. AEC Translation pp. 672–682).

Padamsee, H., Knobloch, J., and Hays, T. (1998) *RF Superconductivity for Accelerators*, John Wiley & Sons, Inc., New York.

Stokes, R.H. and Wangler, T.P. (1988) *Radiofrequency Quadrupole Accelerators and their Applications*, Annual Review of Nuclear and Particle Science, Vol. 38, Annual Reviews, Inc., Palo Alto, CA.

Vladimirskij, V. (1956) *Prib. Tekh. Eksp.*, **3**, 35.

Watson, J.M. (1985) The Los Alamos free-electron laser. *IEEE Trans. Nucl. Scin*, **NS-32** (5), 3363 (Proceedings 1985 Particle Accelerator Conference).

Wiedemann, H. (2007) *Particle Accelerator Physics*, Springer, London.

Wilson, P.B. (1989) Introduction to wakefields and wake potentials physics of particle accelerators. AIP Conference Proceedings 184, New York, p. 526.

5
Exotic Nuclear Beam Facilities

Hans Geissel, Mark Huyse, Gottfried Münzenberg, and Piet Van Duppen

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

5.1 Introduction

The knowledge of atomic nuclei, the building blocks of elementary matter, is fundamental for the understanding of our universe. Stable nuclei have a detailed balance between their constituents, protons, and neutrons, determined by the nuclear force and the Coulomb repulsion, and their interference. Radioactive, short-lived nuclei are called *exotic nuclei*. They are characterized by extreme proton-to-neutron ratios, compared to stable isotopes. Exotic nuclei reveal novel properties of the strong interaction. The investigation of exotic nuclei extends our knowledge of nuclear structure and reactions and explores the nuclear properties at the limits of the existence of matter as well as the creation of the chemical elements in stars [1–5].

Recent progress in studying exotic nuclei has been achieved by new accelerator developments, powerful separators possibly combined with high-resolution spectrometers and storage devices, and efficient detector systems [6]. All these components are essential parts of a modern exotic nuclear beam facility. Modern accelerators provide high-intensity energetic projectiles to create the exotic nuclides far away from the valley of stable isotopes. Besides the high intensity, the maximum kinetic energy and the time structure are decisive requirements for the accelerated ion beams. The isotope separators have to be fast, selective, efficient, and sensitive down to separate nuclei produced with rates of a few ions per day or even per week from the overwhelming amount of unwanted isotopes. From these challenges, it directly follows that the detector system must be capable of recording the properties of exotic nuclides with high efficiency, sensitivity, and resolution. These key components applied in the research of exotic nuclei have been substantially improved in the past decade such that new isotopes, new decay modes, and other novel nuclear properties have been discovered.

An overview of the experimentally known exotic nuclides and the landmarks for the investigations is illustrated in the chart of nuclides in Figure 5.1. The stable nuclides form a narrow band (black squares) between the calculated lines of nuclear particle stability, the proton and neutron driplines, beyond which the nuclei are unbound and decay by spontaneous particle emission. For the heaviest elements, the onset of spontaneous fission sets the corresponding borderline. The limits of the discovered isotopes are indicated in the figure by the outskirts of the chart of nuclides (grey area). Detailed experimental

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

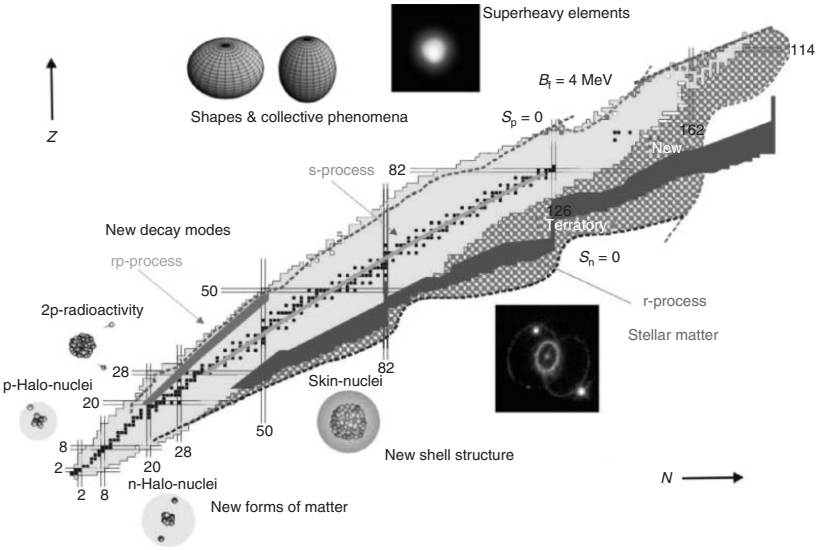


Figure 5.1 Research goals with exotic nuclei, and limits of nuclear matter. The experimentally discovered isotopes Ref. [108] are plotted in the chart of nuclides (grey area) enclosed by the calculated nuclear driplines ($S_p = 0$, $S_n = 0$) and the fission barrier ($B_f = 4$ MeV, dashed lines). The stable nuclides are represented by black squares and form the valley of stability. The astrophysical synthesis paths, the *r*-, *s*-, and *rp*-path are also shown demonstrating the synergy between nuclear and astrophysics. New properties of exotic nuclides, such as decay modes, nuclear matter distributions (halo, skin, cluster formation), and the large area of neutron-rich nuclei (New Territory) waiting to be discovered before the dripline is reached, are indicated. The horizontal and vertical lines in the chart of nuclides represent the magic number of the nuclear shell structure.

information on the structure and decay properties is available in the neighborhood of the stable nuclei. Only theoretical extrapolations can be made for the isotopes far off stability. Therefore, exotic nuclei are the most interesting species to test and to improve the understanding of basic nuclear properties [7].

In this Chapter, we discuss the basic features and modern developments of the production, separation, and investigation of exotic nuclei. We present experimental developments, future directions, and highlights in the field but have no intension of covering all the existing and planned facilities or activities. Open questions and future developments are discussed briefly.

5.2 Production of Exotic Nuclei

The discovery of the isotopic nature of chemical elements was made by J.J. Thomson [8] and F.W. Aston [9]. They discovered and measured systematically the stable isotopes of almost all chemical elements by applying a mass spectrograph [9]. Singly charged atoms extracted from an ion source were mass analyzed by subsequent electric and magnetic fields, which revealed that most stable elements are a mixture of nuclides, with the same number of protons (Z) but with different numbers of neutrons (N). Indeed, these pioneering experiments used an instrumentation that represents the roots of the first exotic nuclear beam facility built many decades later [10]. From Aston's mass measurements, the first knowledge of the nuclear binding energy and a basic theoretical nuclear model [11] were deduced. Soon it became clear that a deeper understanding of nuclear matter can be achieved only if the investigations go beyond the stable nuclides and include species that cannot be found on Earth but have to be produced in nuclear reactions [12] (Figure 5.2).

5.2.1 Reactions

5.2.1.1 Fusion

Near and at the Coulomb barrier, central collisions of heavy ions lead to complete fusion, the coalescence of target and projectile nuclei. The fusion reaction is a two-step process, where in the first step, an excited intermediate nucleus is formed. This compound nucleus (m_{cn}) equilibrates and may deexcite via neutrons, protons, α-particles, cluster evaporation, or fission, besides the emission of γ-rays. The amount of excitation energy is determined by the Q-value of the reaction and the kinetic energy available in the center-of-mass system. The *Q-value* is defined as the mass difference of the collision partners and reaction products:

$$Q = (m_1 + m_2 - m_{cn})\, c^2 \qquad (5.1)$$

where m_1 and m_2 are the masses of the projectile and target nuclei, respectively.

In cold fusion reactions, the excitation energies are (10–20) MeV, which means that only one or two nucleons will be emitted during the de-excitation phase of the intermediate compound nucleus. The excitation energies in hot fusion are much higher [30–50] MeV. Therefore, the average number of emitted nucleons is higher, about 3–5. In Figure 5.3, measured cross sections for the formation of superheavy elements (SHEs) in cold and hot fusion reactions are presented. The small energy window for the production of a certain isotope by heavy-ion fusion is of the order of 10 MeV. The cross section (unit $1\ \mathrm{b} = 10^{-24}\,\mathrm{cm}^2$) is a measure of the

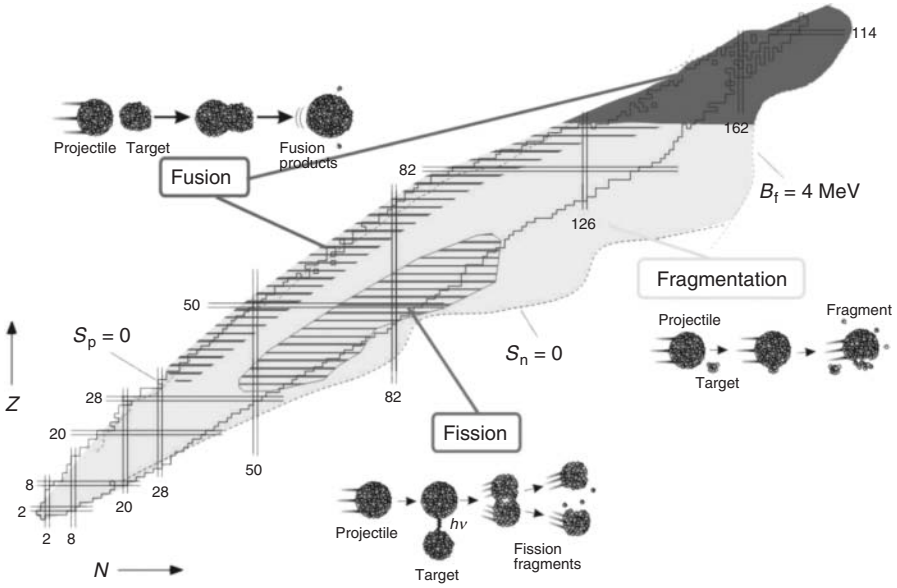


Figure 5.2 Nuclear reactions [12] to create exotic nuclei. Fission and fragmentation (spallation) are used to reach the most neutron-rich isotopes up to uranium, whereas the heaviest elements have been synthesized with cold and hot fusion reactions. Fission and fragmentation (spallation) can be induced by different projectiles (n, H to U, e^-) or by high-energy photons. Certain regions of the nuclear chart, for example, the heavy neutron-rich nuclei close to the neutron dripline, are difficult to be produced with stable beam and target combinations. However, the use of radioactive neutron-rich projectiles has recently been proposed to study the species that are probably formed in stars via neutron capture and subsequent β-decay at extremely high neutron densities.

probability to create a selected isotope. Using present accelerators and separators, a cross section of 10^{-12} b means only one atom detected in a few days. The excitation function, that is, the cross section as a function of the bombarding projectile energy, determines basically the optimum target thickness, which takes into account the atomic slowing down of the projectile before it fuses with a target nucleus. The heaviest neutron-rich isotopes can probably only be formed via reactions with radioactive target and projectile nuclei.

5.2.1.2 Multinucleon Transfer

The transfer of nucleons is a binary reaction where protons, neutrons, or clusters are exchanged between the colliding target and projectile nuclei. The selected energies are close to the Coulomb barrier, and the process mainly occurs in non-central collisions. Transfer reactions represent at these energies a large part of the total reaction cross section and are considered to produce efficiently new neutron-rich isotopes near the shell closures at $N = 126$ and in the superheavy region, areas that are difficult to reach with complete fusion with stable projectile and target nuclides.

5.2.1.3 Fission

With increasing proton number the heaviest nuclides undergo fission owing to the Coulomb repulsion of the protons. Heavy nuclei are characterized by a large neutron excess. This nuclear matter distribution

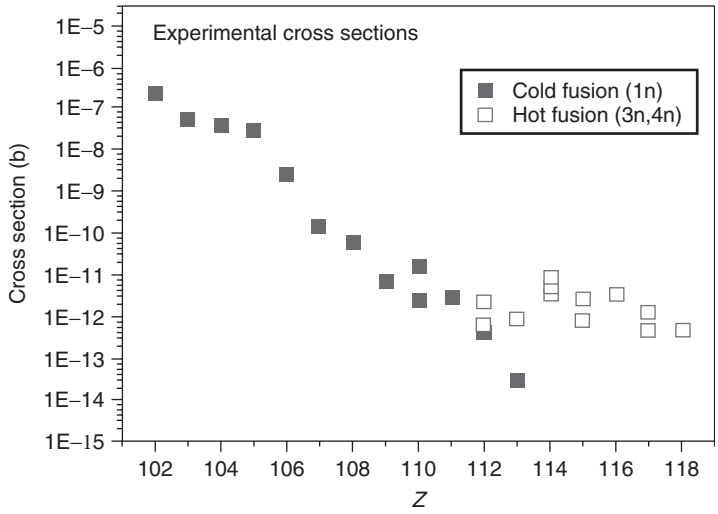


Figure 5.3 Measured cross sections to create SHEs via cold fusion reactions using ^{208}Pb and ^{209}Bi targets bombarded with projectiles from ^{48}Ca to ^{70}Zn. The cross sections are shown for the 1n channels (filled squares) and for the hot fusion reactions with ^{48}Ca projectiles and actinide targets (open squares). (Courtesy [13].)

makes fission reactions, in the element range of uranium and above, an important source for the production of very neutron-rich medium-mass nuclides. Fission can be initiated via many different processes. Thermal neutron- and photon-induced fissions are normally characterized by low excitation. Photon excitation can be practically achieved with Bremsstrahlung, created by stopping of an intense electron beam (real photons) or via electromagnetic excitation of fissile ions interacting with a high-Z target in peripheral collisions at relativistic velocities [3]. The latter excitation with virtual photons has been successfully applied to fission relativistic uranium projectiles via collisions with lead target nuclei. In this case, mainly the giant dipole resonance (GDR, 12 MeV) in uranium nuclei is excited before the fission process takes place. Higher excitation energies of the fissile nuclei can be achieved by reactions with neutrons or ions. This type of reaction can be realized either by bombarding or accelerating fissile materials. At relativistic energies, abrasion fission is a dominant reaction with ^{238}U projectiles for the production of neutron-rich new isotopes [14, 46, 88].

5.2.1.4 Fragmentation and Spallation

When the projectile energies exceed the intrinsic velocity of the bound nucleons (Fermi velocity $\approx$ 30 MeV u^{-1}), the probability for transfer reactions quickly decreases and one enters into the region where nucleon–nucleon collisions dominate the reaction mechanism. This is the domain of fragmentation that can also be described by a two-step process called *abrasion* and *ablation* [15]. In the first step, the overlapping zone from target and projectile is abraded. The non-interacting parts of the target–projectile system are called *spectators*, and the overlapping hot

zone of the collision represents the participants. In the second step, the moderately excited spectators, also called *pre-fragments*, reach their ground state by particle and γ-ray emission. These residual fragments are a rich and universal source of rare isotopes [16, 17], which can cover a large part of the chart of nuclides up to the mass of the heaviest collision partner (i.e., usually the heaviest element in the domain of fragmentation is uranium).

Depending on which spectator is investigated or applied, one uses the term *target* or *projectile fragment*. The physical models and descriptions are of course applicable to both, because the difference is only the kinematical reference system. Historically, one has called the case of target fragmentation induced by high-energy light ions, such as protons, as *spallation*.

In Figure 5.4, we exemplarily demonstrate for tin isotopes the different creation probabilities in terms of cross sections of fusion, projectile fragmentation, and fission. The most neutron-deficient tin isotopes, including the doubly magic ^{100}Sn, can be produced via fusion and projectile fragmentation. Fusion was successfully employed down to ^{101}Sn, but the doubly magic ^{100}Sn has been discovered via fragmentation. In a very recent ^{100}Sn experiment also, ^{99}Sn was unambiguously identified [18]. The most neutron-rich tin isotopes, interesting owing to their relevance for shell structure studies and the predictions of the astrophysical *r*-process path, can be very efficiently created via fission reactions.

5.2.2 Kinematics

The kinematic properties of the rare isotopes emerging from the production target can be crucial for the layout of in-flight separators and for their separation quality and efficiency [19, 20]. The energy-angular distributions and the atomic charge-state population are characteristic for the type of nuclear reactions used to produce rare isotopes. Owing to their kinematic properties, a number of production schemes, such

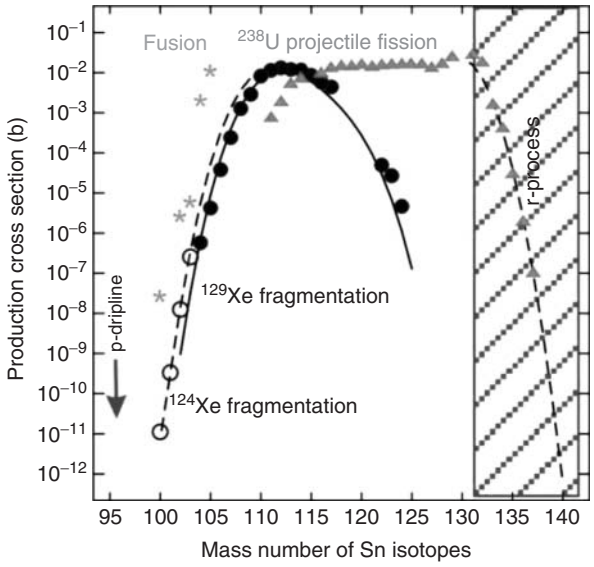


Figure 5.4 Calculated and measured cross sections for the production of tin isotopes. The fission cross sections are represented by triangles. The proton dripline and the onset of *r*-process nuclei are marked for orientation. The fusion cross sections (indicated by stars) have been calculated by extrapolation of experimentally known values in the neighborhood.

as high-energy-proton-induced fission and spallation, cannot be used with in-flight separators. For isotope separation on line (ISOL) systems [21, 22], the kinematics is not that crucial as the reaction products must be stopped. This can be done in the (thick) production target or in a catcher when thin targets are used.

Heavy reaction products emerging from relatively thin targets in forward direction with a high density in phase space are well suited for in-flight separation. This condition is well fulfilled for fusion products at Coulomb-barrier energies around $5\,\mathrm{MeV\,u^{-1}}$ and also for projectile and fission fragments at high energies (typically between [0.1–1.5] $\mathrm{GeV\,u^{-1}}$). The nuclear reaction kinematics is convoluted by the inevitable atomic interactions with the target atoms, such as angular scattering, energy-loss straggling, and electron capture and loss collisions. The latter interaction spreads out the rare isotopes over an extended phase-space volume particular at dispersive focal planes of electromagnetic devices. As a consequence, atomic collision processes limit the useful target thicknesses predominantly at low energies near the Coulomb-barrier.

5.3 Basic Separation Principles

The production is only the necessary first part of the experimental research with exotic nuclei. The next important task is to separate the exotic nuclides from the primary beam and the other abundant contaminants. Only after this phase, the exotic nuclides can be efficiently investigated with detector systems or used for reaction studies.

There exist two complementary ways to separate good quality exotic nuclear beams: the ISOL technique and the in-flight separation technique. Both methods transport the nuclei of interest away from their place of production, where the intense projectile beam and a large background from other nuclear reaction products are present, to a well-shielded experimental set-up, where the different nuclear and atomic properties can be explored. Apart from creating low-background conditions for the experiment, the transport serves at the same time to purify the beam and to prepare it in the necessary conditions with respect to energy, time structure, and ion optical properties for the experiments. Figure 5.5 shows a schematic comparison between the different techniques.

The ISOL technique [21, 22] uses either thick production targets of typically a few tens of grams per square centimeter where the nuclear-reaction products are stopped, or thin production targets of typically a few milligrams per square centimeter, where the reaction products recoil out and are collected in a solid or gas catcher [23]. The nuclei that are stopped in a heated target or a solid catcher are desorbed by diffusion and effusion processes and transported to an ion source for ionization before they are accelerated to a few tens of electronvolts. The target–catcher system can also be integrated into the ion source. The diffusion and effusion processes can, depending on the physical and chemical properties of the isotope of interest, be rather slow. It gives limitations to access very short-lived exotic nuclides of some elements, especially those with low volatility. When using a gaseous catcher (typically helium or argon), the nuclei are thermalized to a 1^+ state and then brought by the gas flow and sometimes supported by electric fields toward the acceleration parts. This method is much less selective but applicable for many elements. It is

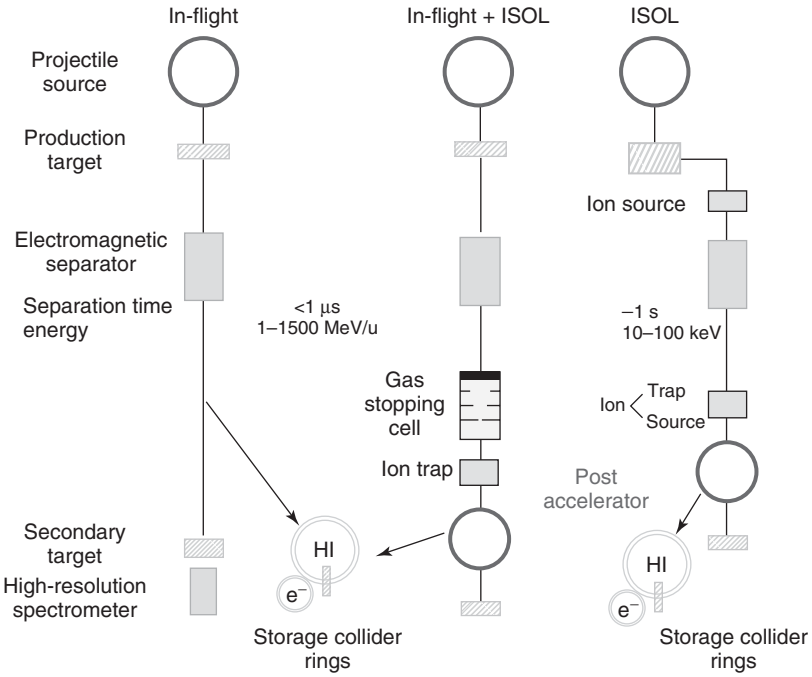


Figure 5.5 Principles of ISOL and In-Flight Facilities. Both are powerful and complementary methods for the production and separation of exotic nuclides. The combination of an in-flight facility with a gas-filled stopping cell provides advantages of both methods.

also possible to let the nuclei neutralize and then apply the ultra-selective resonant laser ionization technique. The gas-catcher technique is fast and is thus mainly used when the solid target–catcher systems are inefficient. It is now also used at the focal plane of in-flight separators (see relevant text in the following paragraphs 5.4.3).

Electromagnetic fields separate the nuclei according to their mass-over-charge ratio. The low-energy ion beam can be further accelerated in a tandem, linear, or cyclotron accelerator. Tandem accelerators need negative-ion injection, whereas linear accelerators and cyclotrons need highly charged, positive ion beams, as these make the acceleration process efficient and cost effective. To accomplish the latter, the 1^+ ion beams are injected in a charge breeder, or, one or more stripping stages are implemented in the accelerator sections. In ISOL systems, one obtains in general, for the well-suited nuclides, a much higher intensity compared to the in-flight method because of the performance of the high-intensity light ion driver accelerators and the much thicker primary production targets. Another advantage for some categories of experiments is that the phase-space volume is much smaller for a secondary ion beam emerging from the ISOL ion source compared to reaction products produced and separated in-flight.

The second method is in-flight separation [19, 20, 24]. Using this method, one

takes advantage of the specific reaction kinematics to separate selected nuclear reaction products from contaminants with a high suppression factor ($<10^{-12}$). In Figure 5.5, the principle of the in-flight method is compared with the ISOL technique. Because the target used in an in-flight separator is much thinner than that in an ISOL facility, the reaction products penetrate the ion-optical system with high velocities determined by the reaction kinematics. A major advantage of an in-flight separator is that the limits in half-lives for the separated exotic nuclides are determined only by the *time-of-flight* (TOF) through the ion-optical system. In-flight separators have been applied from low energies near the Coulomb barrier up to relativistic energies far above the Fermi energy [25, 26]. This means that already in the pioneering experiments the separated exotic nuclei have been used in studies with secondary reactions [27]. Novel studies are possible when in-flight separators are directly coupled with ion storage and cooler rings [28, 29] at high energies. In storage-cooler rings, the phase space of the exotic nuclear beam can be efficiently reduced by large factors of 10^5 and more. The decay of stored exotic nuclides can be studied with bare and few electron ions, a condition that prevails in hot stellar plasmas.

The combination of an in-flight separator with a gas-filled stopping cell, the so-called *gas catcher*, is a recent development that is based on the installation of the gas cell in the focal plane of the separator. Compared to the gas catcher approach described earlier, the primary beam does not enter the gas cell, while the exotic nuclei often already spatially separated and identified by A and Z do. Gas catchers are implemented at low- as well as at high-energy in-flight separators using heavy-ion fusion reactions or fragmentation reactions, respectively [30–33].

5.3.1 Intensities and Efficiencies

The intensity (I) of a beam of reaction product is determined by the cross section, the number of incident projectiles and target atoms, and the losses in the separation process:

$$I = \sigma N_{\mathrm{proj}} N_{\mathrm{tar}} \varepsilon$$

with σ being the reaction cross section (cm^2), N_{tar} the number of target atoms per surface area (cm^{-2}), and N_{proj} the primary beam intensity (s^{-1}). Since the cross section is energy dependent and the primary beam looses energy while passing through the target, the intensity has to be calculated by integrating over the target thickness, taking into account the energy loss of the beam. ε represents the efficiency of the whole separation process. This efficiency is defined as the ratio of the final secondary beam intensity that arrives at the experimental set-up versus the intensity of the created reaction products. ε is a product of a series of partial efficiencies depending on the technique applied for the production of the radioactive ion beam. These partial efficiencies are listed in Table 5.1 and Table 5.2.

5.3.2 Ion Transport in Static Electromagnetic Fields

In heavy-ion separators electromagnetic fields are employed to separate the desired nuclear species from the projectile beam or from other unwanted products and to transport them from the source to the focal planes for investigations with dedicated

Table 5.1 Partial efficiencies for ISOL.

$\varepsilon_{\text{delay}}$: probability of survival against radioactive decay during the time needed to extract the ion from the target-ion source or from the gas catcher system
ε_{ion}: ionization efficiency as defined for a stable isotope of a specific element
$\varepsilon_{\text{transport}}$: transport efficiency for the ion beam through the separator system
$\varepsilon_{\text{manipulation}}$: efficiency of ion beam manipulation in a trap
$\varepsilon_{\text{breeding}}$: charge-state breeding efficiency
ε_{TOF}: probability of survival against radioactive decay during the TOF through ion manipulation devices and charge breeders
$\varepsilon_{\text{acceleration}}$: efficiency of the post-accelerator
$\varepsilon_{\text{total}} = \varepsilon_{\text{delay}}\ \varepsilon_{\text{ion}}\ \varepsilon_{\text{transport}}\ \varepsilon_{\text{manipulation}}\ \varepsilon_{\text{breeding}}\ \varepsilon_{\text{TOF}}\ \varepsilon_{\text{acceleration}}$

Table 5.2 Partial efficiencies for in-flight.

$\varepsilon_{\text{transport}}$: acceptance of the separator times transport efficiency for the ion beam through the separator system
ε_{nuc}: losses due to secondary nuclear reactions in the matter (degrader, detectors)
$\varepsilon_{\text{injection}}$: efficiency of injection into another coupled ion-optical system (e.g., spectrometers, storage rings)
ε_{TOF}: probability of survival against radioactive decay during the TOF through the separator, ion manipulation devices, or storage rings
$\varepsilon_{\text{total}} = \varepsilon_{\text{transport}}\ \varepsilon_{\text{nuc}}\ \varepsilon_{\text{TOF}}\ \varepsilon_{\text{injection}}$

detector systems or, in case of energetic products, for further studies via secondary reactions or experiments in a storage ring.

The motion of the energetic ions in static electromagnetic fields is described by the Lorentz force:

$$\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \tag{5.2}$$

where q, $\boldsymbol{p}$, and $\boldsymbol{v}$ are the charge, the momentum, and the velocity of the analyzed ions, respectively. $\boldsymbol{E}$ and $\boldsymbol{B}$ are the static electric and magnetic field strengths governing the ion motion.

The ensemble of ions moving in electromagnetic fields can be described by Hamiltonian methods. It provides a framework for the ion motion within the allowed phase space using generalized coordinates and the Hamiltonian function including in this case the Lorentz force:

$$H = q\Phi + c\left\{m_0^2c^2 + \left(\boldsymbol{p} - q\boldsymbol{A}\right)^2\right\}^{1/2} \tag{5.3}$$

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}, \quad -\nabla\Phi = \boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t}$$

with $\boldsymbol{A}$, Φ, and m_0c^2 representing the vector and scalar potentials of the electromagnetic field and the rest energy of the ions, respectively. From the Hamiltonian function and the corresponding canonical equations, the motion of ions in the phase-space volume can be calculated. It follows directly that the particle density in the phase space under the action of the Lorentz force is invariant in time, that is, the time derivative of

the particle density is 0. A particle beam behaves as same as an incompressible liquid. This fact is called *Liouville theorem* and has many practical consequences for the description of ion motion in electromagnetic fields [34]. In a good approximation also sub-phase space volumes are preserved and the transfer function in matrix form has unit determinants (see relevant following text).

The situation is quite different when the ions interact with matter placed within ion-optical systems. The ions lose energy and change their original direction of motion via stochastic collisions with the atoms in matter. In this case, Hamilton's equations of motion have to be modified under the action of a non-conservative force, the stopping power. This results in an additive term in the generalized Hamilton equations that cause that now the phase-space density not being conserved. A non-conservative force is defined that it cannot be represented by a potential in the Hamilton equations. For such an interaction, the Liouville theorem is not valid and the phase-space density can either decrease or most often increase. The positive effect of a decrease in the phase-space volume can be achieved with different cooling processes, for example, electron cooling where the hot rare isotope beam is merged with a low-temperature electron beam in a storage ring [35] or when ions interact in a trap with the atoms of a buffer gas in a storage potential.

A deeper insight of the performance of a separator system can be obtained by elaborated ion-optical calculations [36, 37]. The goal in this calculation is to find the transfer function, which images the initial phase space to any desired position in the system. The transfer function is expanded in a Taylor series using the coefficients as elements in a matrix calculation. In these transformations, it is established to use a coordinate system moving with the reference charged particle on the central trajectory, the optical axis. The coordinates of a particle trajectory are given by (x, x',y, y', s, $\Delta p/p$), where x is the direction of the dispersion, and y is perpendicular to both x and the path length s along the central trajectory, which forms a curvilinear coordinate system. x' and y' are the derivatives of these coordinates with respect to s and represent the angles. $\Delta p/p$ represents the relative momentum deviation.

Using these definitions, the equation of motion in static magnetic fields is given by the linear second-order differential equations:

$$x'' + k_x(s)x = \frac{1}{\rho}\frac{\Delta p}{p}, \quad y'' + k_y(s)y = 0 \tag{5.4}$$

with the derivatives

$$x'' = \frac{\partial^2 x}{\partial s^2}, \quad y'' = \frac{\partial^2 y}{\partial s^2}$$

The solution of these trajectory equations can be evaluated by Taylor series where the coefficients represent the matrix elements, which represent, for example, the magnification and dispersion at a given focal plane. The imaging condition of such a system can be represented by a 6×6 matrix in first order, and a system consisting of several stages can be simply described by multiplication of the different matrices. The phase-space volume is formed for an ensemble of ions by the three position and momentum coordinates, which form in a good approximation an ellipsoid. For a Liouvillian system, the phase space is conserved throughout the ion-optical system and the matrices have unit determinants, even for the subsystems the corresponding

area is constant. For example, the transverse phase space in x-direction, the emittance, is given by

$$\varepsilon_x = \pi \sigma_x \sigma_{x'} \tag{5.5}$$

The emittance is conserved in static electromagnetic fields. In the case of de- or acceleration with electromagnetic fields (conservative forces), the emittance multiplied by the momentum of the ions is also constant. Non-Liouvillian properties of the ion beam are inevitably introduced with atomic and nuclear collision processes.

5.3.3 Separation in Electric and Magnetic Sector Fields

The spatial separation of the selected rare-isotope beam in homogeneous sector fields is governed by their magnetic rigidity

$$B\rho = \frac{p}{q} = \frac{mv}{q} \tag{5.6}$$

and their electric rigidity:

$$E\rho = \frac{pv}{q} \tag{5.7}$$

It follows directly from the Eqs. (5.6) and (5.7) that a separator based on pure electric or pure magnetic fields cannot focus ions of different velocities. However, new possibilities are opened for a combination of magnetic and electric sector fields. In case that the magnetic field is counterbalanced by an electric field leads to a pure velocity separation given by

$$\frac{\Delta v}{v} = \frac{\Delta(E\rho)}{E\rho} - \frac{\Delta(B\rho)}{B\rho} \tag{5.8}$$

Such separators are called *velocity filters* or *Wien filters*; they transmit for perpendicular field arrangements only one velocity, which is given by the ratio of the electric and magnetic field strength $v = E/B$. Owing to the stochastic nature of the nuclear reactions creating exotic nuclei, it is desirable to eliminate the influence of the inevitable velocity spread from the possible mass-over-charge resolution. This goal can be achieved with double focusing spectrometers, where a twofold $B\rho$ resolution is counterbalanced by $E\rho$ resolution

$$\frac{\Delta(m/q)}{m/q} = \frac{\Delta(E\rho)}{E\rho} - 2\frac{\Delta(B\rho)}{B\rho} \tag{5.9}$$

Separators employing electrical fields are operated at relative low velocities and are mostly used in the near-Coulomb-barrier region owing to technical limitations of the electric field strength.

A basic building block of a separator is illustrated with a simple one-stage dispersive system consisting of static magnetic fields (Figure 5.6). The system consists of a target area at the entrance of the ion-optical system (object plane) and the incident phase space is transported with quadrupole lenses in front and behind a dipole magnet to the focal plane at the exit of this unit. The envelopes are depicted in x-direction for an emittance of 20 π mm mrad. Such a one-stage dispersive separator will spatially deflect the reaction products according to the magnetic rigidity; that is, the system requires that the $B\rho$ values of the selected exotic nuclides are different from the one of the primary beam or of abundant contaminants. This means that this dispersive system must have a sufficient large $B\rho$-resolving power. The ion-optical resolving power (R_{io}) is determined by the normalized magnetic flux, $\int \vec{B}\,\mathrm{d}\vec{f}$, enclosed between the central and outermost trajectories and the

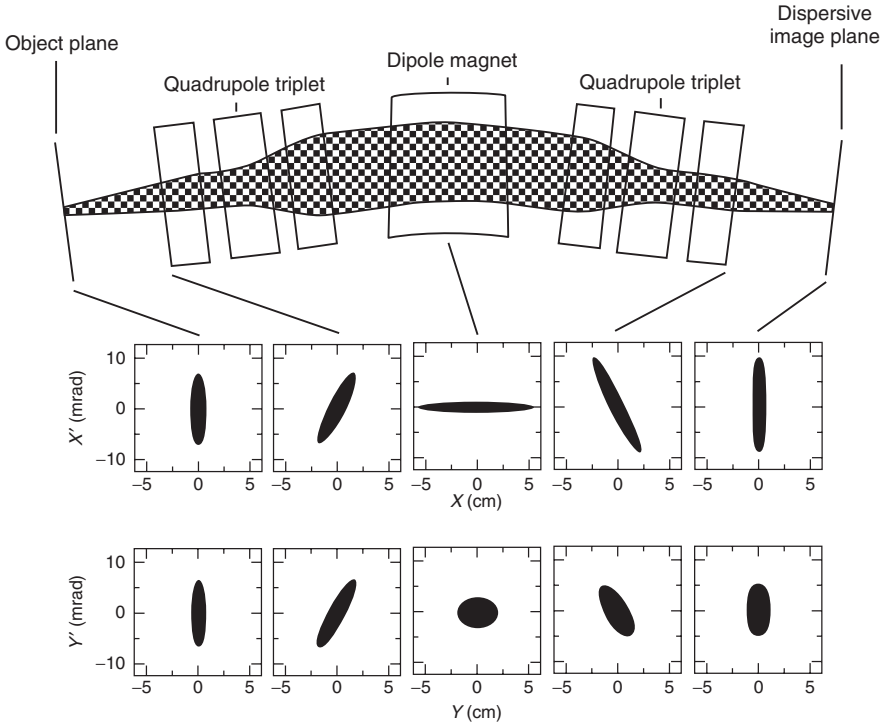


Figure 5.6 A one-stage dispersive magnet system, with ion-optical elements and calculated envelopes in *x*-direction. The system consists of a lens system (quadrupole triplets) in front and behind of a dipole magnet. The position and angular sub-phase spaces in *x*- and *y*-direction are given for $\delta p = 0$ at five positions.

emittance in the *x*-plane:

$$R_{\mathrm{io}} = \frac{1}{2x_0\,2x'_0}\;\frac{1}{B\rho}\int B\mathrm{d}f \qquad (5.10)$$

The corresponding resolving power of a separator with electrical fields is quite similar, that is, the magnetic parameters are replaced by the electric field strength and the electrostatic rigidity.

The entrance quadrupole triplet illuminates the dipole magnet (see marked area in Figure 5.6) to provide the magnetic flux which determines the possible resolution. This dispersive stage can in principle represent the first stage of an in-flight separator or can be used for analysis behind the ion source of an ISOL system. The ion transport in the presented magnetic system is governed by conservative forces which means that the phase-space volume formed by all particle trajectories is conserved according to the theorem of Liouville, that is, the transfer matrix has a unit determinant. The phase-space volumes are even constant in the $(x{-}x', y{-}y')$ subspaces, for $\delta p = 0$, because the coordinates in *x* and *y* are not coupled in the first-order ion-optical approximation. The conservation of the sub-phase space is illustrated by the plotted emittance ellipses in *x*- and *y*-direction at different positions

along the ion-optical system. A practical application of this phase-space conservation used in separator systems are strong focusing lenses in front and behind the target acting as condenser structures. They rotate twice the enlarged emittance (upright in position and angle) to preserve the resolution and to simultaneously achieve a larger phase-space acceptance.

The condition for the required difference in $B\rho$ of the selected fragment, the primary beam and the contaminants is not a severe restriction for a dispersive entrance stage particularly if the kinetic energy of the projectile beam is high enough to emerge from the target as fully ionized ions. The situation at low energies is clearly more difficult, for example, for a 6 MeV u^{-1} ^{238}U primary beam populating about six different charge states. In this case, a finger system must be employed at the dispersive focal plane to stop the primary beam, and the reaction products are transported further only through the remaining areas. In this context, one has also to consider that not only the primary beam but also the formed exotic nuclei travel in many different charge states through the separator system, which makes the separation even more complex.

Behind this first dispersive magnetic stage, there are many other electromagnetic field variants possible to form a complete separator system. For example, an electrostatic stage can be added, and then the overall system will separate according to Eq. (5.9). Because such a combination usually focuses independently of the incident velocity and angular spread, the performance is called *double focusing*. As mentioned earlier, the practical realization of electric fields is limited to low and intermediate energies. At high energies the required length and field strength exceed the present technical possibilities. However, in the domain of projectile fragmentation (E $\gg$ 30 MeV u^{-1}) another opportunity is created and successfully used.

The principle works as an overall achromatic ion-optical system, which can be realized by mirror imaging of the one-stage dispersive system, see Figure 5.7. Ion-optically, it requires that waist conditions in x- and y-directions and a parallel dispersion line are realized at the dispersive intermediate focal plane. In the calculations provided earlier (Figure 5.6), these conditions are obtained as illustrated with the upright phase-space ellipsoids and the dispersion function. The achromatic system has the advantage that the spatial resolution at the final focus is independent of the velocity – and angular spread of the reaction products created at the entrance. The latter property is illustrated with three different velocities ($\delta p = 0$, $\pm 1\%$) and an angular range of 7 mrad for a selected fragment, they are well separated at the intermediate focus and form a single beam spot, determined by the chosen optical magnification, at the final focal plane. The principle of the described achromatic system is successfully applied in many in-flight facilities worldwide.

The described separators focus in angle and are lateral dispersive. They use the high kinetic energies of the reaction products emerging from relatively thin targets characterized by thicknesses much smaller than the atomic range of the reaction products. In this way, they yield access to all exotic nuclides with lifetimes down to the flight time through the optical system. Depending on the energy domain and the nuclear reaction type, the transit time can be well below 1 μs.

A special extension of an achromatic separator system is the application of a so-called energy-loss spectrometer. In this

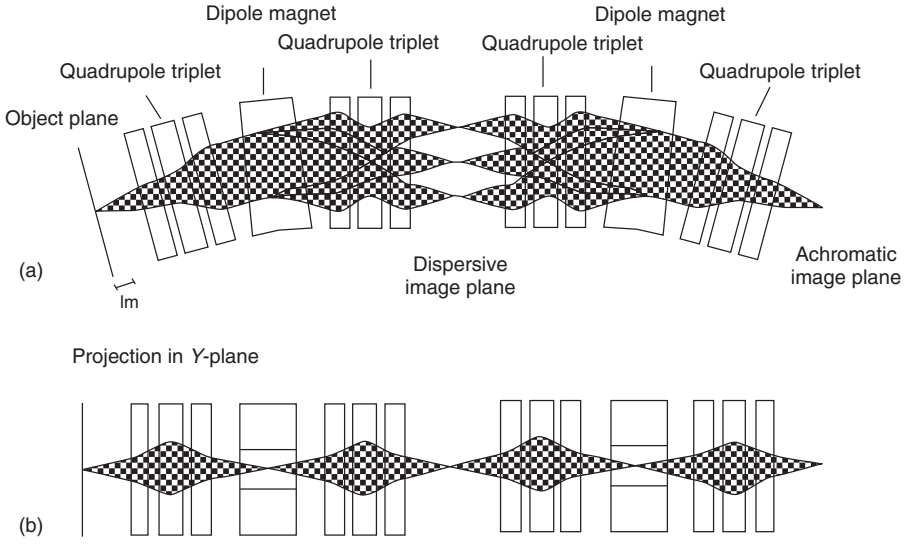


Figure 5.7 (a) Calculated achromatic system. The ion transport is shown in the *x*-plane for three different $B\rho$-values deviating by $\pm 1\%$ from the reference particle at the optical axis. (b) The optical elements and the envelopes in the *y*-direction.

case, the nuclear reaction takes place at the dispersive intermediate focal plane of the two-stage system consisting of two parts with the same resolution. For such a system the influence of the incident phase space is canceled in the final analysis of the reaction products. This means that one can obtain a higher resolution for the analyzed exotic nuclear beam than the momentum spread and emittance of the incident primary beam would allow.

5.3.4
Separation via Time-of-Flight

Separation of energetic ions created in nuclear reactions can also be performed on the basis of high-resolution TOF and frequency measurements in isochronous spectrometers [38] or in storage devices (storage rings, traps, and cyclotrons) [35, 39, 40]. A lower resolution separation can be obtained with RF deflectors (*radio-frequency*), not discussed here. The highest background suppression can be obtained with accelerator mass spectrometry (AMS) which is also not discussed here.

The isochronous conditions for an electro-magnetic separator are such that the transit or revolution time for an ion with a given mass-to-charge ratio is independent of the velocity. This can be achieved if ions with a higher velocity travel on longer trajectories so that the velocity difference is completely canceled in the flight time or revolution frequency. In some device, this separation in time can be converted into spatial separation, for example, with a Bradbury–Nielsen–Gate [40].

On the basis of these ion-optical methods, the resolving power of such a system can be in the order of 10^6; thus even isobars and ground and isomeric states can be resolved. An obvious consequence

is that such high resolution isochronous storage devices are ideally suited to perform direct mass measurements of exotic nuclides via recording the revolution time. The described properties and resolution can be illustrated by simple first-order formulas for a storage ring. The revolution frequency f is given by the ratio of the ion velocity v and the circumference of the closed ion orbit C: $f = v/C$. Furthermore, the length of the closed orbit depends only on the magnetic rigidity of the circulating particle:

$$\frac{\Delta C}{C} = \alpha_p \frac{\Delta(p/q)}{(p/q)} = \alpha_p \frac{\Delta(B\rho)}{(B\rho)} \tag{5.11}$$

where p and q are the momentum and the electric charge of the stored ion, respecitvely. The momentum compaction factor $\alpha_p = 1/{\gamma_t}^2$ is determined by the ion optical setting of the storage ring. From Eq. (5.11), one can deduce the differential dependence of the revolution frequency on the velocity and the mass-to-charge ratio of the stored ions:

$$\frac{\Delta f}{f} = -\alpha_p \frac{\Delta(m/q)}{(m/q)} + \left(\frac{1}{\gamma^2} - \frac{1}{{\gamma_t}^2}\right)^2 \frac{\Delta v}{v} \tag{5.12}$$

where γ is the Lorentz factor $\{\gamma = [1 - (v/c)^2]^{-1/2}\}$. The highest resolution for the revolution frequency measurement of circulating hot exotic nuclides can be obtained when the second term of Eq. (5.12) becomes negligible, that is, when the isochronous condition is fulfilled ($\gamma = \gamma_t$), that is, at the so-called transition energy. Longitudinal dispersive devices achieve the required resolving power only if the measured time for an ensemble of ions, at a selected focal plane, is not dependent on the incident velocity (momentum) spread or on the positions or angles. These ion-optical conditions and requirements include that the system is, after one turn, also achromatic.

Time-of-flight mass spectrometers (TOF-MS are well suited for broadband measurements with characteristic times in the millisecond range or less. Relativistic ions in a storage ring are even faster. The revolution time they have is typically 500 ns. For many decades, the resolving power of the isochronous devices was only modest, but recently, this situation has been changed by enlarging drastically the path length via multi-turn modes. In a storage ring this means to aim at many revolutions for the circulating ions and for the low-energy TOF system to use electrostatic mirrors to achieve many reflections without losing the time-focusing properties.

For both systems, a mass resolving power of more than 100 000 has been reached. Both devices are schematically shown in Figure 5.8. The geometrical dimensions differ by a factor of more than 100. But in the same direction is the operating energy, which is for ion storage rings, in the megaelectronvolt per nucleon range and for multiple TOF-MSs in the sub-kiloelectronvolt range. Both devices are especially fast because the beam preparation via cooling is generally not needed. However, if the lifetime of the investigated exotic nuclei allows, different cooling methods can be applied, which increase the phase-space density and thus the resolution without removing ions from the stored ensemble. Cooling is discussed in Section 5.3.8.

5.3.5
Separation Employing Atomic Interaction

As demonstrated earlier, pure electromagnetic fields lead to spatial separation of energetic ions from the projectile beam and background from other reaction products

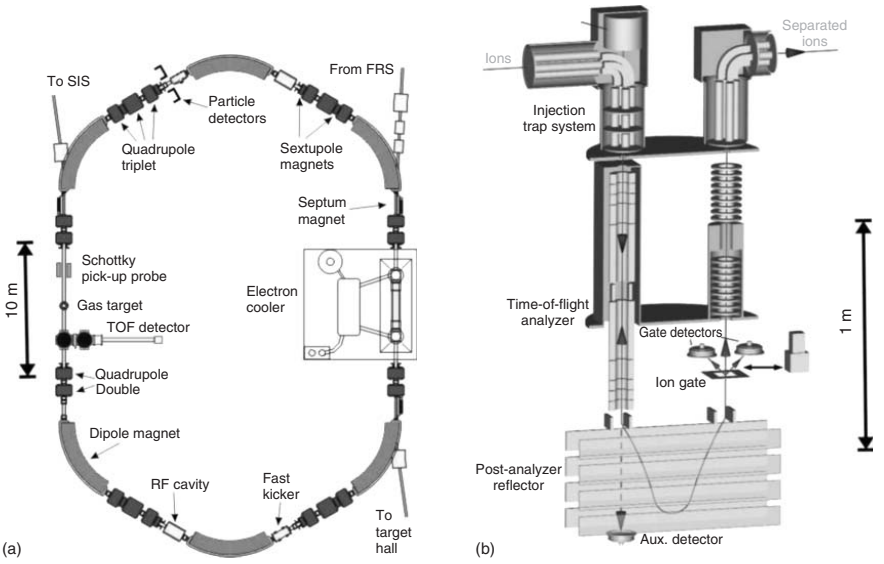


Figure 5.8 High-resolution isochronous devices. (a) Scheme of the layout of an ion storage-cooler ring; here, the ESR at GSI [41] is presented. The ring can be operated in the isochronous mode or for long-lived fragments ($T_{1/2} > 1$ s) the stochastic and electron cooling can be employed. In the isochronous mode, no cooling is needed. Therefore, only the flight time is a principal lower limitation. The velocity spread can be reduced to 10^{-7} with cooling. The revolution time can be measured by Schottky pick-up probes or by TOF detectors. The circumference of the ring is 108 m, and for the isochronous revolution time measurement typically several hundred turns are used. (b) Isochronous multiple reflection time of flight spectrometer [40]. The RFQ mass filter at the entrance is used to remove contaminants and to apply buffer-gas cooling. The high resolution is achieved by enlarging the flight path via multiple reflections in the TOF analyzer. After separation and analysis in the TOF analyzer, the ions can either be detected in a multi channel plate (MCP) detector, or be spatially separated with an ion gate for further investigations.

according to their mass-over-charge ratio, which is often not sufficient. Introducing non-Liouvillian elements, (e.g., the atomic interaction with matter), for instance, in a gas atmosphere, or a piece of solid material, placed within the dispersive optical elements, can be an excellent solution.

At relatively low velocities where the exotic nuclei carry still many bound electrons during their flight through the electromagnetic fields, atomic charge-exchange processes are applied in gas-filled separators [42] to increase the transmission and to provide an additional separation criterion. The atomic interactions with the gas atoms in a dipole section lead to charge-exchange collisions, which yield, after an equilibrium thickness, a mean charge state that depends on the atomic number Z_1 of the ions and is proportional to their velocity v:

$$\bar{q} = Z_1^{1/3}\frac{\boldsymbol{v}}{\boldsymbol{v}_0}, \quad B\rho = 0.0227\, mZ_1^{-1/3}Tm \tag{5.13}$$

where $v_0 = 2.18 \times 10^6\,\mathrm{m\,s^{-1}}$ is the Bohr velocity. Inserting this relation in Eq. (5.6) shows that for ions with a selected mass, the magnetic deflection is dependent on the atomic number but no longer on the velocity. The formula for the mean charge is a result of the Bohr stripping criterion and the Thomas–Fermi model. The gas layer must be thick enough to create charge-state equilibrium and optimized such that the advantage of frequent charge-changing collisions is not canceled out by the influence of the inevitable angular and energy loss straggling.

Two main features are advantageous at high and relativistic velocities for the application of atomic interactions of heavy ions with matter placed within ion-optical systems: first, the ions emerge from the target completely ionized, $q = Z$, and secondly, the energy-loss and angular straggling result only in small phase-space enlargements. These basic properties were deduced from accurate energy-loss measurement with the spectrometer FRS [43] in an energy range of 80–1000 MeV u^{-1} [44]. These accurate experiments contribute to a substantial improvement of the theoretical description beyond the first-order Born approximation for the atomic collisions.

The separation principle is that a specially shaped piece of matter (energy degrader) is placed in a dispersive intermediate focal plane of an achromatic magnetic separator system. A twofold $B\rho$ analysis is combined with energy loss ΔE in between which is called *$B\rho$–ΔE–$B\rho$ method* [43]. The shape of the degrader is chosen according to the goals of the separation, for example, the shape can be adjusted to preserve the achromatic condition of the ion optics (achromatic degrader) or it can be optimized to bunch the energy distribution (monoenergetic degrader). The optimum thickness of such an energy degrader reaches the order of half of the atomic range of the selected penetrating exotic nuclear beam. The separation scheme with an achromatic degrader is illustrated by ion-optical calculations in Figure 5.9a. The first stage of the spectrometer separates the projectile fragments according to A/Z because the velocity of the primary beam is approximately conserved in these peripheral nuclear collisions. The energy required to yield 90% fully ionized fragments is the lower limit for an effective operating domain. In the second magnetic dipole stage, the energy loss ($\Delta E \sim Z_1$) dissipated in the thick degrader is measured with subsequent separator stages thus providing a Z separation.

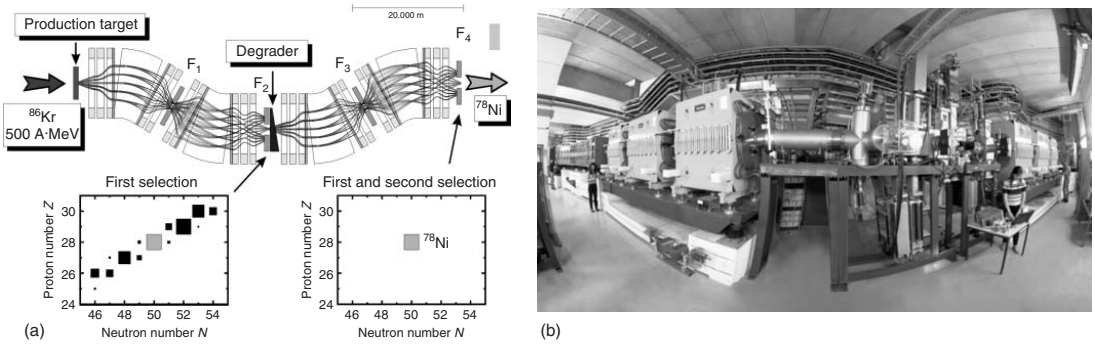


Figure 5.9 (a, b) The fragment separator FRS as an example of an achromatic separator system, consisting of two dispersive dipole magnet stages with quadrupole and hexapole focusing elements in front and behind the degrader stage. The ion-optical elements and calculated envelopes in *x*-direction are shown. (a) An achromatic degrader is positioned at the central dispersive focal plan. In this simulation, ^{78}Ni ions are produced via projectile fragmentation of ^{86}Kr projectiles and ^{78}Ni ions are isotopically separated at the final focal plane. (b) A Photograph of the central part of the FRS, in which the 30° dipole magnets and focusing lenses are seen surrounding the central focal plane with the degrader station. Please note that the photograph has been taken with a large-angle "fish eye" lens. (Photo A. Zschau.)

5.3.6 Delay Times and TOF

An important advantage of an in-flight separator is that the losses due to the radioactive decay of the exotic nuclei during the whole separation process are only determined by the TOF through the ion-optical system. The TOF is in the order of microseconds in devices using reactions near the Coulomb barrier and decreases to a few hundred nanoseconds with separators for relativistic nuclei. Therefore ε_{TOF} is close to 100% for isotopes and isomers having half-lives slightly above these values. Only when further ion manipulation like, for example, accumulation of ions in a storage ring, is performed, ε_{TOF} might be reduced or be used to let contaminants with shorter half-lives decay before the actual measurement starts.

The situation is different for ISOL systems [22, 47] as the time scale the atoms reside in the target-ion source system is in most cases much longer compared to the TOF. ε_{delay} expresses the relative amount of atoms that survive, in spite of their radioactive decay, during the time elapsed from the moment of their production to the moment of their extraction from the ion source. In case of solid and liquid catcher systems this so-called delay time is determined by the diffusion from the target/catcher material, the desorption from the material surface and the effusion to the ion source exit hole. For gas catchers delay times are determined by the gas flow and, in case electrical fields are used to obtain a faster evacuation of the gas cell, by the mobility of the ions. ε_{delay} depends on the half-life of the exotic nucleus of interest and on its physical and chemical properties. The differences in physico-chemical properties can be used in a number of cases to suppress contaminating isobars. To quantify the loss factor, a delay-time distribution ($P(t)$) is defined, whereby $P(t)\mathrm{d}t$ is the probability for an atom of a given element created at $t = 0$ to be released from the ion source between time t and $t + \mathrm{d}t$. It is important to note that this function does not depend on the half-life of the isotope. One can calculate ε_{delay} by folding $P(t)\mathrm{d}t$ with $\exp(-\lambda t)$:

$$\varepsilon_{delay}(T_{1/2}) = \int_0^\infty P(t)\mathrm{e}^{(-\lambda t)}\mathrm{d}t \quad (5.14)$$

with $\lambda = 1/\tau = \ln(2)/T_{1/2}$. Note that $P(t)$ is normalized thus that $\varepsilon_{delay}(T_{1/2} = \infty) = 1$, which means that all produced atoms will in the end have escaped from the target-ion source system. One obtains the average delay time as

$$\tau_{delay} = \int_0^\infty tP(t)\mathrm{d}t \quad (5.15)$$

An illustrative release process is described by a delay time distribution equal to a properly normalized single exponential function:

$$P(t) = \lambda_{delay}\mathrm{e}^{(-\lambda_{delay}t)} \quad (5.16)$$

with a release rate constant equal to $\lambda_{delay} = 1/\tau_{delay} = \ln(2)/T_{delay}$. It is easy to show that:

$$\varepsilon_{delay}(T_{1/2}) = \frac{T_{1/2}}{T_{1/2} + T_{delay}} \quad (5.17)$$

In this case half of the radioactive atoms will escape from the ion source if the delay time equals the mean life time but still a fraction of 10% escapes even if the delay time is nine times longer than the half-life. The latter is due to the exponential nature of the release process in this case. The effect of the delay times for a series of

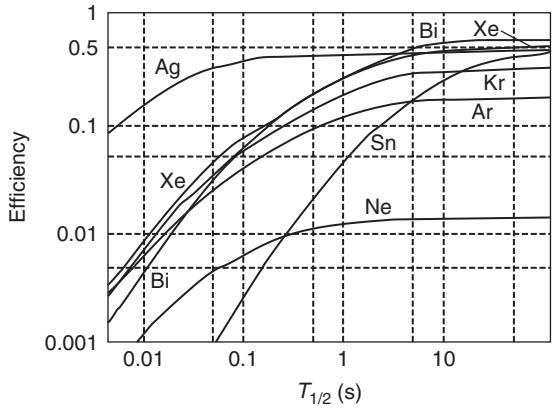


Figure 5.10 The products of ε_{delay}, ε_{ion}, $\varepsilon_{transport}$, and ε_{TOF} are shown as a function of half-life for an ISOL system. For silver, a graphite catcher was used, while for the other cases niobium foils. For long half-lives, ε_{delay} evolves to unity and the total efficiency depends on the product of the three remaining efficiencies. Very fast diffusion/effusion is observed for the silver isotopes, ε_{delay} is still close to 100% for isotopes with a half-life down to 100 ms. Isotopes of tin, however, encounter a strong loss owing to the increased delay times, ε_{delay} is close to 0.5% for isotopes with a half-life of 100 ms. (Adapted from [48].)

elements and catcher materials is shown in Figure 5.10.

Moreover, the times involved in cooling or charge-state breeding will induce losses for short-lived radioactive isotopes and can be described in a similar manner.

Although in most cases delay losses should be avoided, sometimes the difference in delay times between different elements is used as a separation mechanism. A recent example studied at ISOLDE was the production of a ^{68}Ni ($T_{1/2} = 29$ s) beam while purifying it from the overwhelming ^{68}Ga ($T_{1/2} = 67.6$ m) contamination. As the gallium atoms diffuse much faster from the target ion source system compared to nickel, a dedicated time gate after the impact of the primary proton beam was applied to increase the ^{68}Ni to total beam intensity by several orders of magnitude without a substantial loss in the ^{68}Ni beam intensity. The same procedure was used to suppress short-lived francium isotopes to obtain isobaric pure beams of bismuth.

5.3.7 Ionization

Different ionization mechanisms, which can be applied as an effective separation tool, are implemented in the ion sources of ISOL systems. Their use depends essentially on the ionization potential (W_i) of the element of interest, the required charge state, and the needed selectivity. Three main ionization mechanisms are implemented to create the ions of interest. In gas catchers, ion survival is of importance.

5.3.7.1 Electron Impact Ionization

For isotopes of elements with $W_i > 7$ eV and for the creation of multiply charged ions, electron impact ionization is mostly used. The atoms or ions are bombarded by energetic electrons, thereby losing one

or more of their outer electrons. The bombardment of atoms by mono-energetic electrons and the related evolution of the atom and ion density as a function of time can be described by the following equation:

$$\frac{dn_i}{dt} = (n_{i-1}\sigma_{i-1\to i} - n_i\sigma_{i\to i+1})j_e \qquad (5.18)$$

with n_i being the ion density with charge state i, j_e the electron current density, and $\sigma_{x\to y}$ the cross section for impact ionization from charge state x to y. This equation is valid for i from 1 to $i_{max} - 1$. For the neutral atom density n_0, the first term of the right-hand side vanishes, whereas for the fully stripped ion n_{imax}, the last term vanishes. Note that this expression does not take into account charge exchange, multiple charge ionization, or recombination effects. Electron impact ionization is used in high temperature gaseous discharge ion sources for 1^+ charge-state ion beams (hot plasma sources), and in electron cyclotron resonance ion sources (ECR) and electron beam ion sources (EBISs) for multiple charged ion beams. In the gaseous discharge ion sources and the EBIS, the electrons are accelerated over a DC potential whereas in ECR ion sources the electrons are heated by injecting RF power. Because of the unselective nature of this ionization process, ion sources based on this principle do not offer much chemical selectivity, in general. However, in some cases, high-temperature gaseous discharge ion sources are ideal to obtain through chemical techniques a very high degree of purity. By adding minute amounts of chemical compounds to a hot plasma ion source, the isotope of interest can form molecules. Once these molecules get ionized, molecular ion beams can be extracted and mass separated. As the molecular formation probability depends strongly on the chemical properties of the elements of interest, often only one element forms molecules, while all the isotopes from the other elements remain in their atomic form. As a consequence, the molecular ion beams is situated in another mass region that is free from isobaric contamination. For example, SnS molecules are easily formed by adding sulfur to the ion source while all neighboring elements such as for example, indium, antimony, and cesium do not form molecules. The formation of SeF is another example to form pure beams of neutron-deficient selenium isotopes. These molecules can easily be dissociated in an ECR or EBIS charge breeder prior to injection in the post-accelerator, and in this way, pure energetic ionic beams are obtained. Another example is cooling the transfer line between the target and the ion source. This prevents the condensable elements to reach the ion source. This technique was already applied in the very first ISOL facility to produce pure beams of krypton isotopes [49] and is now readily used to produce pure beams of gaseous elements (e.g., the noble gaseous) or molecules (e.g., nitrogen or oxygen molecules).

5.3.7.2 Surface Ionization

Surface ionization is a frequently used ionization mechanism. When an atom interacts with a heated surface, it can lose or gain an electron before leaving the surface as a positive or negative singly charged ion. This technique can be used efficiently for elements with an ionization potential $(W_i) < 7\,\text{eV}$ for the creation of positive ions (positive surface ionization) or with electron affinity $E_A > 1.5\,\text{eV}$ for the creation of negative ions (negative surface ionization). The ratio between the ion density (n_i) and the neutral density (n_0) of a certain element, with W_i or E_A, at a

heated surface, with temperature (T) and work function (ϕ), is given by the Langmuir equation:

$$\frac{n_i^+}{n_0} = \left(\frac{g_i^+}{g_0}\right) e^{[(\phi - W_i)/kT]}$$
$$\frac{n_i^-}{n_0} = \left(\frac{g_i^-}{g_0}\right) e^{[(E_A - \phi)/kT]} \qquad (5.19)$$

with g_i and g_0 being the statistical weight of the ionic and atomic ground state. The positive and negative sign indicate positive and negative ions, respectively. The efficiency is then described as

$$\varepsilon_{ion} = \frac{n_i}{(n_i + n_0)} \qquad (5.20)$$

depending strongly on the temperature and the difference between the work function of the surface material and the ionization potential or the electron affinity of the element of interest. Materials with high work functions, such as tantalum ($\phi = 4.19$ eV), tungsten (4.53 eV), and rhenium (5.1 eV), which can be heated to high temperatures, are used to construct positive surface ionization sources. Compounds with low work function such as LaB_6 ($\phi = 2.6$ eV) or BaO (1.7 eV) are used for negative surface ionization. Using this ionization technique, extreme selectivity can be obtained if the elements of which isotopes are produced in the same nuclear reaction have very different ionization potentials. For example, the neighboring elements krypton ($W_i = 14.0$ eV), rubidium (4.18 eV), and strontium (5.70 eV) are often produced in the same nuclear reaction, but their ε_{ion} differs by several orders of magnitude owing to the exponential dependence of Eq. (5.19). In general, the alkali elements (group IA of Mendeleev's table: lithium, sodium, potassium, etc.) are efficiently ionized using positive surface ionization, whereas the halogen elements (group VIIA of Mendeleev's table: fluorine, chlorine, bromium, etc.) are good candidates for negative ionization. When the surface ionization mechanism is applied in a hot cavity, a quasi-neutral plasma in thermal equilibrium can be created. This so-called thermo-ionization mechanism is applied in hot-cavity ion sources and results in higher ε_{ion} for elements with slightly higher ionization potentials without reducing substantially the selectivity.

5.3.7.3 Laser Resonant Ionization

Resonant photo-ionization using lasers has recently been implemented in ISOL systems. It is based on a stepwise excitation of the outer atomic electron by laser photons, leading finally to the continuum, to auto-ionizing states or to highly excited states close to the continuum. In the latter case, the ionization step is achieved through infra-red irradiation, electrical fields, or atomic collisions. The ionization process consists of typically two or three steps and, because of the resonant nature of most of these steps, resonant laser ionization is very efficient and element selective, resulting in isobaric and, if the spectral resolution is high enough, even isomeric pure beams. Figure 5.11 shows an example of a three-step ionization scheme used to produce isomeric beams of lead isotopes.

As the cross section for the first step excitation is large, the atomic ground and excited state are in equilibrium during laser irradiation. Under this condition the evolution of the number of neutral atoms (N(t)) can be obtained from

$$-\frac{dN(t)}{dt} = (\sigma_I F + \beta)N(t) \qquad (5.21)$$

With σ_I (cm^2) being the cross section for ionization from the excited state to

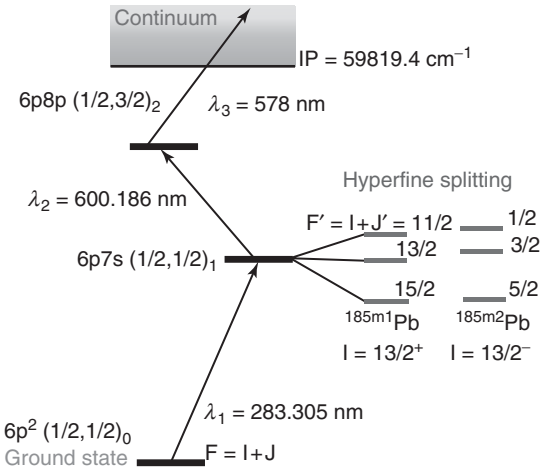


Figure 5.11 A three-step laser ionization scheme applicable for ^{185}Pb atoms. The nuclear-spin $I = 13/2+$ or $3/2-$. The atomic spin of the first excited state ($J' = 1$) leads three possible values for the total spin (F') of the energy levels whose degeneracy is lifted. The hyperfine splitting, and the isotopic and isomeric shift of the atomic levels are used to measure charge radii; magnetic and quadrupole moments can be used to ionize the isomer or the ground state separately, thereby producing isomeric nuclear beams.

the continuum, F (cm^{-2} s^{-1}) the flux of photons, and β the depopulation rate of the excited state due to natural decay or other loss mechanisms except ionization. Moreover, it is assumed that the laser pulse fires at $t = 0$. Conditions on the flux (the number of photons per second and unit area) and on the fluence (the number of photons during the laser pulse per unit area) of the laser photons are obtained. The flux condition requires that the depopulation rate of the excited state by ionization into the continuum is much larger than that of the natural decay toward, for example, a metastable state. Thus

$$\sigma_I F > \beta \tag{5.22}$$

The fluence condition requires that in total (i.e., in one laser pulse) enough photons irradiate the atom to induce the ionization step:

$$\sigma_I \phi \frac{g_2}{g_1 + g_2} > 1 \tag{5.23}$$

with ϕ (cm^{-2}) being the fluence and g_i the statistical weight factor of the ground state ($i = 1$) and excited state ($i = 2$), respectively. When both conditions are fulfilled, complete saturation is reached; that is, essentially all the atoms irradiated by the laser light are ionized. These conditions can nowadays be reached by pulsed lasers, whose photon wavelength must be tunable in order to obtain high efficiency, and element and/or isomer selectivity. As radioactive nuclei are produced in most cases in a quasi continuous mode, the low-duty cycle of the pulsed lasers limits ε_{ion} to about 10^{-4}. Therefore, apart from the laser development as such, much effort has been devoted to construct an effective storage device for radioactive atoms in between laser pulses. Laser ionization has been applied in hot cavities [50] and gas-cell systems [51]. Typical dwell times in hot cavities are of the order of 0.1 ms; in gas cells, they are above tens of milliseconds. Therefore, for laser with a pulse repetition rate of 10 kHz, 100 Hz are used. Note that, due to thermo-ionization in a hot cavity (see relevant earlier-given text) or ion survival in a gas cell (see relevant further text), the selectivity is reduced. Resonant laser ionization can result in a selectivity of 10^4 per resonant step, whereas the selectivity varies between 10 and 1000 when applied in the ion source systems. In order

to improve the selectivity and the spectral resolution of the laser ion source, ionization outside the atom storage device (hot cavity or gas cell) has recently been performed [52, 53]. In this so-called "laser Ion source trap (LIST)" approach, when the atoms escape, for example, from the gas cell through the exit hole, a cooled and homogeneous gas jet is formed. Contaminating ions leaving the gas cell can be stopped by a repelling voltage. By counter propagating the laser light with this gas jet and capturing the photo-ions in an RF ion trap, a high level of selectivity and spectral line width is achievable. This allows us to produce pure isomeric beams of medium heavy and heavy elements as well as detailed laser spectroscopy measurements.

5.3.7.4 Ion Survival in a Noble Gas Environment

The creation of radioactive ion beams can also be accomplished using gas catchers. The reaction residues are thermalized in a gas cell filled with a noble gas (helium or argon). Together with the gas, the atoms and ions are evacuated through the exit hole of the gas cell and injected into the first acceleration stage of an isotope separator or into a RF ion guide. Ionization is not created by external means (electron bombardment, surface ionization, or laser ionization) as in the earlier-mentioned mechanisms; ions after thermalization, the noble gas environment prevents part of the to be neutralized, as charge exchange with the noble-gas atoms is hampered. The way the 1^+ ions are formed and survive the transport through the gas cell depends strongly on the experimental conditions, namely evacuation time, purity of the buffer gas, electron and ion density created in the cell, and, in some cases, the chemical properties of the isotopes. In the so-called ion guide isotope separator on line (IGISOL) systems, fast evacuation is the key issue, as it prevents the ions from recombining and neutralizing [54]. In gas catcher systems, after in-flight separations, electrical RF and DC fields are used to collect electrons and to increase the transport time of the ions through the gas cell [30–32], thereby reducing neutralization and diffusion losses.

5.3.8 Ion Beam Cooling, Accumulation, and Bunching

The ion-optical resolving power of a separation device depends on the longitudinal and transverse emittance of the ion beam. Maximizing the phase-space density will also increase $\varepsilon_{\text{transport}}$ of any separator system and the injection efficiency into, for example, multi-reflection TOF spectrometers or ion traps for mass measurements, or charge-state breeders for further post acceleration. The phase space of the secondary ion beam from in-flight separation mainly determined by the reaction kinematics (see Chapter 5.2.2.) is inferior to the one from ISOL systems, which depend on the ion-source properties. The phase space of in-flight beams can be improved by cooling in a storage ring (see relevant further text), or the energy spread can be reduced by the so-called range bunching, as explained in Chapter 5.3.5 and Ref. [34]

To maximize the phase-space density for low-energy ISOL beams, buffer-gas cooling is applied as implemented in Paul and Penning traps [55]. Schematic drawings of both ion trap principles are shown in Figure 5.12.

The ion confinement capability of an RF trap is based on the creation of an inhomogeneous RF field that results in a so-called pseudopotential. For example, for the linear Paul trap with four rods as electrodes

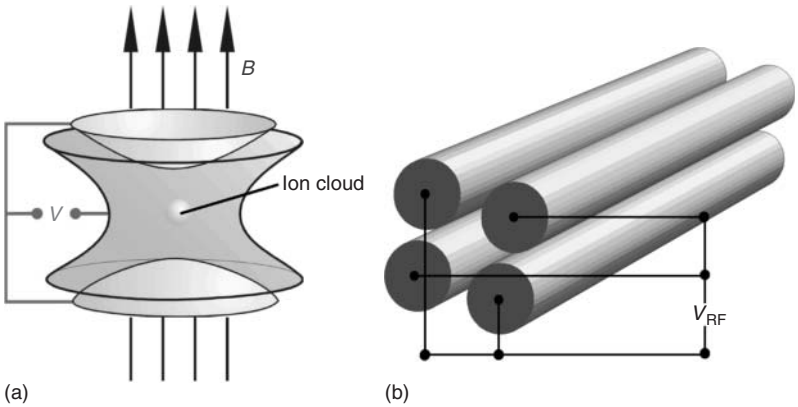


Figure 5.12 Principles of a Penning (a) and a linear Paul trap (b). Ion confinement in a Penning trap is achieved by a combined static magnetic dipole and electric quadrupole field configuration. In the Paul trap a confinement potential is created by the RF electric quadrupole field. (Adapted from [55].)

(Figure 5.12) the pseudopotential (V_{pseudo}) that confines the ions in the radial direction can be expressed as

$$V_{\text{pseudo}}(r) = \frac{e}{m}\left(\frac{U_{\text{rf}}}{r_0^2\omega_{\text{rf}}}\right)^2 r^2 \qquad (5.24)$$

where e and m are the ion electrical charge and mass respectively, U_{rf} and ω_{rf} are the RF amplitude and frequency, $2r_0$ the separation between the surface of the opposite electrodes, and r the distance from the central axis.

A Penning trap combines an electrical quadrupole DC field with a homogeneous magnetic field in which the ions are confined in radial direction by the magnetic field and axially by the electrical field. Ion movement in a Penning trap is characterized by three oscillating frequencies ω_+, ω_-, and ω_z. The former two depend on the well-known cyclotron frequency $\omega_c = (e/m)B$ of particle with electrical charge (e) and mass (m) in a magnetic field (B):

$$\omega_\pm = \frac{\omega_c}{2} \pm \sqrt{\frac{\omega_c^2}{4} - \frac{\omega_z^2}{2}} \qquad (5.25)$$

whereby ω_z is determined by the electrical field potential (U) and characteristic trap dimension (d)

$$d = \sqrt{\frac{\rho_0^2}{4} + \frac{z_0^2}{2}}$$
$$\omega_z = \sqrt{\frac{eU}{md^2}} \qquad (5.26)$$

Although the ions are confined in these traps, they undergo numerous collisions with the room temperature (or even cooled) buffer gas atoms, leading to a process called *buffer-gas cooling*. In most cases, helium is used as a buffer gas. The effect of the buffer gas on the ion motion can be parameterized as a viscous drag force that depends on the gas pressure, ion mobility, and temperature. The net result is that the ions lose their energy, resulting in a substantial increase in the phase-space density when they are ejected from

the trap. The efficiency for ion trapping and cooling ($\varepsilon_{\text{manipulation}}$, see Section 5.3.1) reaches above 50%, and the techniques are readily used to accumulate ions and to release them in a short bunch, to reduce the energy spread of the ion beam, or to store the radioactive ions in a smaller volume, which makes measurements less sensitive to trap imperfections. An increased phase-space density, especially a reduced energy spread, increases the mass resolution of a dipole magnet-based mass separator (see above) [42, 47]. Bunching of the ions and reducing the energy spread is also of interest for collinear laser spectroscopy measurements to increase the peak-to-background and the Doppler broadening, for mass measurements and further beam purification in Penning traps and multi-reflection TOF devices. EBIS charge breeders also profit from the possibility of pulsed injection, as the pulse structure can be matched to the duty cycle of the EBIS and thus for subsequent injection in storage rings after post acceleration (see relevant further text). As mentioned earlier, in the LIST mode RF, structures are integrated in target-ion source systems based on laser ionization to capture the photo-ionized ions from atoms leaving a hot cavity or a gas cell.

Other means for ion or atom cooling that are currently being implemented for radioactive ions is laser cooling in a magneto optical trap. Temperatures below millikelvin can readily be achieved however at an efficiency cost: $\varepsilon_{\text{manipulation}}$ is currently below to the per mille level. Further optimization might make these devices applicable for RIB manipulation.

Cooling of a circulating ion beam in a storage ring is, besides the discussed isochronous operating mode, the second experimental method to achieve a high resolving power via revolution-time (frequency) measurements. Second, cooling is a prerequisite in collider experiments to efficiently perform phase-space compression, or it is used to simply prevent beam losses caused by intra-beam scattering or scattering at an internal target. Cooling is a method to overcome the limits given by the Liouville's theorem discussed earlier.

There are different ways of ion-beam cooling in a storage ring, which can be subsequently applied depending on the experimental requirements. Electron cooling was invented [56] to prepare antiproton ring experiments. Electron cooling is realized by merging the stored hot ion beam with a cold electron beam in a straight section of a ring. It reduces, via atomic collisions, the longitudinal and transverse phase space; that is, the velocity spread and the divergence of the ion beam are by several orders of magnitude reduced. While the same stored and cooled ion beam interacts with the cold electrons at each revolution, the electron beam is constantly being renewed. Unfortunately, it takes a certain time until temperature equilibrium is reached, usually of the order of a few seconds. The cooling time t_{cool} [57] strongly depends on the incident velocity spread (velocity deviation β) and the intensity of the merged electrons (n_{e^-}), the ion mass (m_{ion}), and charge (q):

$$t_{\text{cool}} \propto \frac{m_{\text{ion}}}{q^2 n_{e^-}} \beta^3 \gamma^5 \theta^3 \tag{5.27}$$

The mass and charge dependence is in favor of heavy bare ions. In addition, the cooling time and final equilibrium depends on the intensity of the stored ions [58]. At low intensities of several hundred stored atomic nuclei, a phase transition was observed, which led to a velocity spread of about 10^{-7}, which is the basis for accurate

mass measurements with resolving powers of about 10^6. The long cooling time in the beginning of the process when the exotic ions have the largest longitudinal and transverse spread can be overcome with fast pre-cooling by applying stochastic cooling. Stochastic cooling is well suited to compress the phase-space volume of the coasting ion beam. It uses the fact that the phase space volume is not occupied by a continuous medium of invariant density but by individual ions with a finite volume, which are points with an empty space in between. Stochastic cooling has the task to reduce this empty phase volume and thus to achieve a phase-space compression of the ions. This is performed via recording of the particle coordinates with sensitive pick-up probes and powerful amplifiers with a large bandwidth (W). This signal is then sent to special kicker devices to correct for the measured displacement. Analysis and correction are positioned at appropriate ion-optical positions in the ring with respect to the betatron oscillation. It is obvious that the particle TOF from the place of analysis to the location of the kicker has to be chosen such that it exactly matches the time for signal propagation. The time for stochastic cooling depends on the number of stored ions (N_{ion}) and the electronic bandwidth W:

$$t_{\mathrm{cool}} \propto \frac{N_{\mathrm{ion}}}{W} \tag{5.28}$$

The advantage is that the cooling time does not depend, in first approximation, on the large emittance of the incident injected exotic nuclei. Therefore, stochastic cooling can be advantageously applied as a fast pre-cooling stage before the electron cooling performs the final work. An example of stochastic pre-cooling followed by electron cooling is shown in Figure 5.13 for a hot uranium beam at 400 MeV u^{-1}. The total cooling time for this example is about

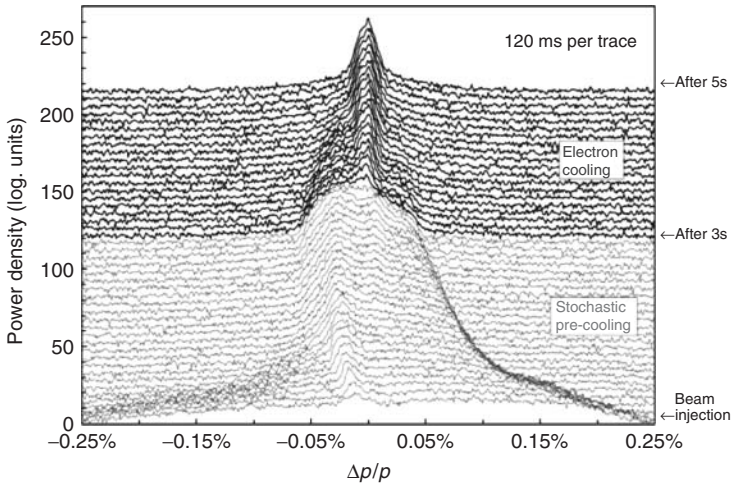


Figure 5.13 Measured cooling time and performance for a hot uranium beam circulating in the ESR. The stochastic pre-cooling reaches an equilibrium after 2–3 s; subsequent electron cooling reduces the velocity spread from 10^{-3} to better than 10^{-6} after another 2 s [35].

4–5 s. Laser cooling is not frequently used in experiments with exotic nuclei.

5.4 Facilities

5.4.1 In-Flight Facilities

In-flight separators are best suited for the investigation of short-lived nuclides at the limits of stability. They are fast because the separation time is only the flight time through the separator, which is in the submicrosecond range. In-flight separation is not sensitive to the chemical properties of the separated species. Modern in-flight separators with their detection system are sensitive down to single atoms and have provided new physics phenomena with individual atomic nuclei.

The large primary beam intensities, up to $10^{14}\,s^{-1}$, planned for future heavy-ion accelerators represent a great challenge for the next-generation separators. A method chosen to cope with this is to have separators consisting of several stages (e.g., pre- and main-separators) [59], with different separation criteria to reach the necessary background suppression at the final focal plane and to combine the separator with a high-resolution system such as a magnetic spectrometer, a storage ring, or an ion-trap system. In general, low-energy systems use the decay characteristics for identification or high-resolution recoil mass separators for isotopic separation. High-energy systems combine the in-flight separators with detector systems for direct Z and A identification.

5.4.1.1 Separators for Low-Energy Exotic Nuclei

In-flight separators for low-energy nuclear reaction products, near the Coulomb barrier ($E < 6\,MeV\,u^{-1}$), are often called *"recoil separators,"* where the term *"recoil"* is related to the reaction kinematics; that is, the produced fragments of interest recoil from the collision process and emerge from the thin target toward the entrance of the device. In this sense, a velocity filter or a double-focusing mass spectrometer, as described in Chapter 3, belong to the category of recoil separators. Recoil separators consist of electric and magnetic dipoles, focusing lenses, and multipole fields for the correction of image aberrations. Such recoil separators can also include a complete velocity filter as a modular block. High vacuum conditions ($\approx 10^{-7}$ mbar) are necessary in recoil separators in order to avoid losses of the reaction products due to atomic interactions, such as slowing down, angular scattering, and charge-changing collisions. Furthermore, this vacuum condition is required to operate the high electric field strengths.

As described in Chapter 2, a velocity filter is well suited to separate in-flight the selected exotic nuclei, formed via fusion, from other reaction products and the scattered primary beam. For example, the velocity filter SHIP (separator for heavy ion reaction products) is successfully applied at GSI to separate the heaviest fusion evaporation products and has enabled the discovery of the elements $107 \leq Z \leq 112$ [60–62]. It is a two-stage velocity filter with spatially separated electric and magnetic fields. SHIP separates fusion products emerging from a thin target, with high efficiency and low background. Reaction products are accepted in an angular range of 3° in horizontal and vertical directions and a velocity range of 10%. They are implanted into a silicon detector with their full energy, where their decay is observed *in situ*. The unambiguous

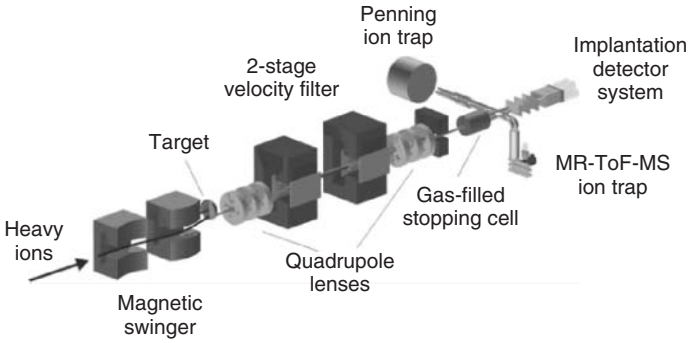


Figure 5.14 A new generation of velocity filter for the investigation and discovery of exotic nuclei, especially for neutron-rich isotopes of the heaviest elements. The separator elements are: a magnetic swinger in front of the target, a large acceptance two-stage velocity filter with large aperture superconducting quadrupole magnets, a low-resolution $B\rho$ deflection, atomic range separation in solids, before the beam is stopped in the gas cell. The fast extracted beam is then transported to a high-resolution trap system. Depending on the half-life range, a Penning or a multiple reflection TOF spectrometer is applied. A new feature will also be the recycling of the non-reacted projectiles. The components are only schematically shown and not drawn to scale.

isotope identification is performed via α spectroscopy using the position correlation between implanted evaporation residue and the daughter generations until the detected α chains reach characteristic α lines of well-known isotopes. This method has worked extremely well for very neutron-deficient isotopes till Cn ($Z = 112$) and, as noted in RIKEN Japan, also for the element 113 [63]. In hot fusion reactions, more neutron-rich elements have been produced up to $Z = 118$ [64]. However, their decay chains end in fission before the area of known α-emitters has been reached. This means a direct identification of the mass and proton numbers is not possible by the method of parent–daughter correlation. An experimental solution can be a fast combination of stopping in a gas cell and an ion trap with mass resolving powers of 10^5 and better. In this way, the isotope can be unambiguously identified. Depending on the lifetime of the produced SHE, a conventional Penning trap or a fast multiple reflection time-of-flight mass spectrometer (MR-TOF-MS) can be applied, as indicated in Figure 5.14. The described scenario is a next-generation solution when the primary beam intensity and thus the background have been increased by a factor of about 100. In this case, a magnetic swinger system in front of the target, the combination of a high-acceptance velocity filter, with the atomic range selection in front of the gas cell, and an effective high-resolution trap system represent a powerful future facility. The magnetic swinger allows the acceptance of reaction products that have the highest production cross sections different from the 0° direction with respect to the incident primary projectile beam, for example, from transfer reactions. Targets mounted on a fast rotating target wheel must withstand the high-intensity primary beam, which is certainly a challenge and requires cooling and a continuous constant intensity distribution in time. The two-stage velocity filter removes most of the scattered primary beam before a separation according to atomic ranges in solids and the stopping in the gas cell follows. The final purification

will be performed after extraction from the gas cell with radio frequency quadrupole (RFQ) systems and in the last step with the ion traps. The RFQs at the exit of the gas cell can be easily removed and inserted in the beam axis by means of remote-controlled vacuum feedthroughs; thus, a conventional semiconductor implantation detector system can also be alternatively employed. For this operation mode, the gas cell will be evacuated.

With gas-filled separators, an experimental development has started more than five decades ago with investigations of fission products [65]. These separators (now used in heavy-element research) are characterized by large transmission of the fusion products, while the primary beam and the contaminants have been successfully suppressed to enable discoveries of new isotopes down to femtobarn production cross sections [63]. The separation principle is explained in Chapter 3. Ion-optically, the device is very simple and compact; it consists of a gas-filled (He, H_2) dipole magnet followed by a quadrupole magnet system, which focuses the selected reaction products on the decay detectors [42, 66]. A possible next-generation extension of such a system would consist of a magnetic swinger system in front of the target to be able to measure also different from 0°, which is advantageous if transfer reactions are used to create neutron-rich heavy elements. Furthermore, a large superconducting quadrupole system in front of the dipole magnet would determine with the corresponding field flux illumination the proper resolution needed for the specific separation task. Moreover, the focusing system behind the gas-filled dipole could be more effective if large apertures and superconductivity are employed. The advantage of a gas-filled system is the effective collection of different charge states due to atomic collisions, and an effective separation according to the mean charge (see also Chapter 4.3 (Figure 5.15).

An interesting combination of different low-energy separation criteria could be a velocity filter system followed by a gas-filled separator. Such a combination has, besides the velocity separation, a charge-state focusing, depending on the nuclear charge (see Eq. (5.13)). A novel application is also to apply atomic energy-loss (ΔE) separation in solid degraders also at low energies.

5.4.1.2 Separators for High-Energy Exotic Nuclei

Projectile fragmentation and projectile fission are powerful sources for the production of exotic nuclei at energies exceeding the Fermi velocity up to the relativistic domain (1–2 GeV u^{-1}). At these velocities, the application of electric fields is not practical any longer; thus, the separators are based on magnetic elements only. The in-flight separators at high energies are usually called *"fragment separators"* because of the dominant use of projectile fragmentation and fission reaction for the production of exotic nuclei. All existing and planned projectile fragment separators are based on the described $B\rho$–ΔE–$B\rho$ separation method. The phase-space enlargement due atomic interactions of relativistic heavy ions in matter is very favorable and most of the ions are fully ionized. Therefore, the heaviest projectile fragments were discovered at 1000 MeV u^{-1} at GSI Ref. [46] (see Chapter 5.5.1). The production of exotic nuclei with projectile fragmentation was pioneered at relativistic energies in Berkeley (USA) [16] and has been applied afterwards at intermediate energies at GANIL (France) and RIKEN (Japan) Ref. [88]. The $B\rho$–ΔE–$B\rho$ separation method at relativistic energies has been pioneered

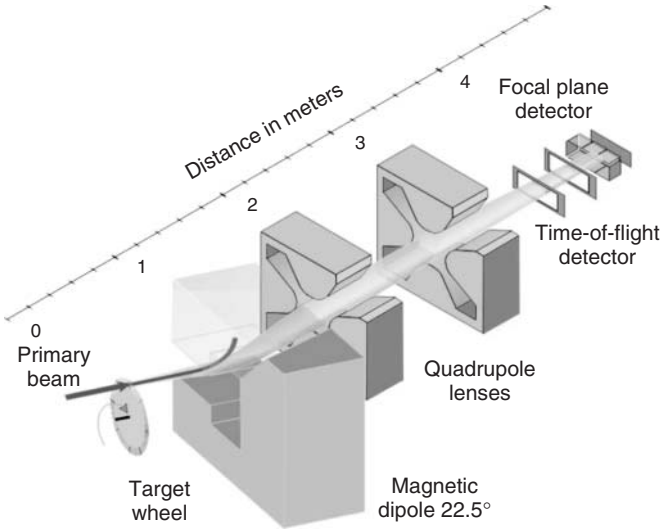


Figure 5.15 Gas-filled separator: a typical gas-filled separator is the Dubna Gas-Filled Recoil Separator (DGFRS) consisting of a rotating target and an entrance window in front of the gas-filled dipole magnet. A quadrupole doublet focuses reaction products on the detector setup. (Courtesy A. Popekov.)

with the FRS [43] at GSI (Germany). After the successful operation of the ion-optical layout of the four magnetic dipole stages, including focusing quadrupole and hexapole magnets, of the FRS, this scheme has been introduced to other facilities worldwide (e.g., the A1900 MSU (USA) [67], the Big-RIPS (RIKEN) [68], and the planned Super-FRS at GSI [59]). A fascinating aspect of fast projectile fragments is that they have high kinetic energies to be investigated in various secondary nuclear reactions.

As an example of a typical next-generation facility, the Super-FRS, presently under construction, is schematically shown in Figure 5.16. It includes the main characteristic features deduced from the long experience of the successfully running devices. The incident high intensity primary beam is guided on the production target by a special versatile focusing system that can vary the shape and area of the beam spot according to the requirements from the optical resolution, the transmission, and thermal endurance of the target material. It consists of a Pre-Separator and a Main-Separator each equipped with at least one degrader system. In the first dipole stage, the non-reacted primary beam will be separated from the selected fragments and implanted in a movable beam catcher system. In the Pre-Separator, a coarse separation will be performed, including the removal of contaminants caused by not fully ionized primary beam ions. In general, the rate of ions penetrating the first degrader might be fairly high; therefore, contaminants created in this thick degrader must be removed in the next stage as well. In the Main-Separator, the resolution for spatial separation is roughly twice higher compared to the Pre-Separator, and thus

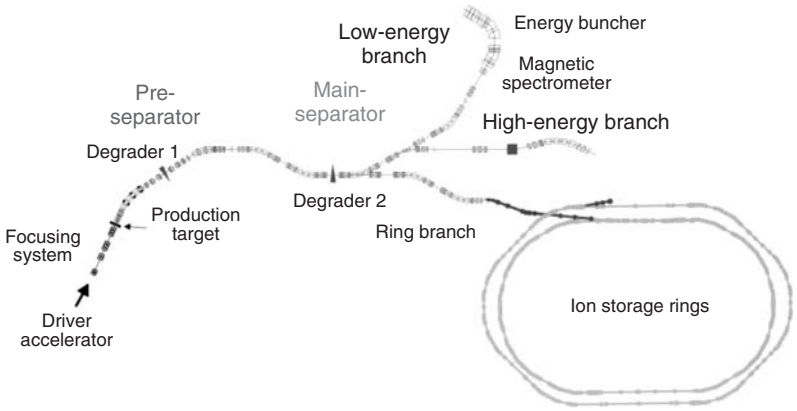


Figure 5.16 The Super-FRS [59], presently under construction, is a large-acceptance in-flight separator for experiments with exotic nuclei produced at relativistic energies up to 1500 MeV u^{-1}. The maximum rigidity is 20 Tm. The Energy Buncher represents a high-resolution versatile spectrometer, which is suited to compress the energy distribution; thus the slowed down exotic nuclei can be stopped in a gas cell of about 1 m length. Separated fragments can be stored and cooled in storage rings for unique mass and decay measurements, or studied in scattering experiments with the internal target.

complete isotopic separation can be obtained. A versatile fragment separator can usually transport the separated exotic nuclei to different ion-optical branches. As illustrated in Figure 5.16, the planed Super-FRS has three branches, similarly to the present FRS.

The Ring-Branch is connected to storage rings, where the stored and cooled exotic nuclei are investigated in precision experiments, including direct reactions at the internal target. The High-Energy Branch provides the separated fragment beam up to the highest energies, corresponding to a maximum magnetic rigidity of 20 Tm, for experiments to study exotic nuclei under the so-called complete kinematic conditions. Since the reaction and particle decay in flight are reconstructed by the method of invariant mass, particle unstable nuclei and resonances can also be investigated in this way. At the Low-Energy Branch, the separated fragment beam can be efficiently slowed down in a gas cell to be fast extracted for precision experiments in traps.

Quite similar to the Super-FRS layout, other in-flight facilities such as the Big-Rips in RIKEN (Japan), the separator A1900 at NSCL (USA), and the RIBLL at IMP in China and in Korea are built or planned.

5.4.1.3 Combination with Spectrometers, Storage-Rings, and Traps

Many new properties of exotic nuclei were discovered in secondary reactions or with high-resolution measurements after the stage of isotope separation. The pioneering experiments with relativistic fragments were performed by Tanihata et al. in 1985 [27] using the BEVALAC accelerators at LBL Berkeley and a large dipole magnet, and the HISS spectrometer. This was the birth of the high-energy rare-isotope facilities now existing at JINR Dubna Russia; INFN Legnaro, Italy; GANIL, France; RARF, Riken, Japan; GSI, Germany; and NSCL at MSU, United States and IMP Lanzhou, China. The relatively small emittance and narrow momentum distribution

allow an efficient transportation of separated fragment beams to experimental setups such as large spectrometers for reaction studies and storage rings. Recently, the developments of efficient and fast gas-filled stopping cells in combination with MR-TOF spectrometers and ion traps open up a new field for experimental studies of beams produced at Coulomb barrier energies as well as for the relativistic regime.

Reaction Setups with a Large Dipole Magnet A typical setup for reaction studies with unstable nuclei separated in flight is schematically shown in Figure 5.17. The experimental scenario with rare isotope reactions is to measure in reversed kinematics with the goal to record all secondary reaction products. This gives access to short-lived species far-off stability separated in-flight. Important for these reaction studies is the possibility to use cocktail beams, which consists of different projectile fragments characterized by the same magnetic rigidity. The advantage of the reversed kinematics technique is that all reaction participants have high energies and are kinematically focused in forward direction. This allows experiments in complete kinematics, that is, the detection of all particles participating in the reaction in the entrance- and exit channels, including γ-radiation. Such an experimental setup is applied at several exotic nuclear beam facilities, such as ALADIN-LAND at GSI, MONA at NSCL, and SAMURAI at RIKEN.

The secondary beam separated in the in-flight separator is directed onto the secondary target with energies typically in a range of several hundred megaelectronvolts per nucleon. The separated fragments are measured event-by-event in A and Z. Position, angle, and momentum are measured with tracking detectors. In the case of exotic nuclei at the limits of stability produced with very low counting rates, measurements with cocktail beams are advantageous. Typically light target materials (H_2, CH_4, and C) are used for scattering experiments; heavy targets (Pb) are advantageous for Coulomb excitation including the GDR. The target is surrounded by a crystal ball for γ-detection. The large-aperture dipole magnet deflects

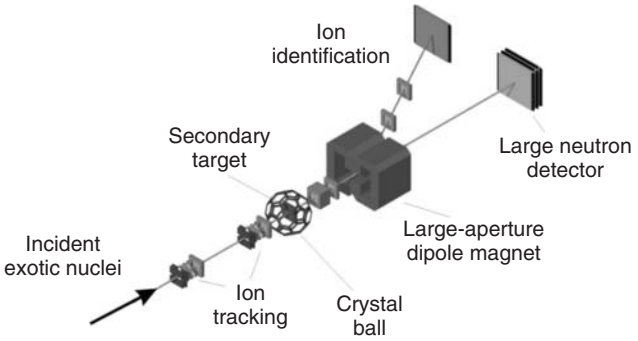


Figure 5.17 Exotic nuclei are produced, separated in-flight, and transported to the reaction setup. A cocktail beam of rare isotopes is focused on the secondary target viewed by a γ-array (crystal ball). Neutrons and charged particles are measured behind the secondary target by employing a large-aperture dipole magnet and a large area neutron detector. The kinematically complete measurements yield spectroscopic results also for unbound nuclei via the reconstruction of the invariant mass. In this way, the unbound systems such as ^{10,12}Li and ^{7}H have been investigated [69].

the charged particles, including protons, α-particles, and heavy fragments. Tracking detectors measure the kinematic properties, a plastic wall is used for TOF information and the energy deposition for *Z*-identification of light fragments, and the heavier ions are identified by multi sampling ionization chambers, which can resolve all elements up to uranium. Neutrons are measured with a large neutron detector, and a combination of scintillators and iron converters (neutron wall). An important feature of such a detection device is that many experiments can be done simultaneously with a cocktail beam of exotic nuclei identified event-by-event in front of the secondary target and behind.

High Resolution Magnetic Spectrometers Spectrometer setups are used in combination with fragment separators for reaction studies with high resolution. An example of such a combination is the A1900 fragment separator with the S800 [70] spectrograph at MSU, USA. It combines high resolution and high acceptance in a single device, and is specially designed for reaction studies with secondary beams. The large angular (20 msr) and momentum (5%) acceptances are well adapted to the emittances of radioactive beams produced by projectile fragmentation. The high resolution is achieved for secondary reaction products, such as core nuclei after knockout reactions, in a *dispersion-matched* operating mode, whereby the image of the secondary target represents a dispersive focal plane, and the final focal plane is achromatic with respect to the incident primary fragment beam. Ion-optically the high resolution of the S800 is achieved via an analytical reconstruction method in which aberrations are calculated a priori from the magnetic field maps and used directly to correct the raw data. It is installed vertically on a carriage that can be rotated from 0° to 60°. Its maximum rigidity is limited to 4 Tm, corresponding to kinetic energies 176 MeV u^{-1} for ions with a mass-over-charge ratio of 2. As in other reaction setups, the secondary target is surrounded by an efficient γ-array.

In-flight separators in combination with high-resolution spectrometers in other laboratories are the α + SPEG at GANIL/SPIRAL (France), BIG-RIPS + SHARAQ at RIKEN, and Grand-RAIDEN at the RCNP Osaka (Japan). A combination of a large acceptance dipole magnet combined with a high-resolution spectrometer is planned for NUSTAR experiments at FAIR-GSI. Such combinations are ideally suited for studies of rare species at the outskirts of the nuclear landscape.

Storage Rings The combination of an in-flight separator with a cooler-storage ring allows precision experiments with exotic nuclear beams in a large energy range, because in general a storage ring has also synchrotron capabilities and thus can vary the energy of the circulating beam with RF cavities. Presently, there are worldwide two facilities based on this combination: GSI (Germany) and IMP Lanzhou (China). An ion storage ring is operated under ultra-high-vacuum conditions (10^{-11} mbar) to prevent interactions with the atoms of the rest gas. The ions circulate in the rings with about 2 MHz at an injection energy of about 400 MeV u^{-1}. The pioneering experiments with relativistic short-lived nuclei have been performed with the FRS and ESR at GSI in the early 1990s Ref. [28].

Depending on the goal of the experiment, pure isotopic or cocktail beams can be injected and stored in ion rings. They have a large acceptance for circulating beams that even changes in the ionic charge states keep the ions within the ring apertures. This large acceptance is also

important for decay measurements with stored ion beams. As the longitudinal emittance (Et) of the fragment beams is small, the bunch structure from the synchrotron is almost preserved, a necessary condition for the kicker system to deflect the fragments into stable storage orbits. With the RF force, the stored ions can be transferred from the injection orbit to a storage orbit where the ions merge the electron beam from the cooler. In this way, multiturn injection can be used to increase the intensity of the stored beam. The stored beam can be pre-cooled with stochastic cooling, and finally cooled with merged electrons at low temperature. Permanent cooling compensates for intra-beam scattering and allows long storage times of up to more than 10 h for light ions at about 400 MeV u^{-1}.

The ESR [41] (Figure 5.8) has two arcs formed by dipole and multipole magnets, and two long straight sections where the electron cooler and experimental equipment (including the internal gas target are installed). The momentum acceptance is $\Delta p/p = 2.5\%$. The maximum magnetic rigidity of 10 Tm allows the storage of fully stripped uranium ions up to 560 MeV u^{-1}. Beam cooling (s.a.) is the prerequisite for beam storage as well as for precision experiments. The momentum spread is detected by Schottky noise probes and subsequent fast Fourier transformation (FFT) of the pickup signals.

Precision experiments can also be performed in the isochronous mode (Section 3.3), whereby no cooling is required. This experimental method is in particular well suited for stored short-lived ions. For these measurements, a TOF detector system has been installed in the ESR and recently also in the Chinese storage ring (CSR) [71]. Mass measurements of exotic nuclei have been performed in the isochronous mode in both rings. Half-life measurements of bare and highly-charged ions can be ideally performed with stored beams. In case, the magnetic rigidity change ($\Delta B\rho$), after nuclear decay, is within the storage acceptance of an ion ring, the daughter can be recorded with sensitive Schottky probes with high mass-over-charge resolution such that the particle identification and also the mass difference of the mother and daughter are obtained. The $\Delta B\rho$ larger than the daughter nuclides leave the closed orbits and have to be measured in detector systems positioned at dispersive locations with respect to where the decay has taken place. Nuclear reaction studies, characterized by low momentum transfer, require very thin targets and can be ideally performed in an internal gas target of 10^{13}–10^{15} atoms cm^{-2}, taking advantage of the circulation frequency of several megahertz. Reaction studies using the internal gas target are under exploration.

Figure 5.18 shows the scheme of the experimental facilities at the IMP in Lanzhou. The injector cyclotrons SFC and SSC feed the synchrotron ring CSRm. An in-flight fragment separator separates and transports the fragment beam to the cooler-storage ring CSRe (which is similar to the ESR at GSI), however, with longer straight sections, which is advantageous for TOF measurements and reaction studies at the internal target.

5.4.2 ISOL Facilities

The Danish physicists O. Kofoed-Hansen and K-O. Nielsen [49] produced short-lived krypton isotopes by bombarding a heated uranium target with neutrons produced using a deuteron beam from the Copenhagen cyclotron. The radioactive krypton isotopes were pumped toward an

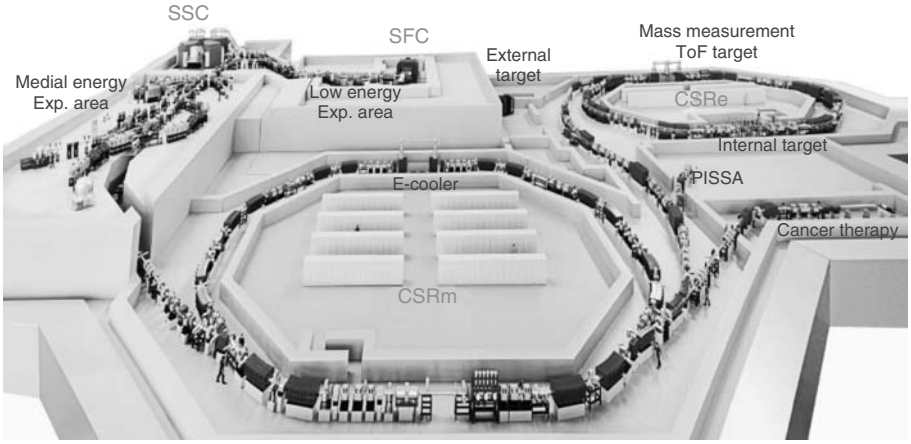


Figure 5.18 Scheme of the accelerator scenario and the experimental storage ring facilities built at the IMP in Lanzhou (China). The institute has based its research with heavy ions on the synchrotron CSRm and the storage-cooler ring CSRe. The connection of the two rings is realized by the fragment separator RIBLL. Courtesy: IMP Lanzhou.

ion source at the entrance of a mass analyzer. This development led to a new instrument for nuclear research, called an *isotope separator on-line*. The success of this ISOL separation technique led, in 1967, to the construction of the online isotope separator ISOLDE [21] at CERN. An intense high-energy (600 MeV u^{-1}) proton beam from the Synchro-Cyclotron (SC) produced exotic nuclei via spallation and fission reactions in thick targets connected to an ion source at the entrance of a mass separator called ISOLDE. After several upgrade programs, the ISOLDE facility is still in successful operation and has been the prototype of ISOL facilities worldwide. A wealth of ISOL facilities started all over the world, all with their specific characteristics, depending on the primary driver producing the radioactive nuclei (electron, proton (low or high energy), or heavy-ion accelerator, reactor, spontaneous fission source), the type of targets (thin or thick; solid, liquid, or gaseous), ion sources, and instrumentation. A new era started with the post acceleration of ISOL beams, enabling to perform reaction studies with radioactive nuclei. This was for the first time realized at Louvain-La-Neuve (Belgium), where two cyclotrons were coupled through an ISOL system [73, 74] and an important reaction for nuclear astrophysics using the radioactive ^{13}N beam was measured [75]. Since then, several post-accelerated ISOL facilities are operational or planned. As it is impossible to give an overview of all ISOL-based facilities worldwide, a number of facilities will be presented as examples of the different approaches that have been developed in order to provide high-quality beams of exotic nuclei all over the nuclear chart.

5.4.2.1 Low-Energy ISOL Facilities

ISOLDE at CERN (Switzerland) is the archetype ISOL facility: although it is the longest-running facility its many transformations, based on innovative developments, makes it still the place with the

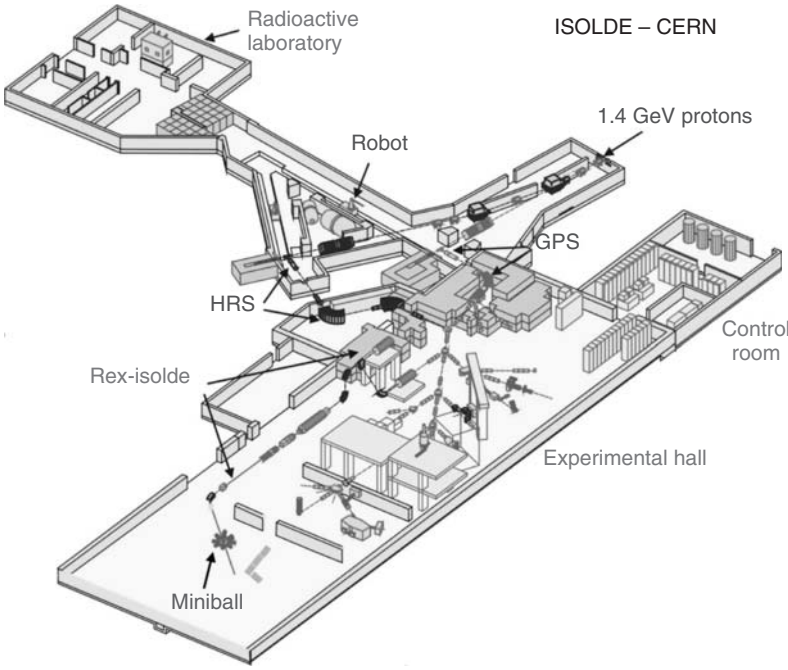


Figure 5.19 A typical ISOL facility, the ISOLDE at CERN [21].

largest catalog of different radioactive ion beams. Its extensive user community is interested in the low-energy beams as well as in the post-accelerated beams [76] (Figure 5.19).

The radioactive nuclei are produced in reactions of high-energy protons with intensities up to 2 μA from the PS-Booster accelerator on thick targets. The typical proton energies are between 1 and 1.4 GeV. These beams enable three different production mechanisms: spallation, fragmentation, and fission (see Section 5.2.1). More than 25 different target materials are used, which means that the neutron-deficient as well the neutron-rich region of the nuclear chart from hydrogen up to uranium is covered. The target material is kept at an elevated temperature so that the produced radioactive atoms diffuse out of the target into different dedicated ion sources. Ionization can take place in a hot plasma, on a hot surface, or by laser excitation (see Section 5.3.7). By judicious combinations of target-ion source systems, chemical selectivity may be obtained. This has resulted in selective production of isotopes of more than 70 of the chemical elements. In this way, ISOLDE has been able to deliver more than 700 beams of pure isotopes with intensities ranging from 1 to more than 10^{10} ions s^{-1} (Figure 5.20).

The future of ISOLDE is secured through the HIE-ISOLDE project, which contains three major elements: improvements in the beam quality and flexibility, higher beam intensities (up to a factor 10), and higher energies. The ISAC facility at TRIUMF, Vancouver (Canada) uses the same approach as ISOLDE, and although the proton energy is slightly less (600 MeV), its intensity can mount up to 100 μA [77].

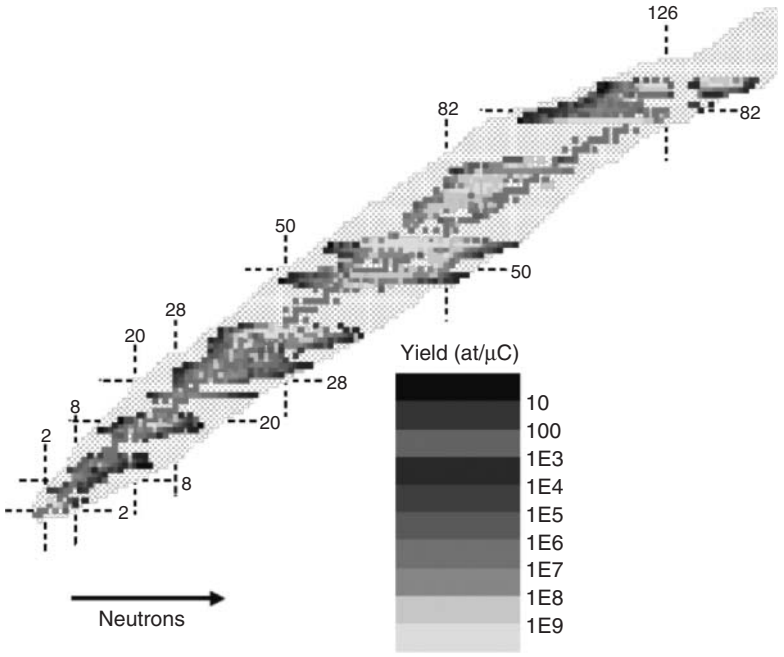


Figure 5.20 Overview of the beams produced at ISOLDE. The achieved yields in atoms per micro Coulomb of proton beam are presented with the color code.

Thin-Target Approach: the IGISOL Project Owing to their physico-chemical properties, a number of elements are difficult to extract efficiently or swiftly from the thick target or solid/liquid catcher (see Sections 5.3.6 and 5.3.7). An alternative is to use thin production targets and let the reaction products recoil out of the target and stop them in a cell filled with a noble gas. The IGISOL technique was developed in 1981 at the University of Jyväskylä cyclotron laboratory in Finland and continuously extended to different production mechanisms such as light- and heavy-ion fusion reactions and fission [78]. Although the primary production rate with the thin-target approach is considerably lower than with the thick-target approach, the IGISOL technique is certainly competitive for a number of elements (the so-called refractory elements) and for short-living nuclei with half-lives in the tens of millisecond range as the evacuation time of the gas cell can be very fast. The IGISOL mass separator facility is now moved to a new experimental area next to a new high current MCC30/15 light ion cyclotron. The beam from the JYFL heavy ion K-130 cyclotron can also be directed at the IGISOL target in the new location. This allows doubling the run time of the IGISOL facility to about 4000 h a year. Recently, a special type of gas catcher has been developed to make low-energy beams of fission products from a strong spontaneous fission source (^{252}Cf). This so-called Caribu project developed at Argonne National Laboratory (USA) has recently been commissioned and first beams were extracted.

The gas cell approach is not strongly dependent on the chemical properties of

the nuclei of interest, which also means that this technique is not element selective. Next to the beneficial aspects of this *universality*, it can also have drawbacks when the contamination of isobars is extremely high. A way to overcome this problem is to let the reaction products neutralize and then selectively re-ionize the element of interest by resonant laser light. This technique was pioneered at the LISOL facility of the KU Leuven (Belgium) and is discussed further.

5.4.2.2 High-Energy ISOL Facilities

The beam quality of the low-energy ISOL beams is very well adapted for further acceleration provided the 1^+ charge state is converted to a higher charge state in order to make efficient use of the accelerating process. Post acceleration can be provided by cyclotrons, tandems, or LINACs putting different constraints on the charge-state conversion process. This section follows the different approaches for charge-state conversion. Always one facility is selected to describe the possibilities, putting special emphasis on new developments. In 2011, a Focus Section of the *Journal of Physics G* was devoted to the "Physics with Reaccelerated ISOL beams" [79].

Charge-State Breeding in Electron Cyclotron Resonance Sources: the SPIRAL Project

ECR sources are extensively used at stable heavy-ion accelerators in order to reach the highest charge state with the highest intensity. Electron impact ionization where the electrons are heated by injecting RF power delivers multiple-charged ion beams. However, ECR sources are not that widely used at low-energy ISOL facilities, as their high efficiency is mainly for gaseous species and severe losses can occur for radioactive nuclei in non-gaseous form owing to sticking at the source walls. Furthermore, ECR sources are producing intense stable beams throughout the whole mass range, possibly contaminating the weak radioactive beams of interest. Therefore, they need high-resolution mass separation to suppress this background. Cyclotrons can do this job. It was a logical choice at Louvain-la-Neuve to use such a source for the production of the first radioactive ion beam especially as the aimed nuclei (^{13}N, ^{19}Ne, ...) came in gaseous form out of the target and the post accelerator was a cyclotron.

A running project using a cyclotron as post accelerator combined with a strong upgrade project are the so-called SPIRAL1 and SPIRAL2 projects carried out at GANIL, Caen, France [80]. The CIME cyclotron accelerates radioactive beams in the energy domain 2–25 MeV u^{-1}. These secondary beams are produced by the ISOL method using the very intense primary GANIL beam impinging on a thick production target. An R&D program on the target ion source systems is presently in progress, both for SPIRAL1 and for the future SPIRAL2 facility under construction. It involves the characterization of the 1^+ to n^+ conversion in an ECR source in order to bring the 1^+ beams from standard ISOL sources to a higher charge state. The driver of the new SPIRAL2 facility is a high power, CW, superconducting LINAC, delivering up to 5 mA of deuterons at 40 MeV (200 kW, the highest power ever delivered by this type of accelerator) directed on a Carbon converter + Uranium target. Production of the radioactive nuclear beams is based essentially on the neutron-induced fission of the uranium target. The expected radioactive ion beams intensities, in the mass range between $A = 60$ and $A = 140$, reaches, before post acceleration, up to 10^{10} particles per second for some species.

Charge-State Breeding in EBIS: the HIE-ISOLDE Project In 1994, a proposal was submitted to post accelerate the low-energy beams of ISOLDE called REX-ISOLDE, the acronym coming from ''Radioactive beam EXperiments at ISOLDE.'' Its key ingredient was an innovative charge breeding scheme combining bunching in a Penning trap with an EBIS. The philosophy was to separate as much as possible the production of singly-charged radioactive ions from the subsequent ion manipulation to provide bunches of multiply-charged ions needed for an efficient acceleration independent of their mass. The REX-Penning trap in front of the EBIS accumulates, bunches, and phase space cools the semi-continuous beam, leading to a considerably higher injection and trapping efficiency into the EBIS than obtained with continuous injection. In the EBIS the 1^+ to n^+ ion conversion takes place by electron bombardment by a magnetically compressed energetic electron beam accelerated over a DC potential. The efficiency for trapping, beam transport, charge breeding, and mass selection is typically between 5 and 15%. Sub-picoampere beams can be handled, and the upper throughput limit is a few 10^9 ions per second. A room-temperature accelerator boosts the beam energy to a maximum of about 3 MeV u^{-1}. Isotopes as light as ^{6}He and as heavy as ^{224}Ra have been used for Coulomb excitation, few-nucleon transfer reaction, or fusion evaporation studies [81].

The most important ingredient of HIE-ISOLDE, the major upgrade going on at ISOLDE, will come from replacing most of the existing REX accelerating structure by a superconducting (SC) linear accelerator with a maximum energy of 10 MeV u^{-1}. This will allow all ISOLDE beams to be accelerated to energies significantly above and well below the Coulomb barrier, facilitating a broad program of nuclear structure and nuclear astrophysics studies using different classes of nuclear reactions. A proposal is now under study to install the Heidelberg heavy-ion, low-energy ring TSR at the HIE-ISOLDE facility TSR. Such a facility will provide a capability for experiments with stored secondary beams that is unique worldwide.

The Use of Stripping Foils and Charge Exchange Cells: the ISAC and HRIBF Facilities Post acceleration of 60 keV 1^+ ions is difficult. At the ISAC facility in TRIUMF, this is done in two stages with a normal conductive cw RFQ followed by an IH structure room temperature DTL. A stripping section is present in the medium energy transport line between the two stages of acceleration. The beam energy that can be reached for $A \leq 30$ is fully variable between 0.153 and 1.53 MeV u^{-1}. The construction of a superconducting heavy ion LINAC was started in 2006 to add a further 40 MV to the ISAC beam energy. The installation known as *ISAC-II* reached its goal to accelerate $A/q = 6$ ions to energies around and above the Coulomb barrier. In order to increase the mass range of possible nuclei, up to $A < 150$ developments to add a charge state booster (1^+ to n^+) are ongoing. TRIUMF has recently started with the construction of ARIEL, the Advanced Rare Isotope Laboratory. The ARIEL complex comprises a new Superconducting 50 MeV 10 mA cw electron LINAC photo-fission driver and a number of new beamlines and target stations.

Another possibility to post accelerate ISOL beams is to use a tandem. Such a scheme has been developed at HRIBF, Oak Ridge using the Oak Ridge Isochronous Cyclotron as primary driver, and the 25 URC Tandem Accelerator for post

acceleration. As the Tandem will only accelerate negative ions, positive ions from the ISOL separator are charge exchanged. Although a successful research program was realized [82], the facility ceased operation on 15 April 2012.

5.4.3 Gas Catchers after In-Flight Separators

For a number of applications such as mass measurements in traps, laser spectroscopy, or decay studies with pointlike sources, the conversion of the high-energy radioactive ion beam from in-flight separators into a low-energy small emittance beam is required. Therefore, a gas catcher is placed at the focal plane of the in-flight separator, and the ions leaving the gas catcher can be subsequently treated in the standard ISOL way. The gas cell requirements for low-energy in-flight separators are very similar to those developed at ISOL systems, although the entrance window can be larger and thicker owing to the beam properties. For high-energy in-flight separators, the requirements of the gas cell have to be adapted to the range and range straggling of the energetic ions. This can be done in linear gas cells or circular ones.

5.4.3.1 Gas Catcher for Rare Fusion Reaction Products: IGLIS@S^3

At the LISOL facility, a laser ion source for the selective production of radioactive isotopes that is based on resonance ionization of radioactive reaction products thermalized in a buffer gas cell was developed. Recently the resonance ionization technique was extended to also perform laser-spectroscopy studies in and out of the gas cell [53, 83]. This new method is called the in-gas laser ionization and spectroscopy (IGLIS) technique.

The super separator spectrometer (S3) [108] is a device designed for experiments with very high intensity stable beams from the superconducting linear accelerator of GANIL, which is under construction in the framework of SPIRAL2. These beams can reach intensities exceeding 100 pμA for lighter ions. Such unprecedented intensities open new opportunities in several physics domains, for example, super-heavy and very-heavy nuclei, spectroscopy at and beyond the dripline. After A and Z selection, the fusion reaction products will be stopped in the gas cell and then resonantly re-ionized in broadband inside the gas cell or in narrow band in the supersonic gas jet for decay and laser spectroscopy studies. The improvements of this new approach in intensity (primary beam intensity and improved laser-gas-cell efficiency), selectivity, and spectral resolution should enable laser spectroscopy studies up to the nobelium region (Figure 5.21).

5.4.3.2 Gas Catcher for High-Energy Reaction Products

The combination of a separator with a gas-filled stopping cell is a recent development that is based on the relatively narrow range straggling of heavy ions. At high energies, where the contribution due to charge-exchange straggling is still negligible, the range straggling is about 1‰ of the mean range. It is obvious that an efficient stopping of energetic ions is only possible if the range straggling matches the thickness of the gaseous matter in the cell. In Figure 5.22, we illustrate with calculated range distributions of 300 MeV u^{-1} ^{132}Sn fragments the reduction of the range straggling with a monoenergetic degrader as a function of the ion-optical resolving power [109].

The procedure is to slow down the separated relativistic exotic nuclei in a

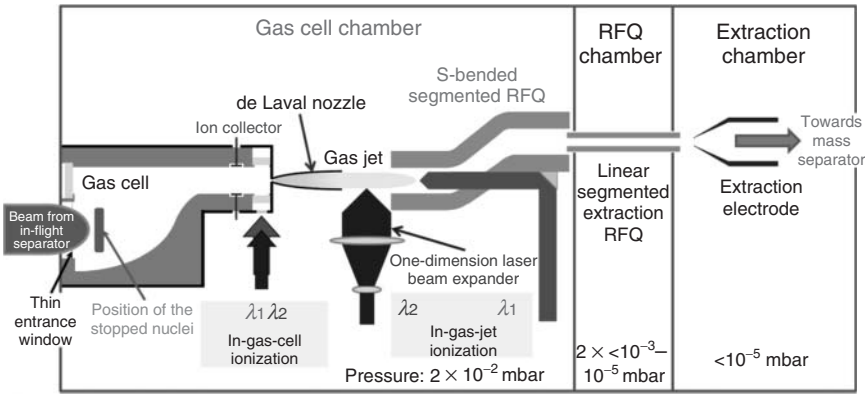


Figure 5.21 Schematic presentation of the IGLIS technique applied in the gas jet, outside the gas cell.

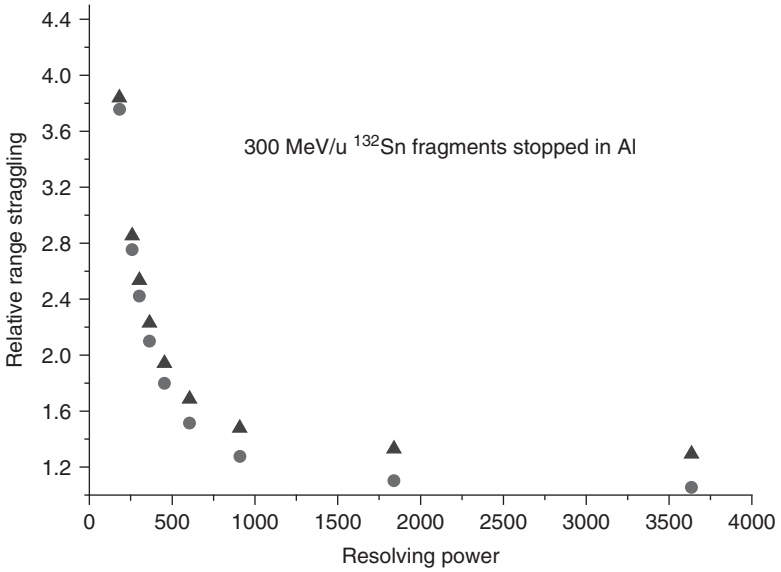


Figure 5.22 Calculated range straggling for a ^{132}Sn fragment beam at 300 MeV u^{-1}. The relative range straggling is normalized by the corresponding straggling of an ideal monoenergetic beam. The circles represent a first-order ion-optical calculation, whereas the triangles are a result of third-order calculation.

solid degrader to energies where the range distribution matches the thickness of the helium gas in the stopping cell. The time to slow down the relativistic ions in a solid degrader in front of the gas cell takes <1 ns, which is much less than the transit time through the ion-optical system. However, the fragments have a wide momentum distribution owing to their statistical formation process in the production target. Therefore, first the momentum distribution has to be bunched

before the fragment beam enters the gas volume of the stopping cell. Second, the ions must stay at least singly charged to be transported by electromagnetic fields through the gas cell within a few tens of milliseconds. The gas flow caused by pressure difference would be too slow and inefficient because of the large volume needed to stop the reaction products. Therefore, high electrical fields (DC) are applied to increase the mobility of the ions traveling through the gas volume. Special RF fields (RF carpet) are used in addition to guide the ions with minimal losses toward the exit nozzle. The goal to preserve the thermalized fragments ionized during their transport through the gas cell is achieved by choosing high-purity helium gas, characterized by a high ionization energy (24 eV).

The first challenge can be solved with a high-resolution dispersive spectrometer in combination with a monoenergetic degrader [34]. Considering the required ion-optical resolution, one has to take into account the large emittance of the incident fragment beam. It can have an emittance of about 300π mm mrad and a momentum spread of more than 5%, especially if projectile fission fragments are considered. In Figure 5.22, we present the calculated momentum compression resulting in a corresponding range bunching as a function of the ion-optical resolving power of the spectrometer. An ion-optical resolving power of 600–1000 is a cost-effective optimum with respect to the gain, because one can achieve under these conditions for a 300 MeV u^{-1} a straggling compression which approaches within (30–40)% of an ideal monoenergetic $\delta p_0 = 0$ incident beam. Of course the image aberrations of spectrometer have to be corrected with multipole fields, and the monoenergetic degrader shape has to be shaped to adapt to the residual focal plane distortions.

The second challenge is the limitation for the gas cell efficiency if a high-intensity beam enters the gas volume which would create a dense plasma during the stopping process. This situation can occur if the exotic nuclei of interest are accompanied by many contaminants or if relative high intensity fragment beam is prepared for post-acceleration experiments. A possible technical solution could be a gas-filled weak focusing cyclotron magnet equipped with collection electrodes. An RF carpet is employed for an efficient extraction through the orifice. As the cyclotron stopper has a large volume and low gas pressure, the space-charge effects by ionization of the helium atoms are strongly reduced and extraction times in the 5 ms range should be possible [33].

5.4.3.3 Novel Instrumentation coupled with Gas Catchers

After the successful implementation of Penning trap systems for exotic nuclei at ISOL facilities [84], such systems were also installed at low-energy in-flight separators in many laboratories. Recently, a cryogenic gas-filled ion-catcher combined with a MR TOF system has been installed also at relativistic energies [40].

The Penning trap at SHIP is a typical ion-trap setup combined with an in-flight separator. It is used for direct investigation of heavy-ion fusion products, including direct mass measurements, spectroscopy, and in future also for chemical reaction studies. After extraction out of the stopping cell, the ions are manipulated in a linear Paul trap where they are accumulated, bunched, and cooled in buffer gas. The exotic nuclei are then injected in a double Penning trap system. In the purification Penning trap with a mass resolving power

of about 100 000, the ions of interests are spatially separated and injected into the high-resolution trap, where precision mass measurements are performed. The preparation and measuring cycle of a Penning trap limit its use for very short-lived exotic nuclei.

Quite analogous to the operation of a heavy-ion storage ring with cooling or in the isochronous mode, we have the alternative for short-lived nuclei to use an isochronous MR-TOF system instead of the Penning trap. The stopping and extraction time of a fast ion catcher can be in the 10 ms range, and even below this time, the MR-TOF system can perform isobar separation, particle identification, high-resolution mass measurements, or guide the beam to another spectroscopy setup. The MR-TOF system consists of an RFQ and TOF sections, in which the ions can be cooled, bunched, mass separated with a resolution of several hundred thousand (presently achieved 600 000 max). It is ideally suited for experiments with cocktail beams because it can do simultaneous measurements over a large range, which a Penning trap cannot perform. Details of the complex device can be seen in Figure 5.8.

5.5 Milestones and Outlook

The vigorous development of intense radioactive ion beams, at low and high energies, combined with novel techniques for detection and identification has led to important new insights into the structure of the atomic nucleus and the stellar nucleosynthesis and also to applications in different fields such as condensed matter physics, accelerator physics, and life sciences. For example, new microscopic approaches taking into account previously neglected contributions to the nucleon–nucleon interaction, such as the tensor force, are developed to explain these new observations. To go beyond the shell model, developments have been recently initiated to link the nuclear force at work in atomic nuclei with quantum chromodynamics (QCD), a fundamental theory that describes the structure of the nucleons and their mutual interaction. It becomes obvious that new experimental and theoretical work is needed to understand the strong interaction on the microscopic level. Nuclear-structure research is driven by a continuous exploration and extension of the nuclides in the Segrè chart (Figure 5.1) to challenge the paradigms developed near stability, and to look for the limits of nuclear existence and for new phenomena of loosely bound quantum systems at and beyond the drip line.

5.5.1 Mapping the Boundaries of Nuclear Existence

The proton dripline has been reached within a few isotopes, whereas the neutron dripline has been experimentally verified only up to oxygen (see particle identification spectrum in Figure 5.23. Each dot indicates an individual atomic nucleus, each cluster an isotope. Such an explorative spectrum gives already important results: the existence of an isotope, its formation cross section and, in certain cases, the lifetime. The upper boundary of the nuclear chart is determined by fission and is called the *Super Heavy Elements* region. The key question of heavy-element research is the existence of a shell-stabilized island. This island is theoretically predicted to be at $Z = 114$ or 120 and $N = 184$ for the next

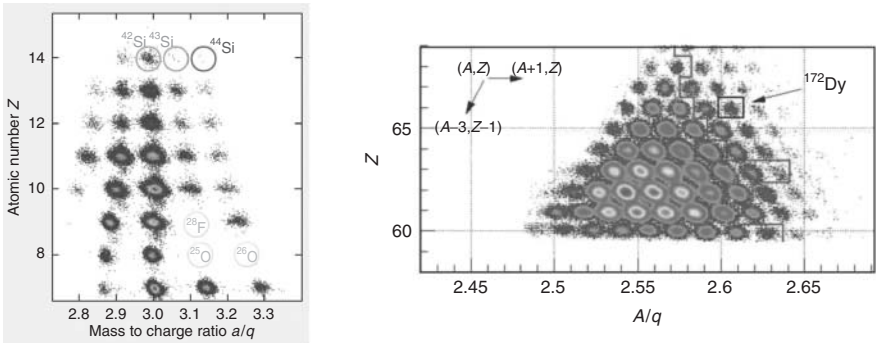


Figure 5.23 Measured particle identification for fragments created with a 142 MeV u^{-1} ^{48}Ca beam [45] (a) and a 1000 MeV u^{-1} ^{238}U projectile beam (b). The reaction products are spatially separated in-flight with the $B\rho-\Delta E-B\rho$ method, and were identified by TOF, energy-deposition, and magnetic rigidity measurements. The experiment (a) clearly demonstrates that the neutron dripline is reached with ^{24}O and that the ^{25,26}O and ^{28}F ions are particle instable. New isotopes discovered with the uranium beam are on the right-hand side of the red line in figure (b) [46] Ref. [88].

doubly magic spherical shell above lead. The present situation is that the elements from 106 up to 113 have been discovered via cold fusion reactions [85]. The unambiguous identification has been performed via α-decay spectroscopy following the mother–daughter relation until the chains reach known isotopes. The heavier isotopes up to $Z = 118$ have been created in hot fusion reactions [64]. For these nuclides, the identification is more difficult because their α-decay chains do not enter known territory but fission. The main goal in the SHE research still is to reach the magic proton number (which seems to be beyond $Z = 118$) and study the properties of the discovered isotopes in more detail.

A new era of the discovery and investigation of heavy isotopes has been opened up with fragmentation and fission of relativistic uranium projectiles [14]. The first spatial isotopic separation of relativistic uranium projectile fragments [86] has given access to new fission studies at low excitation energies, practically, for all fragments created with the relativistic uranium beam [87]. Uranium beams have been a rich source for discovering and investigating heavy neutron-rich isotopes, as demonstrated in recent experiments [46, 88].

5.5.2 Selected Milestones

Guided by important technical developments in production methods and instrumentation, a choice of pioneering efforts and subsequent advances sketching their influence on the progress of the field of radioactive ion beam research will be given.

- New short-lived Kr isotopes identified through their characteristic decay [72].

  This experiment started decay spectroscopy studies of ISOL and in-flight produced radioactive ion beams revealing exotic decay modes such as one- and two-proton decay, neutron decay, cluster decay, and β-delayed fission [89–91]. By coupling in-flight projectile fragment separators to ion

storage and cooler rings for the first time, radioactive decay of bare and highly-ionized atoms have been studied [92, 93], and it was shown that nuclides that are stable as neutral atoms can become radioactive when all the bound electrons are removed, the so-called bound β-decay. A novel basic feature of nuclear electron capture has been found and explained in decay studies of hydrogen- and heliumlike heavy nuclides [94, 95].

- Odd-even staggering of neutron deficient Hg isotopes studied using optical spectroscopy evidencing shape coexistence in heavy nuclei [96]

These pioneering efforts triggered a large program of ground-state property investigations (size, electromagnetic moments, and spins) through collinear laser spectroscopy and resonance ionization spectroscopy [97]. The latter technique is now a standard approach for producing pure, even isometrically pure, radioactive ion beams using the ISOL technique, while laser cooling is implemented in storage rings to reach ultimate phase space conditions.

- Projectile fragmentation at relativistic energies and reaction studies led to the discovery of many new isotopes and nuclear haloes [27, 98].

A new production method and a new class of reaction studies involving radioactive isotopes were initiated. These include total interaction cross section measurements leading to matter radii, knock-out reactions revealing single-particle structures, and Coulomb excitation resulting in collective characteristics. Halo matter distributions have been discovered for weakly bound neutrons [27, 99] and protons [100]. Measurements along isotope chains have clearly revealed the skin structure of neutron-rich nuclides [101]. Subsequently, also ISOL beams were post accelerated, allowing for the first time capture reaction studies of interest to nuclear astrophysics [75]. These energetic beams near and slightly above the Coulomb barrier allowed furthermore Coulomb excitation, few-nucleon transfer, and fusion reactions. Energetic beams from both in-flight and ISOL facilities were used to map the island of inversion around ^{32}Mg, evidencing the weakening of the $N = 20$ shell closure far from stability.

- Precise mass measurements using a Penning trap coupled to an ISOL beam [102] and a storage ring after an in-flight facility [28].

Precision mass measurement programs are currently ongoing to map the nuclear mass surface to observe shell closures. While Penning traps can lead to ultra-high precision measurements for fundamental interaction studies, storage rings represent a complementary approach, ideally suited for very short-lived bare and few-electron isotopes. Ion trapping and cooling is now readily used for beam purification and manipulation, and applied to in-trap assisted decay spectroscopy.

5.5.3 Relevance to Other Fields of Science

Radioactive beams are relevant for many different research fields. Masses, decay properties, and reaction cross sections involving far unstable nuclei form together with nuclear model predictions direct input in network calculations of stellar evolution, cosmo-chronometers, and other astrophysical applications [4, 5]. A number of exotic nuclei are also ideally suited for studying

the low-energy sector in fundamental interaction research, such as the unitarity CKM, neutrinoless double β-decay, and electric dipole moment searches [103]. SHEs, produced in small quantities, challenge atomic physics and chemistry at the end of Mendeleev's table, where relativistic effects are strong [104]. In minute amounts and at low energy, radioactive nuclei are used as probes in different materials ranging from condensed matter to biological systems, without altering the basic properties of substrate [105]. A number of isotopes have radioactive decay and chemical properties that make them suitable for innovative radiopharmaceuticals and could be produced in large amounts directly as beams at ISOL facilities or indirectly being extracted from beam dumps and collimators. Overcoming the challenges in the quest for ever increasing the intensity of radioactive ion beams has also a direct effect in other fields such as heavy-ion hadron therapy, spallation neutron sources, and accelerator-driven systems.

5.5.4 Outlook

The open questions that are currently under study at different exotic nuclear beam facilities have been defined and extensively discussed by the scientific community. The most essential questions are:

- How can we describe the rich variety of low-energy structure and reactions of nuclei in terms of the fundamental interactions between individual nucleons?
- How can we predict the evolution of nuclear collective and single-particle properties as functions of mass, isospin, angular momentum, and temperature?
- How and where are the elements made in the universe?
- Can we recreate on Earth, and understand, the critical reactions that drive the energy generation and the associated synthesis of new elements in the stars?

A brochure explaining the general aims of nuclear-physics research and radioactive ion beams particularly can be downloaded from *www.nupecc.org/lrp2010/Documents/lrp2010_booklet_final.pdf* [106]. The subjects include ''inside the nucleus,'' ''where do the elements come from,'' ''exploring the nuclear landscape,'' ''exploring the first level of complexity,'' ''from nuclear liquid to quark-gluon plasma,'' ''the forces of nature,'' and ''rewards from nuclear science.''

Glossary

α particle: Nucleus of the helium atom; it consists of two protons and two neutrons.

Barn: A barn (symbol b) is the area to characterize the cross sections of any scattering or reaction in nuclear reactions; 1 b $=10^{-28}\,m^2$ Picobarn: 1 pb $= 10^{-12}$ b $= 10^{-40}\,m^2$ Femtobarn: 1 fb $= 10^{-15}$b $= 10^{-43}\,m^2$

β-particle: Name of electrons or positrons emitted from nuclei in the β-decay.

Chart of nuclides: Sometimes it is also called *''Segrè chart.''* The experimentally known nuclei are arranged in a matrix with the proton number on the vertical axis and the corresponding neutron number on the horizontal axis. Each nucleus is presented by a square box, which can include the basic properties such as the decay mode and half-life.

Driplines: There are two driplines: the proton and the neutron dripline. The driplines are indicated on the chart of nuclides. The last isotopes of an element that can form a nucleus, an additional or one neutron less, lead to an unbound system.

Exotic nucleus: A short-lived nucleus whose proton-to-neutron ratio is very different from those nuclei in the valley of β-stability. Synonymously used terms are *"rare isotope"* and *"nucleus far from stability."*

Gas catcher: Gas-filled stopping cell where the reaction products are slowed down and extracted as high-quality (small emittance) beams for precision experiments with traps, spectroscopy detectors, or for reacceleration with an accelerator facility. Such gas-catcher ion sources are applied with ISOL and in-flight facilities. Energetic, spatially separated nuclei are slowed down in solid degraders before they enter the gas volume (rare gas, usually helium) to be thermalized. In this fast slowing-down process, the exotic nuclei remain ionized and thus can be extracted with DC and RF fields.

In-flight separator/method: A production and separation method in which the exotic nuclei emerge from relatively thin targets with high kinetic energies. The access to nuclides with short lifetimes is only limited by the time of flight through the electromagnetic fields of the separator, which is in the microsecond range and below.

Ion trap: A small electromagnetic device that stores the injected low-energy ions in stable orbits. A trap can be used for precision experiments based on frequency determinations of the stored ions. A trap can also be used as a final separator stage that can provide isobar and even isomer separation.

Isobar: Nuclides with the same atomic mass number (A). A is the sum of the number of protons (Z) and neutrons (N). $A = N + Z$.

ISOL separator/method: Isotope separator on line is a method for producing exotic nuclei in thick targets. The interesting reaction products have to be extracted from the target, which is usually part of the ion source. The exotic nuclei have to effuse from the heated target, and diffuse into the ion source where the atoms will be ionized and extracted by electromagnetic fields.

Isomer: A nucleus that is in a long-lived excited state and cannot easily decay via γ-emission.

Isotopes: An element can have many isotopes up to the driplines. They all have the same number of protons but different numbers of neutrons.

Magic numbers: Some nuclides have an increased stability owing to a particular number of protons and neutrons. The magic neutron numbers are 2, 8, 20, 28, 50, 82, and 126; for protons, they are the same, but the element $Z = 126$ has not been discovered yet.

Nucleon: Common term for either a proton or a neutron.

Storage rings: Energetic ions are injected and stored in the ultra-high vacuum of an electromagnetic device with a closed orbit, that is, the stored ions circulate "undisturbed" in the ring lattice for further investigations. The phase-space density of the stored exotic nuclei can be dramatically increased with different cooling procedures. A frequency measurement of a stored and cooled beam can be ideally used for direct mass measurements of exotic nuclei. A circulating beam has normally

a revolution frequency in the megahertz range such that the interaction with very thin internal targets can be used for reaction studies with low momentum transfer. The non-reacted ions are constantly cooled and in this way recirculated for reactions.

References

1. Mackinstosh, R., Al-Khalili, J., and Jonson, B. (2001) *Nucleus: A Trip into the Heart of Matter*, Canopus Publishing Ltd, Bath.
2. Jolie J. (2012) this encyclopedia. Nuclear Structure.
3. Bertulani C. A. (2012) this encyclopedia. Nuclear Reactions.
4. Schatz H. (2012) this encyclopedia. Nuclear Astrophysics.
5. Thielemann F. K. (2012) this encyclopedia. Supernovae and Their Nucleosynthesis.
6. EMIS-15 (2008) XVth International Conference on Electromagnetic Isotope Separators and Techniques Related to their Application. *Nucl. Instrum. Methods Phys. Res. B*, **266**, 4067–4726.
7. Geesaman, D.F. *et al.* (2006) *Annu. Rev. Nucl. Part. Sci.*, **56**, 53.
8. Thomson, J.J. (1912). *Phil. Mag.*, **24**, 209.
9. Aston, F.W. (1921) *Nature*, **107**, 334.
10. Armbruster, P. and Münzenberg, G. (2012) *Eur. Phys. J. H*, **37**, 237–309. doi: 10.1140/epjh/e2012-20046-7
11. von Weizsäcker, C.F. (1935) *Z. Phys.*, **96**, 431.
12. Bertulani, C.A. and Danielewicz, P. (2004) *Introduction to Nuclear Reactions*, IOP Publishing, Bristol ISBN: 07503 09326.
13. Münzenberg, G. *et al.* (2011) in *Handbook of Nuclear Chemistry*, 2nd 2011 edn, Vol. 2 (eds A. Vértes, S. Nagy, and Z. Klencsár), Springer.
14. Bernas, M. *et al.* (1994) *Phys. Lett. B*, **331**, 19.
15. Gaimard, J.J. and Schmidt, K.-H. (1991) *Nucl. Phys. A*, **531**, 709.
16. Heckman, H.H. *et al.* (1972) *Phys. Rev. Lett.*, **28**, 926.
17. Sümmerer, K. and Blank, B. (2000) *Phys. Rev.*, **C61**, 034607.
18. Hinke, C. *et al.* (2012) *Nature*, **486**, 341.
19. Morrissey, D.J. and Sherrill, B.M. (2004) *Lect. Notes Phys.*, **651**, 113.
20. Geissel, H., Münzenberg, G., and Riisager, C. (1995) *Ann. Rev. Nucl. Part. Sci.*, **45**, 163.
21. Ravn, H. and Allardyce, B. (1989) in *Treatise on Heavy Ion Science*, Vol. 8 (ed D.A. Bromley), Plenum, New York, p. 363.
22. Van Duppen, P. (2006) *Lect. Notes Phys.*, **2**, 37 ISBN-13 978-3-540-33786-7 Springer New York.
23. Äystö, J. *et al.* (1984) *Phys. Lett.*, **138B**, 369.
24. Münzenberg, G. (1994) *Nucl. Instrum. Methods Phys. Res. B*, **70**, 265.
25. EMIS (1997) XIIIth International Conference on Electromagnetic Isotope Separators and Techniques Related to their Application. *Nucl. Instrum. Methods Phys. Res. B*, **126**, 1–427.
26. EMIS (2003) XIVth International Conference on Electromagnetic Isotope Separators and Techniques Related to their Application. *Nucl. Instrum. Methods Phys. Res. B*, **204**, 1–796.
27. Tanihata, I. *et al.* (1985) *Phys. Rev. Lett.*, **55**, 2676.
28. Geissel, H. *et al.* (1992) *Phys. Rev. Lett.*, **68**, 3412.
29. Tu *et al.* (2011). *Phys. Rev. Lett. PRL* **106**, 112501.
30. Savard, G. *et al.* (2003) *Nucl. Instrum. Methods Phys. Res. B*, **204**, 582.
31. Scheidenberger, C. *et al.* (2003) *Nucl. Instrum. Methods Phys. Res. B*, **204**, 119.
32. Wada, M. *et al.* (2003) *Nucl. Instrum. Methods Phys. Res. B*, **204**, 570.
33. Bollen, G. *et al.* (2005) *Nucl. Instrum. Methods Phys. Res. A*, **550**, 27.
34. Geissel, H. *et al.* (1989) *Nucl. Instrum. Methods Phys. Res. A*, **282**, 247.
35. Franzke, B., Geissel, H., and Münzenberg, G. (2008) *Mass Spectrosc. Rev.*, **27**, 428.
36. Yavor, M. (2009) *Optics of Charged Particle Analyzers*, Advances in Imaging and Electron Physics, Vol. 157, Academic Press, San Diego, CA.
37. Wollnik, H. (1987) *Optics of Charged Particle Beams*, Academic Press, New York.
38. Wouters, J.M. *et al.* (1987) *Nucl. Instrum. Methods Phys. Res. B*, **26**, 286.
39. Blaum, K. *et al.* (2006) *Phys. Rep.*, **425**, 1.
40. Plaß, W.R. *et al.* (2008) *Nucl. Instrum. Methods Phys. Res. B*, **266**, 4560.
41. Franzke, B. (1987) *Nucl. Instrum. Methods B*, **24/25**, 18.

42. Leino, M. *et al.* (2003) *Nucl. Instrum. Methods B*, **204**, 129.
43. Geissel, H. *et al.* (1992) *Nucl. Instrum. Methods B*, **70**, 286.
44. Scheidenberger, C. *et al.* (1994) *Phys. Rev. Lett.*, **73**, 50.
45. Tarasov, O.B. *et al.* (2007) *Phys. Rev. C*, **75**, 064613.
46. Kurcewicz, J. *et al.* (2012) *Phys. Lett. B*, **717**, 371.
47. Ravn, H.L. (1979) *Phys. Rep.*, **54**, 201.
48. Kirchner, R. (1992) *Nucl. Instrum. Methods B*, **70**, 186.
49. Kofoed-Hansen, O. and Nielsen, K.O. (1951) *Phys. Rev.*, **82**, 96.
50. Mishin, V.I. *et al.* (1993) *Nucl. Instrum. Methods B*, **73**, 550.
51. Kudryavtsev, Y. *et al.* (1996) *Nucl. Instrum. Methods B*, **114**, 350.
52. Blaum, K. *et al.* (2003) *Nucl. Instrum. Methods B*, **204**, 331.
53. Sonoda, T. *et al.* (2009) *Nucl. Instrum. Methods B*, **267**, 2918.
54. Äystö, J. (2001) *Nucl. Phys. A*, **693**, 477.
55. Bollen, G. (2004) *Lect. Notes Phys.*, **651**, 169.
56. Budker, G.I. (1966) Proceedings of the International Symposium on Electron and Positron Storage Rings, Saclay, 1966, p. 11-1-1.
57. Poth, H. (1990) *Phys. Rep.*, **196**, 135–297.
58. Steck, M. *et al.* (1996) *Phys. Rev. Lett.*, **77**, 3803.
59. Geissel, H. *et al.* (2003) *Nucl. Instrum. Methods Phys. Res. B*, **204**, 71.
60. Münzenberg, G. and Schädel, M. (1996) *Moderne Alchemie: Die jagd nach den Schwersten Elementen*, Vieweg, Wiesbaden ISBN: 3-528-06474-9.
61. Hofmann, S. (2002) *On Beyond Uranium*, Taylor & Francis, London.
62. Hofmann, S. and Münzenberg, G. (2000) *Rev. Mod. Phys.*, **72**, 733.
63. Morita, K. *et al* (2004). *J. Phys. Soc. Jpn.*, **73**, 2593.
64. Oganessian, Yu. (2007) *J. Phys. G: Nucl. Part. Phys.*, **34**, R165.
65. Armbruster, P. (1961) *Nukleonik*, **3**, 188.
66. Tsyganov, Yu.S. (2011) *Phys. Part. Nucl.*, **42** (5), 812–845.
67. Morrissey, D.J. (2003) *Nucl. Instrum. Methods Phys. Res. B*, **204**, 90.
68. Kubo, T. *et al.* (2003) *Nucl. Instrum. Methods Phys. Res. B*, **204**, 97.
69. Simon, H. (2010) *Nucl. Phys. A*, **834**, 394c.
70. Bazin, D. *et al.* (2003) *Nucl. Instrum. Methods Phys. Res. B*, **204**, 629.
71. Tu, X.L. *et al.* (2002) *Nucl. Instrum. Methods Phys. Res. A*, **654**, 213.
72. Kofoed-Hansen, O. and Nielsen, K.O. (1951) *Phys. Rev.*, **82**, 96.
73. Darquennes, D. (1990) *Phys. Rev. C*, **42**, R804.
74. Huyse, M. and Raabe, R. (2011) *J. Phys. G: Nucl. Part. Phys.*, **38**, 024001.
75. Decrock, P. *et al.* (1991) *Phys. Rev. Lett.*, **67**, 808.
76. Jonson, B. and Riisager, K. (2010) *Scholarpedia*, **5** (7), 9742.
77. Ball, G. *et al.* (2011) *J. Phys. G: Nucl. Part. Phys.*, **38**, 024003.
78. Topical Collection: Äystö, J. *et al.* (2012) *Eur. Phys. J. A*, **48**.
79. Physics with Reaccelerated ISOL Beams (2011) *J. Phys. G: Nucl. Part. Phys.*, **38**, 020301–24005.
80. Navin, A. *et al.* (2011) *J. Phys. G: Nucl. Part. Phys.*, **38**, 024004.
81. Van Duppen, P. and Riisager, K. (2011) *J. Phys. G: Nucl. Part. Phys.*, **38**, 024005.
82. Beene, J.R. *et al.* (2011) *J. Phys. G: Nucl Part. Phys.*, **38**, 024002.
83. Cocolios, T. (2009) *Phys. Rev. Lett.*, **103**, 102501.
84. Bollen, G. *et al.* (2001) *Hyperfine Interact.*, **132**, 215.
85. Hofmann S. (2012) in this encyclopedia. Superheavy Nuclei.
86. Magel, A. *et al.* (1994) *Nucl. Instrum. Methods B*, **94**, 548.
87. Schmidt, K.-H. *et al.* (1994) *Phys. Lett. B*, **325**, 313.
88. Ohnishi, T. *et al.* (2010) *J. Phys. Soc. Jpn.*, **79**, 073201.
89. Hofmann, S. *et al.* (1982) *Z. Phys.*, **A305**, 111.
90. Pfützner, M. *et al.* (2002) *Eur. Phys. J.*, **A14**, 279.
91. Giovannozzo, J. *et al.* (2002) *Phys. Rev. Lett.*, **89**, 102501.
92. Jung, M. *et al.* (1992) *Phys. Rev. Lett.*, **69**, 2164.
93. Irnich, H. *et al.* (1995) *Phys. Rev. Lett.*, **75**, 4182.
94. Litvinov, Yu.A. *et al.* (2007) *Phys. Rev. Lett.*, **99**, 262501.
95. Patyk, Z. *et al.* (2008) *Phys. Rev.*, **C77**, 014306.

96. Bonn, J. *et al.* (1972) *Phys. Lett.*, **38B**, 308.
97. Kluge, H.J. and Nörtershäuser, W. (2003) *Spectrochim. Acta*, **B58**, 1031.
98. Westfall *et al.* (1979) *Phys. Rev. Lett.*, **43**, 1859.
99. Hansen, P.G. and Jonson, B. (1989) *Europhys. Lett.*, **4**, 409.
100. Schwab, W. *et al.* (1995) *Z. Phys. A*, **350**, 283.
101. Suzuki, T. *et al.* (1995) *Phys. Rev. Lett.*, **75**, 3241.
102. Bollen, G. *et al.* (1987) *Hyperfine Interact.*, **38**, 793.
103. Severijns, N., Beck, M., and Naviliat-Cuncic, O. (2006) *Rev. Mod. Phys.*, **78**, 991.
104. Schädel, M. (2012) *Radiochim. Acta*, **100**, 579.
105. Forkel-Wirth, D. (1999) *Rep. Prog. Phys.*, **62**, 527.
106. NuPECC (2010) *Nuclear Physics: Science and Application*, European Science Foundation *www.nupecc.org/lrp2010/Documents/lrp2010_booklet_final.pdf* (accessed 11 February 2013).
107. Thoennessen, M. (2011) *Nature*, **478**, 160 and *Nucl. Phys. News* **22**, 3, 2012.
108. *http://pro.ganil-spiral2.eu/spiral2/instrumentation/s3*
109. Winfield, J.S. *et al.* (2013) Nuclear Instruments and Methods in Physics Research A, **704**, 76–83.

6
Superheavy Nuclei

Sigurd Hofmann

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

6.1 Introduction

Searching for superheavy nuclei (SHN) is an attempt to answer questions of partly fundamental character: How many elements may exist? How long is their lifetime? Which properties determine their stability? How can they be synthesized? How are the electrons arranged in the strong electric field of the nucleus? What are the chemical properties of the atoms?

The properties of nuclei are not smooth uniform functions of the proton and neutron numbers, but show nonuniformities as evidenced by variations in the measured atomic masses. Similar to the electrons in an atom, also the nucleons in a nucleus – described by quantum mechanical laws – form closed shells called *magic* numbers. At the magic proton (Z) or neutron (N) numbers 2, 8, 20, 28, 50, and 82, the nuclei have an increased binding energy relative to the average trend. For neutrons, $N = 126$ is also identified as a magic number. However, the highest stability is observed in the case of the *doubly magic* nuclei with a closed shell for both protons and neutrons. Among other special properties, the doubly magic nuclei are spherical and resist deformation. In the chart of nuclei (Figure 6.1), all known nuclei are plotted in a plane spanned by the number of protons and neutrons. The magic nuclei are emphasized.

The magic numbers were successfully explained by the nuclear shell model, and an extrapolation into unknown regions was reasonable. The numbers 126 for the protons and 184 for the neutrons were predicted to be the next shell closures. Instead of 126 for the protons also 114 or 120 were calculated as closed shells. The term superheavy elements (SHEs) was coined for these elements. More appropriate, however, is the term SHN, because the lifetime of the elements is determined by the whole atomic nucleus and not only by the number of protons. In some cases the number of neutrons is even more decisive for the stability of the atomic nucleus than the number of protons.

The prediction of magic numbers, although not unambiguous, was less problematic than the calculation of the stability of those doubly closed shell nuclei against fission. As a consequence, half-lives predicted based on various calculations differed by many orders of magnitude. Some of the half-lives approached the age of the universe, and attempts have been made to discover naturally occurring SHN. Although discoveries were

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

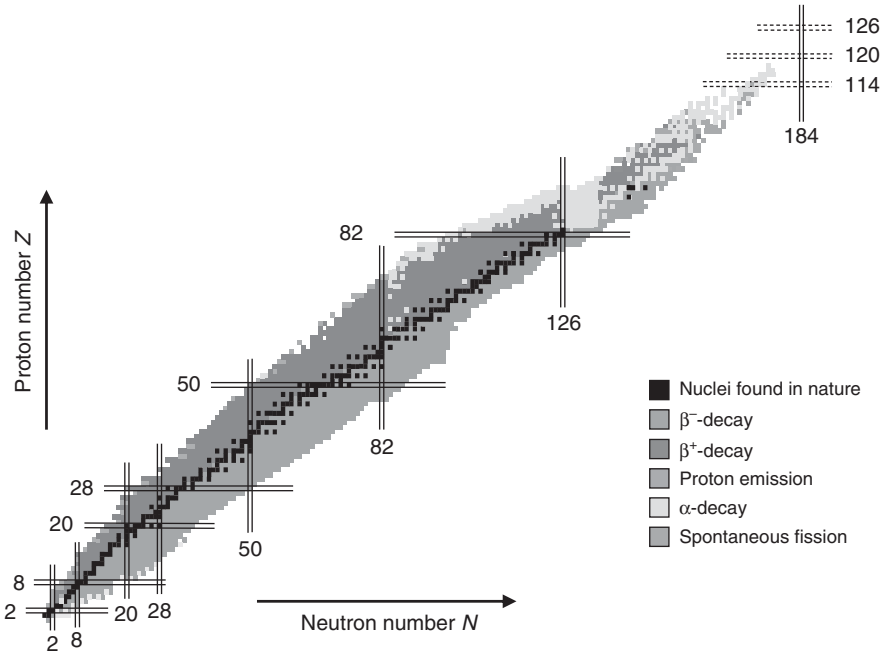


Figure 6.1 The chart of nuclei shows all known nuclei in a coordinate system spanned by the number of protons (Z) and the number of neutrons (N). Black squares mark the nuclei that exist in nature. Most of them are stable. Squares in lighter shade characterize the decay modes of radioactive nuclei. The magic numbers are marked by horizontal and vertical double lines. Superheavy nuclei are located in the region of the double magic nuclei at $Z = 114$, 120, or 126 and at $N = 184$.

announced from time to time, none could be substantiated after more detailed inspection.

There was also great uncertainty of the production yields for SHN. Closely related to the fission probability of SHN in the ground state, the survival of the compound nuclei (CN) formed after complete fusion of stable lighter nuclei was difficult to predict. Even the best choice of the reaction mechanism, fusion or transfer of nucleons, was critically debated. A drawing of the two different reaction types is shown in Figure 6.2. However, as soon as experiments could be performed without technical limitations, it turned out that the most successful methods for the laboratory synthesis of heavy elements are fusion-evaporation reactions using heavy-element targets, recoil-separation techniques, and the identification of the nuclei by generic ties to known daughter decays after implantation into position-sensitive detectors.

In the following sections, a detailed description of the properties that determine the stability of SHN is given. Included are distinct features that occur only in heavy nuclei. Experimental setups and experiments are presented, in which fusion reactions were used for studying the stability of SHN. Experiments and theoretical attempts related to the synthesis of SHN are outlined. Finally, a summary and an outlook are given.

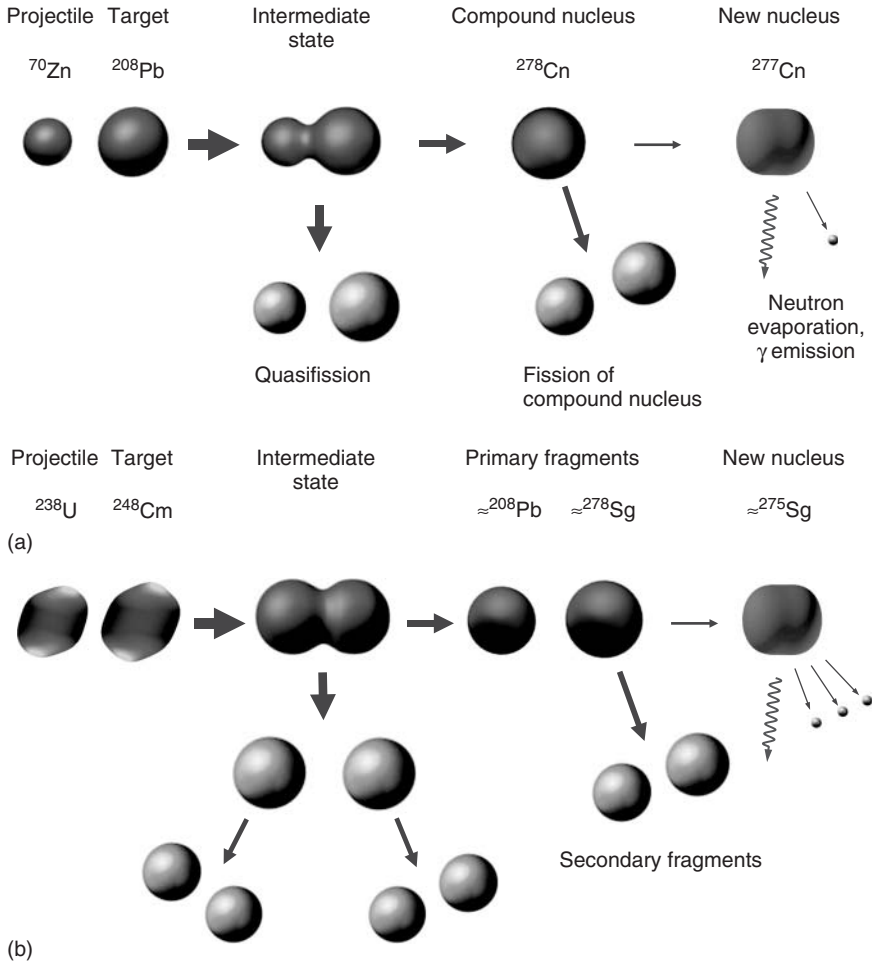


Figure 6.2 Schematic presentation of a fusion reaction (a) and a transfer reaction (b) both resulting in heavy nuclei. The main reaction channels of the fusion reaction are quasifission and fission of the compound nucleus. Only in few cases an ER is formed after emission of one neutron and γ-rays. The shape of the isotope ^{277}Cn is due to its large positive quadrupole and positive hexadecupole deformation. The path resulting in neutron-rich heavy nuclei in the case of the transfer reaction is due to the formation of a double magic ^{208}Pb nucleus as one of the primary quasifission fragments. The thickness of the arrows represents the main reaction channels. The shape of the projectile ^{238}U and the target ^{248}Cm is due to their large positive quadrupole and negative hexadecupole deformation.

6.2 Stability of Superheavy Nuclei

The calculation of the *ground-state binding energy* provides the basic step to determine the stability of SHN. Most accurate predictions are obtained by the *Strutinsky approach* [1], who considered the shell effect as a small deviation from a uniform single-particle energy-level distribution. This deviation, called *shell-correction energy*, was then used as a correction to

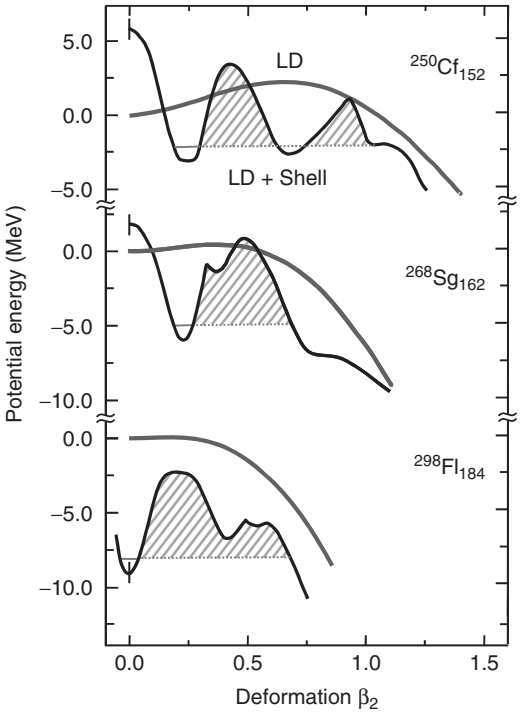


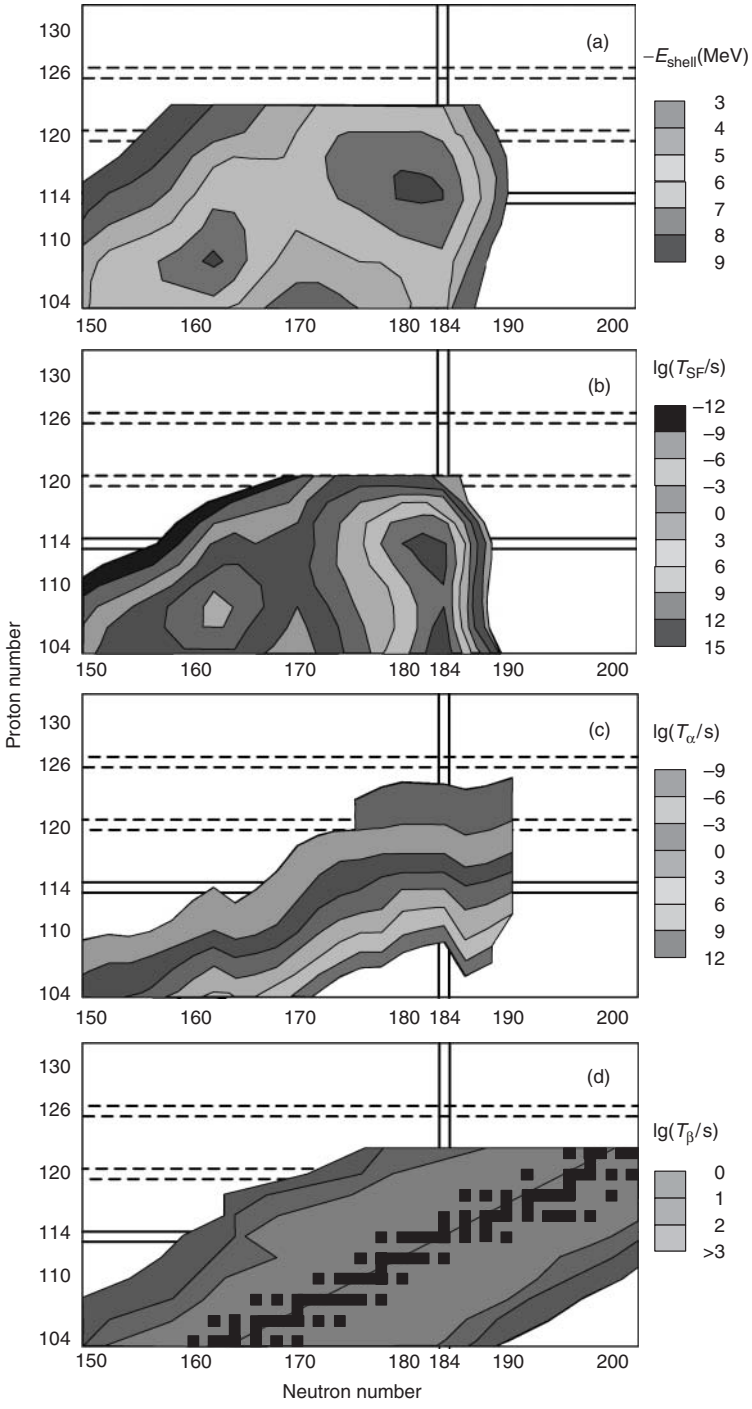
Figure 6.3 The charged liquid-drop model (bold curve) predicts that the nucleus ^{250}Cf (top, $Z = 98$) has a spherical ground state and a wide fission barrier. Taking into account the shell effects, a deformed ground state (first minimum from the left) and a double-humped fission barrier is obtained. The hatched area is a measure of the fission half-life. In ^{268}Sg (center, $Z = 106$) the liquid-drop fission barrier almost disappeared, nevertheless, owing to shell effects, a pronounced barrier exists. Moreover, this nucleus has a deformed ground state. In the superheavy nucleus ^{298}Fl ($Z = 114$), shell effects force a spherical ground state and, still, a wide and high fission barrier exists. (The figure has been provided by courtesy of S. Cwiok and A. Sobiczewski [2].)

the *macroscopic* liquid-drop model energy. The term *macroscopic–microscopic* model was invented for this approach. The shell-correction energies of the ground state of nuclei near closed shells are negative, which results in further decreased values of the negative binding energy from the liquid-drop model – and thus an increased stability. An experimentally accessible signature for the shell-correction energy is obtained by subtracting a calculated smooth macroscopic part from the measured total binding energy.

However, the knowledge of ground-state binding energies is not sufficient for the calculation of partial *spontaneous fission* (*SF*) half-lives. Here it is necessary to determine the size of the *fission barrier* over a wide range of deformation. The most accurate data were obtained for even–even nuclei using the macroscopic–microscopic model. How the macroscopic fission barrier vanishes, but nuclei are still stabilized by shell effects, is shown in Figure 6.3 for three selected heavy and SHN.

The shell-correction energies determining the energy and deformation of the ground states are plotted for heavy and SHN in Figure 6.4a. Two equally deep minima are obtained, one at $Z = 108$ and $N = 162$ for deformed nuclei with quadrupole and hexadecupole deformation parameters $\beta_2 = 0.23$ and $\beta_4 = -0.07$,

Figure 6.4 Ground-state shell-correction energy (a) and partial half-lives for SF, α, and β decay (b–d). For easier comparison of the data given in (a) and (b), the scale in (a) is inverted. The filled squares in (d) mark the β-stable nuclei. (Source: The calculated values in (a–c) were taken from [5, 6] and those in (d) from [7].)


(a)
$-E_{shell}$(MeV)
3
4
5
6
7
8
9
(b)
lg(T_{SF}/s)
−12
−9
−6
−3
0
3
6
9
12
15
(c)
lg(T_{α}/s)
−9
−6
−3
0
3
6
9
12
(d)
lg(T_{β}/s)
0
1
2
>3
130
126
120
114
110
104
150
160
170
180
184
190
200
Proton number
Neutron number

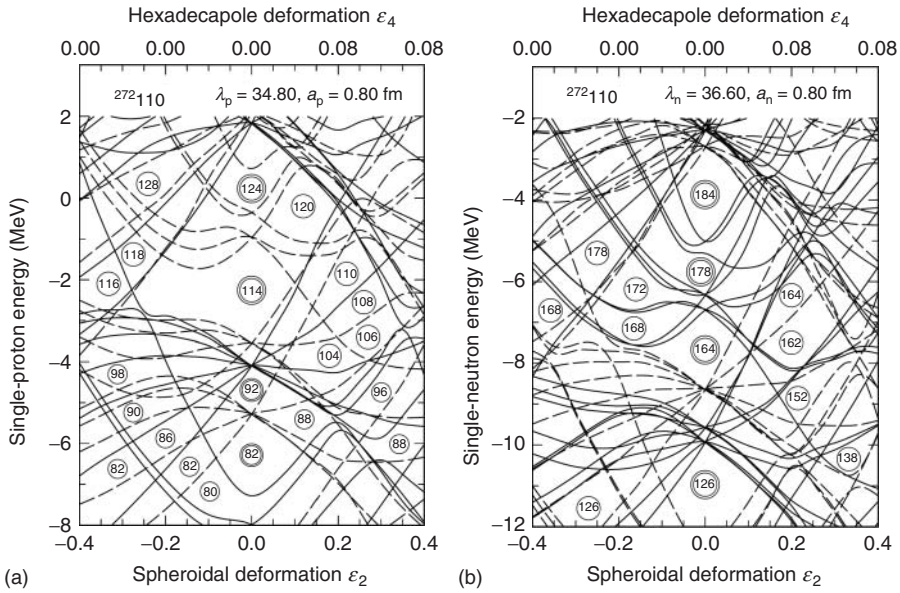


Figure 6.5 Calculated single-particle-level diagrams for spherical and deformed nuclei in the vicinity of $^{272}110_{162}$: (a) protons and (b) neutrons. The nucleus $^{272}110_{162}$ itself is determined to be deformed in its ground state. (Source: The figure was taken from [7].)

respectively, and the other one at $Z = 114$ and $N = 184$ for spherical SHN. Different results are obtained from self-consistent *Hartree–Fock–Bogoliubov* calculations [3] and self-consistent *mean-field* models [4]. They predict for spherical nuclei shells at $Z = 114$, 120, or 126 (indicated as dashed lines in Figure 6.4a) and $N = 172$ or 184. Nevertheless, the figure substantiates the *island of stability* of SHN being located at $Z = 114$–126 and $N = 184$ with a separating borderline from the lighter *heavy nuclei* at $A \approx 280$. This island of nuclei which are exceptionally stable in a sea of surrounding highly unstable nuclei, is the overall theme of SHN research.

The reason for the uncertainty of the correct proton number is given in terms of the shell model by the energy of the major proton shells $^1h_{9/2}$, $^1i_{13/2}$, and $^2f_{7/2}$ below $Z = 114$, and the subshells $^2f_{5/2}$, $^3p_{3/2}$, and $^3p_{1/2}$, which are filled between $Z = 114$ and 126 and $^1i_{11/2}$ and $^1j_{15/2}$ above $Z = 126$. Figure 6.5 shows the level diagrams for nuclei in the region of SHN. The spin–orbit splitting is the main factor that determines the energy of shells and subshells for spherical nuclei. However, at a large nuclear radius, the spin–orbit splitting is not exactly known. Furthermore, it depends on the diffuseness of the nuclear surface and density variations inside the nucleus. Therefore, the calculated energies are model dependent and subject to uncertainties.

Qualitatively, we may expect the minimum of the negative shell-correction energies to be wide and not as deep if the low spin proton orbits $^2f_{5/2}$, $^3p_{3/2}$, and $^3p_{1/2}$ are equally distributed in energy between $Z = 114$ and 126. The fission barriers will then also be flat and narrow, with their

height and width being mainly determined by the shell-correction energy. As a result, the fission half-lives will be relatively short. On the other hand, if there is a wide energy gap just beyond one of the proton numbers 114, 120, or 126, then the shell-correction energy will be pronounced for that element. In combination with the neutron shell effect at $N = 184$, a sharp and deep minimum will be formed, similar to that of the double magic ^{208}Pb, resulting in a high fission barrier and a relatively long fission half-life.

The minimum of the shell-correction energy (Figure 6.4a) is not exactly located at the closed shell numbers but shifted slightly above 114 and below 184. The reason is the low single-particle level density above $Z = 114$ and, reversely, below $N = 184$ (see Figure 6.5).

The low level density for deformed nuclei at $\varepsilon_2 \approx 0.25$ and $\varepsilon_4 \approx 0.08$ for $Z = 108$ and $N = 162$ (see Figure 6.5) is responsible for the second minimum of shell-correction energies and, thus, an increased stability of nuclei in this region (see Figure 6.4a). (The letters ε and β characterize two different parameterizations of the nuclear shape.)

Partial SF half-lives are plotted in Figure 6.4b. The landscape of fission half-lives reflects the landscape of shell-correction energies, because in the region of SHN, the height of the fission barrier is, first, mainly determined by the ground-state shell-correction energy, while the contribution from the macroscopic liquid-drop part approaches 0 for $Z = 104$ and above, and, second, the shell-correction energy at the saddle point is small. Nevertheless, we see a significant increase in SF half-lives from 10^3 s for deformed nuclei to 10^{15} s for spherical SHN. This difference originates mainly from an increasing width of the fission barrier, which becomes wider in the case of spherical SHN.

It is interesting to note that a change of 1 MeV of the ground-state shell-correction energy, which is about the uncertainty of the calculations, alters the partial SF half-life by three orders of magnitude. A similar strong influence of the shell-correction energy on the survival probability of a CN has to be expected, thus, resulting in large uncertainties of calculated fusion-evaporation cross sections.

Partial α half-lives decrease almost monotonically from 10^{12} s down to 10^{-9} s near $Z = 126$ (Figure 6.4c). The α half-lives will be more strongly modulated only in the case of large shell effects, resulting in high Q_α values above and low Q_α values below the magic number and, thus, reduced and increased half-lives, respectively.

The valley of β-stable nuclei passes through $Z = 114$ and $N = 184$. At a distance of about 20 neutrons away from the bottom of this valley, β half-lives of isotopes have dropped down to values of 1 s only (Figure 6.4d). However, already in the neighborhood of the stable nuclei, the half-lives are generally not longer than 100 s.

By combining the results from the individual decay modes, one obtains the dominating partial half-life, as shown in Figure 6.6a for even–even nuclei. The two regions of deformed heavy nuclei near $N = 162$ and spherical SHN merge and form a region of α emitters surrounded by spontaneously fissioning nuclei. The longest half-lives are 1000 s for deformed heavy nuclei and 30 years for spherical SHN. It is interesting to note that the longest half-lives are not reached for nuclei at the doubly magic nucleus ^{298}114, but for $Z = 110$ and $N = 182$. This is a result of continuously increasing Q_α values with increasing atomic numbers. Therefore, α decay becomes the dominant decay mode

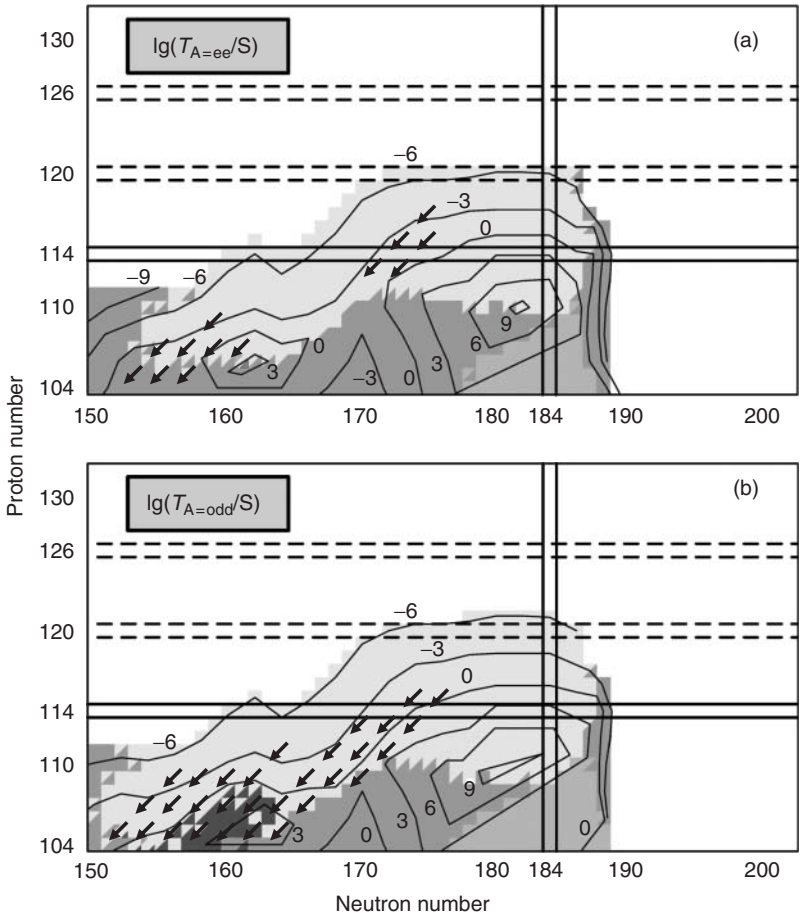


Figure 6.6 Dominating partial half-lives for α, β^+-decay/EC, β^--decay, and SF: (a) for even–even nuclei and (b) for odd-*A* nuclei. Nuclei and decay chains of even elements known at present are marked.

beyond darmstadtium, with continuously decreasing half-lives. For nuclei at $N = 184$ and $Z < 110$, half-lives are determined by β^- decay.

For odd nuclei (Figure 6.6b), partial α and SF half-lives have to be multiplied by a factor of 10 and 1000, respectively, thus making provisions for the odd particle hindrance factors. However, we have to keep in mind that fission hindrance factors show a wide distribution from 10^1 to 10^5; in some cases, even values up to 10^{10} were measured, which is mainly a result of the specific level occupied by the odd nucleon. During deformation, the energy of such levels rises and thus increases the fission barrier. The term *specialization energy* is used for this enlargement. For odd–odd nuclei, the fission hindrance factors from both the odd proton and the odd neutron are multiplied. For odd and odd–odd nuclei, the island

character of α emitters disappears and for nuclei with $N = 150–160$, α decay prevails down to rutherfordium and beyond, with the exception of a few nuclei around $Z = 105$ and $N = 159$, which decay by electron capture (EC). In the allegorical representation where the stability of SHN is seen as an island in a sea of instability, even–even nuclei portray the situation at high tide and odd nuclei at low tide, when the island is connected to the mainland.

The interesting question arises, if and to which extent uncertainties related to the location of proton and neutron shell closures will change the half-lives of SHN? Partial α and β half-lives are only insignificantly modified by shell effects because their decay process occurs between neighboring nuclei. This is different for fission half-lives that are primarily determined by shell effects. However, the uncertainty related to the location of nuclei with the strongest shell effects, and thus the longest partial SF half-life at $Z = 114$, 120, or 126 and at $N = 172$ or 184 is irrelevant concerning the longest *total* half-life of SHN. All regions for these SHN are dominated by α decay.

The line of reasoning is, however, different concerning the production cross section. The survival probability of the CN is determined among other factors significantly by the fission barrier. Therefore, with respect to an efficient production yield, the knowledge of the location of minimal negative shell-correction energy is highly important. However, if it turns out that shell effects in the region of SHN are distributed across the subshell closures at $Z = 114$, 120, and 126, then a wider region of less deep shell-correction energy would exist with corresponding modification to lower values of the production yield of SHN.

If the Fermi level in deformed nuclei is located within a band of close lying levels and levels of both high and low spins exist, then relatively low-lying high-spin *K isomers* can be formed. The regions of deformed nuclei, where K isomers can be expected, are located at $Z = 100$ and $N = 152$, and $Z = 108$ and $N = 162$. A detailed knowledge of the spectroscopic properties of these levels is particularly interesting, because levels that are relevant for the location of the shell closures in the region of spherical SHN could be involved.

K isomers are formed by the alignment of angular momentum of nucleons of one or more broken pairs. Therefore, it is argued that such nucleons could be the reason for a substantial specialization energy and, thus, reduced probability for fission of such isomers. An increased yield of evaporation residues (ERs) would result, which are populated via channels involving K isomers.

A pronounced *second minimum* in the fission barrier as shown for ^{250}Cf in Figure 6.3 exists also in nuclei at $N = 142–150$ of the elements from neptunium to berkelium. In these nuclei the outer fission barrier is high enough so that nuclei can exist at such large deformations, forming *shape isomers* or *fission isomers*. The half-lives are in the range from nanoseconds to milliseconds. The phenomenon was discovered by S. M. Polikanov in Dubna in 1962 [8]. Fission isomers are produced in relatively large quantities when actinide targets are used. They can give reason for misinterpretation when fission is assigned to heavy elements without having a possibility for correlation to α decays.

Fission properties are not only determined by shell effects in the parent nucleus, but also by shell effects in the fission fragments. The former modulate the fission barrier inside the saddle point

(which is determined by the maximum of the barrier), whereas the latter determine the fission path outside the saddle point. In the case of *asymmetric* SF of ^{252}Cf, the heavier fission fragment is near the double magic ^{132}Sn, and the second fragment with mass number A of about 120 is located in the region of deformed nuclei below tin. An elongated fission path occurs, and the fragments separate in a late stage. There, the repulsive Coulomb forces are relatively low, and the fission fragments obtain only low kinetic energy. However, in the case of SF of ^{258}Fm, the formation of both fragments is influenced by the shell effect at $Z = 50$ and $N = 82$. Two nearly symmetric spherical fragments are formed in a compact configuration after the saddle point. They separate with high kinetic energy owing to strong Coulomb repulsion. This *symmetric* SF occurs in a limited area around fermium–nobelium at $N = 158–160$ [9]. The ideal candidate would be the so far not yet known nucleus ^{264}Fm decaying into two double magic ^{132}Sn nuclei. It is interesting to note that nuclei in this region exist, where both fission paths occur, for example, in ^{259}Md [10]. The term *bimodal fission* is in use for this phenomenon.

6.3 Experimental Techniques

Isotopes up to fermium are produced by neutron capture of heavy nuclei and subsequent β^- decay. This method was and is used to produce bigger amounts of transuranium nuclei. Although from neptunium to californium, some isotopes can be produced in amounts of kilograms or at least in grams in high neutron flux reactors, the two heaviest species, ^{254}Es and ^{257}Fm, are available only in quantities of micrograms and picograms, respectively. At fermium, however, the method ends due to the lack of β^- decay and to the too short α and fission half-lives. Sufficiently thick targets both for neutron capture and heavy-ion fusion reactions cannot be manufactured from elements beyond fermium.

The region beyond fermium is best accessible using heavy-ion fusion reactions, the bombardment of heavy-element targets with heavy ions from an accelerator. The cross section is less than in the case of neutron capture, and values are considerably below the geometrical size of the nuclei. Moreover, only thin targets of the order of 0.5 mg cm^{-2} can be used. This limitation arises from the energy loss of the ion beam in the target, which results (using thicker targets) in energy distributions that are too wide for both the production of fusion products and their in-flight separation. On the other hand, the use of thin targets in combination with well-defined beam energies from accelerators results in unique information about the reaction mechanism. The data are obtained by measuring *excitation functions*, with the yield as a function of the beam energy.

Various combinations of projectiles and targets are in principle possible for the synthesis of heavy elements: actinide targets irradiated by light projectiles of elements in the range from neon to calcium, targets of lead and bismuth irradiated by projectiles from calcium to krypton, and symmetric combinations such as tin plus tin up to samarium plus samarium. In addition, inverse reactions using, for example, lead or uranium as projectile are possible and may have technical advantages in specific cases. Mutually, the choice of the reaction determines the design of the accelerators, separators, and detector systems.

6.3.1 Ion Source and Accelerator

Historically, the first accelerators used for the production of heavy elements were the cyclotrons at Berkeley, California, and later at Dubna, Russia. However, they were only able to accelerate light ions up to about neon, with sufficient intensity and up to an energy that is high enough for fusion reactions. Larger and more powerful cyclotrons were built in Dubna for the investigation of reactions using projectiles near calcium. Those cyclotrons include the U300 and U400 (300 and 400 cm diameter cyclotrons, respectively). In Berkeley, a linear accelerator Heavy Ion Linear ACcelerator (HILAC), later upgraded to the SuperHILAC, was built. The shutdown of this accelerator in 1992 led to a revival of heavy-element experiments at the 88 in. Cyclotron. Aiming at the acceleration of ions as heavy as uranium, the UNIversal Linear ACcelerator (UNILAC) was constructed in Darmstadt during 1969–1974.

In order to compensate for the decreasing cross sections of the synthesis of SHN, increasing beam currents are needed from the accelerators. Continuous efforts are aimed at increasing currents of high charge state from the ion source and to simultaneously reduce the consumption of material in the ion source, which usually consists of expensive, isotopically enriched probes such as ^{48}Ca. Further requests are high transmission of the ions through the accelerator, a high-quality beam of low emittance, halo free, and of well-defined energy. Finally, an overall high stability is crucial for all components from the ion source up to the data acquisition.

Many of these requests are achieved with modern Electron Cyclotron Resonance (ECR) ion sources. For example, the following beam intensities could be obtained at the UNILAC on the target: 1.6 pμA for $^{48}Ca^{10+}$; 4.0 pμA for $^{40}Ar^{8+}$; 1.2 pμA for $^{58}Fe^{8+}$; 0.4 pμA for $^{82}Se^{12+}$ (1 pμA = 6.24×10^{12} particles/s). The given values represent mean currents reached on target at a duty factor of 28% (5.5 ms wide pulses at 20 Hz repetition frequency). In the case of rare isotopic abundance, the source material was enriched to reach a concentration higher than 90%. The consumption of material was between 0.2 and 4 mg h^{-1}. The duty factor of 28% is due to power limitation of the 108-MHz radio-frequency tubes at the UNILAC. Using modern superconducting linear accelerators, a 100% duty factor could be reached, which then would increase the beam intensity by a factor of 3.6, and this, is achieved even with a lower consumption of electricity.

6.3.2 Targets

High beam currents, in turn, demand a high resistance of the targets. The present target technology uses target wheels that rotate with high speed through the beam. In general, the beam intensity is limited by the melting point of the target material. In some cases, a higher melting point can be achieved by using chemical compounds or alloys. Already successfully tested is a PbS target (melting point 1118 °C) produced by depositing the target material on a carbon backing. Other examples of high melting point compound targets are BiO_2 and UF_4. In the case of uranium targets, the elementary metallic uranium has a higher melting point (1132 °C compared to 960 °C of UF_4); however, the production of the target by sputtering of the uranium on a thin carbon foil is more complicated than the evaporation of UF_4.

Special safety precautions are needed in the case of highly radioactive actinide targets such as ^{244}Pu, ^{248}Cm, and others. High beam intensities and large beam doses make it necessary to use rotating target wheels also in these cases. Amounts of 1–10 mg of the radioactive material are mounted in several segments on the circumference of a wheel. The generally used production process involves molecular plating of the material from a solvent, which demands a relatively thick and stable backing foil (e.g., a titanium foil of 1 mg cm^{-2} thickness). The final chemical compound used is an actinide oxide having a high melting point.

A crucial item is also the intensity distribution of the beam across the target. Quadrupoles as ion-optical elements allow only for a Gaussian-shaped beam profile with the highest intensity in the center region and tails at the outer areas. The former most likely melts the target in the middle and the latter cause background when hitting the target frame. The intensity distribution can be improved using an octupole doublet in the beam line in addition to the quadrupoles. With these magnets, an almost rectangular intensity distribution is achievable.

The temperature distribution across the target is monitored by an infrared video camera. The monitor system can also be used as a control of the beam position and the beam intensity distribution during the irradiation. The target thickness is controlled online by registration of elastically scattered projectiles. A precise two-dimensional thickness measurement was developed by using a narrow beam of electrons of 20–30 keV energy. In this method the reduction of the electron beam intensity by scattering from the target material is measured.

6.3.3 Recoil-Separation Techniques

The identification of the first transuranium elements was done by chemical means. In the early 1960s, mechanical techniques that allowed for detection of nuclei with lifetimes down to a few tenths of seconds were developed. A substantial improvement of the physical methods was obtained with the development of recoil separators and the implantation of the nuclei into large area position-sensitive detectors.

Recoil-separation techniques use the ionic charge and momentum of the recoiling fusion product obtained in the reaction process. Spatial separation from the projectiles and other reaction products is achieved by magnetic fields or by a combination of electric and magnetic fields. The separation times are determined by the recoil velocities and the lengths of the separators. They are typically in the range of 1–2 μs. Two types of recoil separators have been developed:

1. The gas-filled separators use the different magnetic rigidities of the recoils, and the projectiles, traveling through a low pressure (about 1 mbar) gas-filled volume in a magnetic dipole field [11]. In general, helium or hydrogen, or a mixture of both gases is used in order to obtain a maximum difference in the rigidities of slow reaction products and fast projectiles. An enduring mean charge state of the ions is achieved by frequent collisions with the atoms of the gas.
2. Wien-filter or energy separators use the specific kinematic properties of the fusion products. The latter are created with velocities and energies different from the projectiles and other reaction products. Their ionic-charge state is

determined when they escape from a thin solid-state target into vacuum. A whole charge state distribution is created with a width of about ±10% around a certain mean value. Therefore, ionic-charge achromaticity is essential for high transmission. It is achieved by additional magnetic fields or symmetric arrangements of electric fields. An example of such a separator is the velocity filter Separator for Heavy-Ion reaction Products (SHIP) used in experiments for the investigation of heavy elements at GSI [12]. A scheme of SHIP is shown in Figure 6.7. The principle of separation and detection techniques used in most of the other laboratories is comparable.

6.3.4 Detectors, Signal Processing, and Data Analysis

Recoil separators are designed to filter out those nuclei with a high transmission, which are produced in fusion reactions. Since higher overall yields result in increased background levels, the transmitted particles have to be further identified by detector systems. The detector type to be selected depends on the particle rate, energy, decay mode, and half-life. Experimental as well as theoretical data on the stability of heavy nuclei show that they decay by α emission, EC, or SF, with half-lives ranging from microseconds to days. Therefore, silicon semiconductor detectors are well suited for the identification of nuclei and for the measurement of their decay properties. If the total rate of ions striking the focal plane of the separator is low, then the particles can be implanted directly into the silicon detectors. Using position-sensitive detectors, one can measure the local distribution of the implanted particles [13, 14]. In this case, the detectors act as diagnostic elements to optimize and control the ion-optical properties of the separator.

Given that the implanted nuclei are radioactive, the positions measured for the implantation and all subsequent decay processes are the same. This is the case because the recoil effects are small compared with the range of implanted nuclei, emitted α particles or fission products, and detector resolution. Event-by-event recording of the data allows for the analysis of delayed coincidences with variable position, and time windows allows for the identification of decay chains.

At SHIP, the presently used detector system is composed of three time-of-flight detectors, eight identical 16-strip silicon wafers, and germanium detectors. A schematic view of the detector arrangement in the focal plane of SHIP is shown in Figure 6.7. The three secondary-electron foil detectors in front of the silicon detectors are used for measuring the particle velocity and for tagging ions that are implanted into the stop detector. Three detectors are used to increase the detection efficiency.

A time-of-flight signal and an energy signal from the silicon detector provide the information for switching off the beam after detection of an implanted residue. After a response time of 20 μs a subsequent time window of preset duration opens for counting a preset number of α particles of the decay chain. If the desired conditions are fulfilled, the beam-off period is prolonged up to the expected measurable end of the decay chain by opening a third time window. This improvement considerably reduces the background during the measuring period of the decay chain and allows for the

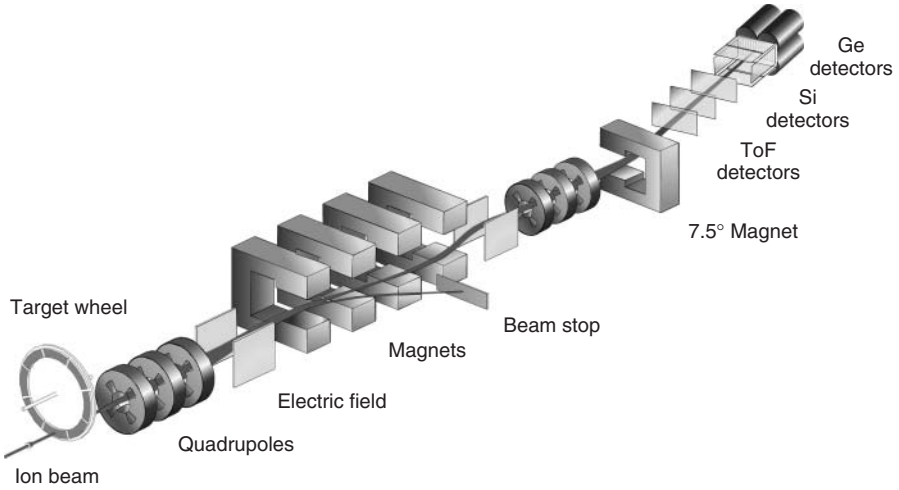


Figure 6.7 The velocity filter SHIP (separator for heavy-ion reaction products) and its target and detection system. The drawing is approximately to scale; however, the target wheel and the detectors are enlarged by a factor of 2. The length of SHIP from the target to the detector is 11 m. The target wheel has a radius up to the center of the targets of 155 mm. It rotates synchronously with the beam macrostructure at 1125 rpm. The flight time of the reaction products through SHIP is 1–2 μs. The filter, consisting of two electric and four magnetic dipole fields plus two quadrupole triplets, was extended by a fifth deflection magnet, allowing for positioning of the detectors away from the straight beam line and further reduction of the background. The detector setup is composed of three time-of-flight detectors, an arrangement of eight 16-strip silicon wafers, and a germanium clover detector. The reaction products are stopped in the main silicon detector (stop detector). Six detectors are mounted upstream for detection of α particles or fission fragments emitted in backward direction. The eighth detector is mounted behind the stop detector. It provides a veto signal for suppression of signals from light, high-energy particles (protons or α particles) emitted from the target, which traverse the stop detector.

safe detection of signals from long-lived decays.

The dead time in systems with conventional pulse-shaping electronics, analog-to-digital converters (ADCs), and data collection is on the order of 10–100 and 20 μs at SHIP. Such relatively long dead times are difficult to accept in experiments, where the nuclei are separated within 1 μs, and some of the heaviest SHN are predicted to decay within microseconds. Therefore, 100 or 250 MHz flash ADCs are installed to overcome this limitation. A further advantage is that the cleanness of the detector signals is assessed and, thus, misinterpretations from pile-up or electronic disturbances are avoided.

The germanium detectors measure X-rays or γ-rays that are coincident with signals from the silicon detectors. This allows for the detection of α transitions to excited levels in the daughter nucleus, which decay by γ emission. In the case of EC, characteristic X-rays may be emitted, which would allow for a clear element identification. Although the probability for detecting coincident events is small, the germanium detectors provide useful spectroscopic information if the cross sections are of the order of nanobarns or

higher. The situation is different in the case of SF owing to the high γ multiplicity. In about 80% of all SF events, γ-rays are measured in coincidence and, thus, help us to distinguish SF from projectile background.

A further extension of the measuring possibilities is achieved with γ-rays, X-rays, or particle (p, n, α, electron) detectors mounted around the target. If these detectors are operated in delayed coincidence with signals from the implantation of reaction products and their radioactive decay in the focal plane of the separator, the sensitivity of *in-beam* spectroscopy is significantly improved. This so-called Recoil-Decay Tagging (RDT) method was first applied in a study of the heavy-ion radiative capture mechanism, a cold fusion process without any evaporation of nucleons, but emission of only γ-rays for cooling of the CN. The reaction studied was $^{90}Zr + {}^{90}Zr \rightarrow {}^{180}Hg$ [15]. Meanwhile the method has become a standard tool in nuclear in-beam spectroscopy.

6.4 Experiments on SHN

Historically, elements beyond uranium up to mendelevium ($Z = 101$) were searched for by chemical means of new elements produced after neutron capture and subsequent β^- decay or irradiation of targets from isotopes of uranium up to einsteinium with light projectiles of ^{2}H and ^{4}He. Only one element, americium (95), was discovered after neutron irradiation in a nuclear reactor. Two elements, einsteinium (^{253}Es) and fermium (^{255}Fm), were found in the debris of the first thermonuclear explosion launched in 1952. The intensive study of the chemistry of these elements led Seaborg to propose the actinide series [16] (filling of the 5f electron shell) beginning with the element actinium ($Z = 89$) in analogy to the lanthanide series (filling of the 4f shell) beginning with lanthanum ($Z = 57$) (see Figure 6.8).

The elements from nobelium (102) to seaborgium (106) were synthesized in fusion reactions of $^{244,246,248}Cm$, ^{249}Bk, and $^{249-252}Cf$ with beams of ^{10}B, $^{12,13}C$, ^{15}N, and $^{16,18}O$ at Berkeley. Targets of ^{238}U, $^{241,242}Pu$, and ^{243}Am were irradiated with beams of ^{18}O and ^{22}Ne for synthesis of elements from nobelium to dubnium at Dubna. In the discovery experiments of these elements, relatively slow mechanical transport systems such as tapes or wheels were used. However, the transport times were fast enough for the detection of the produced nuclei having half-lives on the order of seconds.

The prediction of high stability of SHN and the prospects for producing them with large cross sections in the 1960s generated new experimental efforts in the 1970s and 1980s. Using beams from the new accelerators constructed at that time, a variety of nuclear reactions were studied. These were actinide-based reactions such as $^{65}Cu + {}^{238}U \rightarrow {}^{303}121^*$ or even irradiations of ^{254}Es with ^{48}Ca and other light ions, reactions using ^{208}Pb and ^{209}Bi targets, symmetric reactions such as $^{136}Xe + {}^{170}Er \rightarrow {}^{306}122^*$, and also deep inelastic reactions using the heaviest projectiles and targets available, such as $^{238}U + {}^{248}Cm$. Although SHN were not observed – deduced cross-sectional limits were on the order of tens or hundreds of picobarns depending on expected lifetimes – the negative results revealed that systematic strategies are needed.

Two classes of reactions turned out to be promising: (i) reactions using ^{208}Pb and ^{209}Bi targets and (ii) reactions using actinide targets. The names *cold* and *hot*

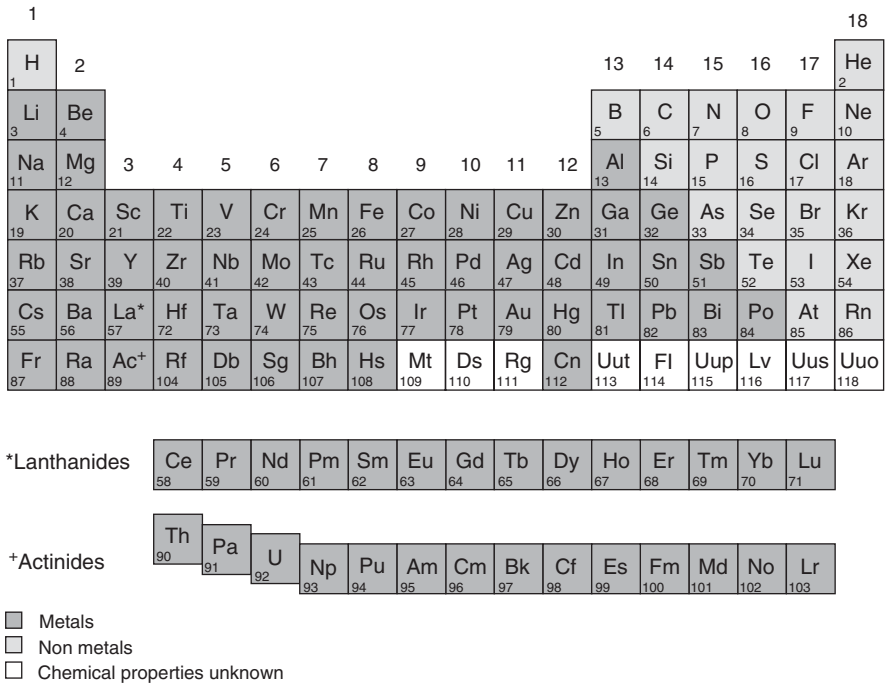


Figure 6.8 Periodic table of the elements. The known transactinide elements 104–118 take the positions from below Hf in group 4 to below Rn in group 18. The elements hassium (Hs) and copernicium (Cn), the heaviest elements chemically investigated, are placed in groups 8 and 12, respectively. The arrangement of the actinides reflects the fact that the first actinide elements still resemble, to a decreasing extent, the chemistry of the other groups: Th (group 4 below Hf), Pa (group 5 below Ta), and U (group 6 below W). All known elements are given with their accepted name or their preliminary name suggested by IUPAC (International Union of Pure and Applied Chemistry).

fusion were coined for these two types of reactions. The reason for this denotation is the different excitation energy of the CN after fusion. As same as chemical reactions, the nuclear reactions can be exothermic or endothermic. Owing to the strong binding energy of the double magic lead nucleus and projectiles of the double magic ^{48}Ca, the excitation energy is rather low (near 10–20 MeV). In this case, only one neutron and γ-rays have to be emitted to cool the nucleus down to the ground state (see Figure 6.2a). For heavier projectiles the excitation energy is further decreasing. For reactions with ^{76}Ge, it is equal to the one-neutron binding energy, and with ^{86}Kr, the reaction becomes even endothermic. In this case, the projectile energy has to be increased so that the reaction becomes energetically possible and that at least one neutron can be emitted. The relations are plotted for stable and radioactive projectile isotopes in Figure 6.9. In general, the most neutron-rich stable isotopes as projectiles result in the lowest excitation energies.

The curves of excitation energies are located considerably higher for reactions with ^{238}U targets. Reactions with other actinide targets show a similar behavior.

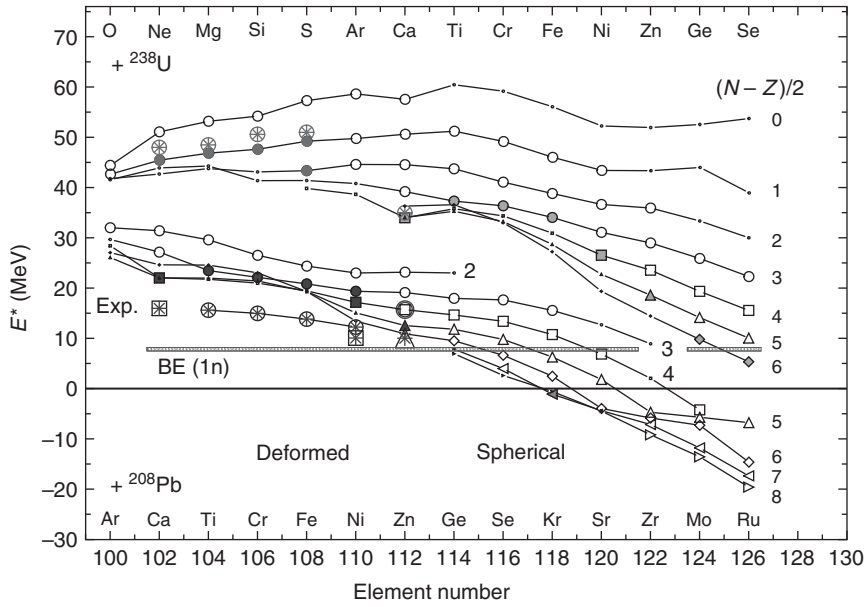


Figure 6.9 Excitation energy of CN produced in fusion reactions at beam energies just high enough for reaching a contact configuration. The upper part shows the trend for hot fusion reactions with ^{238}U targets and projectiles between oxygen and selenium, the lower part shows the trend for cold fusion reactions with ^{208}Pb targets and projectiles from argon to ruthenium. The lines connect projectiles with the same isospin $T_z = (N - Z)/2$. The open symbols mark stable projectile isotopes, and the little symbols radioactive isotopes. The filled symbols of $Z \geq 114$ stable projectile isotopes in the upper part are brought out, because they mark impressively the transition from hot to cold fusion with increasing element number using actinide targets. The other filled symbols up to element 112 mark reactions investigated by cold fusion or by hot fusion. Data for which maximum cross sections were measured (marked by "Exp.") are even below that energy, halfway to the limiting (for the 1n channel) one-neutron binding energy in the case of cold fusion, and slightly above the estimated excitation energies in the case of hot fusion. The 1n-binding energies are in a range from 7.5 to 8.2 MeV for the investigated cold fusion compound nuclei and are marked by the horizontal bar. The excitation energies were calculated using mass values given in [17] for projectile and target and in [18] for the CN.

The hot fusion reactions up to the synthesis of element 110 have excitation energies of about 40 MeV. Remarkable is the relatively low excitation energy for the reaction using a ^{48}Ca beam and a ^{238}U target. Also with the uranium target the reactions get colder with increasing element number. The excitation energy reaches values close to the 1n-binding energy near element 124 using a ^{76}Ge beam. This means that for the heaviest elements, the hot fusion reactions gradually change into cold fusion. This could be a promising aspect for the synthesis of SHN near $Z = 126$ (which is one of the predicted closed proton shells) using actinide targets.

The amount of excitation energy is crucial for the surviving probability of the CN. The emission of only one neutron for reaching energy levels near the ground state needs less time than the emission of three or four neutrons, which results

in a lower probability for fission in the first case. A disadvantage of cold fusion is, however, the stronger electric repulsion between projectile and target, which results in a higher probability for reseparation before fusion.

6.4.1 Nuclei Produced in Cold Fusion Reactions

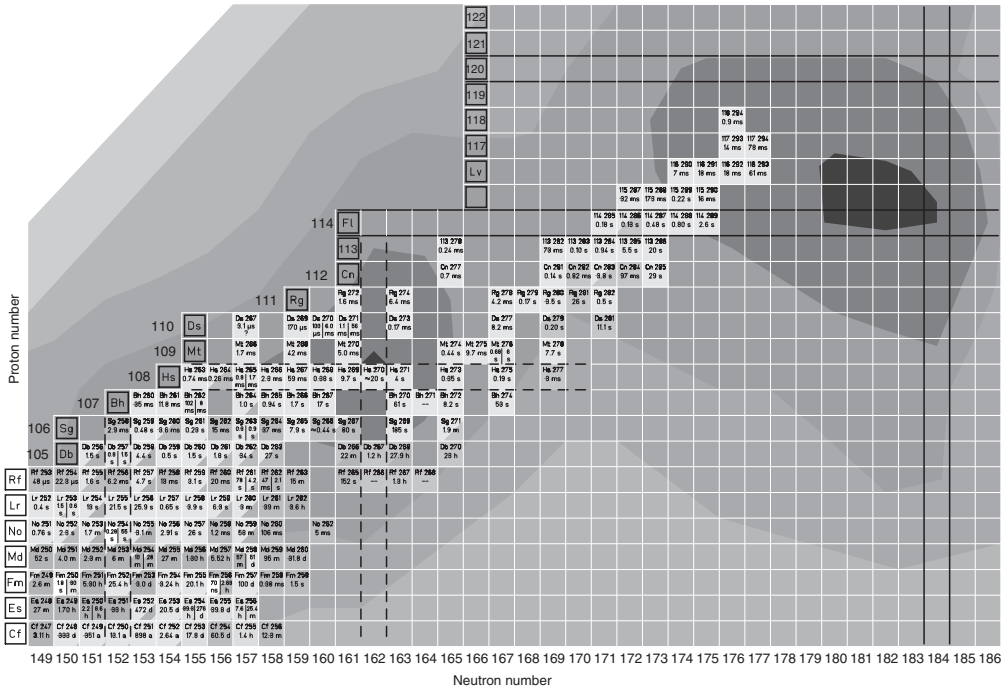
The advantage of cold fusion for the synthesis of heavy nuclei was first realized by Oganessian, and the first successful experiments using ^{208}Pb targets were performed in Dubna in 1973 [19]. In addition, technical considerations were decisive for choosing cold fusion as start-up experiments at SHIP. Targets of stable lead and bismuth isotopes are cheap and easy to handle. Furthermore, in reactions with beams of ^{40}Ar or isotopes of heavier elements that all were available from the UNILAC, the velocities of ERs are much higher than in the case of hot fusion using heavier targets and lighter projectiles. At higher velocities and evaporation of only one neutron, the scattering and angular straggling are much less than in the case of lower velocities and evaporation of three to four neutrons. Consequently, the number of ERs being accepted by a recoil separator is considerably higher in the case of cold fusion. At SHIP, the transmission reaches values of 30–50%.

In first experiments at SHIP, the cold fusion reaction $^{40}\mathrm{Ar} + {}^{208}\mathrm{Pb} \rightarrow {}^{246}\mathrm{Fm} + 2\mathrm{n}$ was studied and previous results obtained at Dubna were confirmed. The search for element 107 was prepared in a study of the expected daughter nucleus produced in the reaction $^{50}\mathrm{Ti} + {}^{209}\mathrm{Bi} \rightarrow {}^{258}105 + 1\mathrm{n}$. The isotope ^{258}105 and its daughter ^{254}Lr were not known at that time. In the main experiment (performed in 1981), the reaction $^{54}\mathrm{Cr} + {}^{209}\mathrm{Bi}$ was used and six decay chains were measured, three at each of the two different beam energies used. The chains were in agreement with the decay of ^{262}107 produced in a 1n channel. Detailed references for this experiment and the results obtained in other studies at SHIP can be found in the review articles [14, 20, 21].

In continuation of the experiments the reaction $^{58}\mathrm{Fe} + {}^{209}\mathrm{Bi}$ was studied in 1982. Only a single decay chain was measured during a beam time of 2 weeks, which, however, was fully in agreement with the decay of the previously studied daughter ^{262}107.

The irradiation of ^{208}Pb with ^{58}Fe was performed in 1984. The study resulted in three α-decay chains with a half-life of 1.8 ms of ^{265}108 and a high upper limit

Figure 6.10 Upper end of the chart of nuclei showing the presently (2012) known nuclei. For each known isotope, the element name, mass number, and half-life are given. The relatively neutron-deficient isotopes of the elements up to proton number 113 were created after evaporation of one neutron from the compound nucleus (cold fusion reactions based on ^{208}Pb and ^{209}Bi targets). The more neutron-rich isotopes from element 112 to 118 were produced in reactions using a ^{48}Ca beam and targets of ^{238}U, ^{237}Np, ^{242}Pu, ^{244}Pu, ^{243}Am, ^{245}Cm, ^{248}Cm, ^{249}Bk, and ^{249}Cf. The magic numbers for the protons at element 114 and 120 are emphasized. The bold dashed lines mark proton number 108 and neutron numbers 152 and 162. Nuclei with that number of protons or neutrons have an increased stability; however, they are deformed contrary to the spherical SHN. The crossing at $Z = 114$ and $N = 162$ reflects the uncertainty, whether nuclei in that region are deformed or spherical. The background structure shows the calculated ground-state shell-correction energy according to the macroscopic–microscopic model.


Proton number
Neutron number
149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186
122
121
120
119
118
117
Lv
114 Fl
113
112 Cn
111 Rg
110 Ds
109 Mt
108 Hs
107 Bh
106 Sg
105 Db
104 Rf
103 Lr
102 No
101 Md
100 Fm
99 Es
98 Cf

of the partial SF half-life. Also in this case the main experiment was preceded by a study of the daughter $^{261}106$ in the reaction $^{54}Cr + ^{208}Pb$.

In 1992, the names bohrium (Bh), hassium (Hs), and meitnerium (Mt) of the new elements 107, 108, and 109, respectively, were officially approved by IUPAC. An overview of nuclei in the region of SHN, which are presently known, is given in the partial chart of nuclides (Figure 6.10). The figure includes all isotopes that were produced in cold and hot fusion reactions so far. Cross sections measured for the synthesis of elements 107–109 are shown in Figure 6.11a,b.

A search for element 110 was performed at SHIP using the cold fusion reaction $^{64}Ni + ^{208}Pb$ in 1985 and using the hot fusion reaction $^{40}Ar + ^{235}U$ in 1986. Both attempts were negative with upper cross-sectional limits of 5 and 8 pb, respectively. The limits are calculated with the assumption that one chain would have been

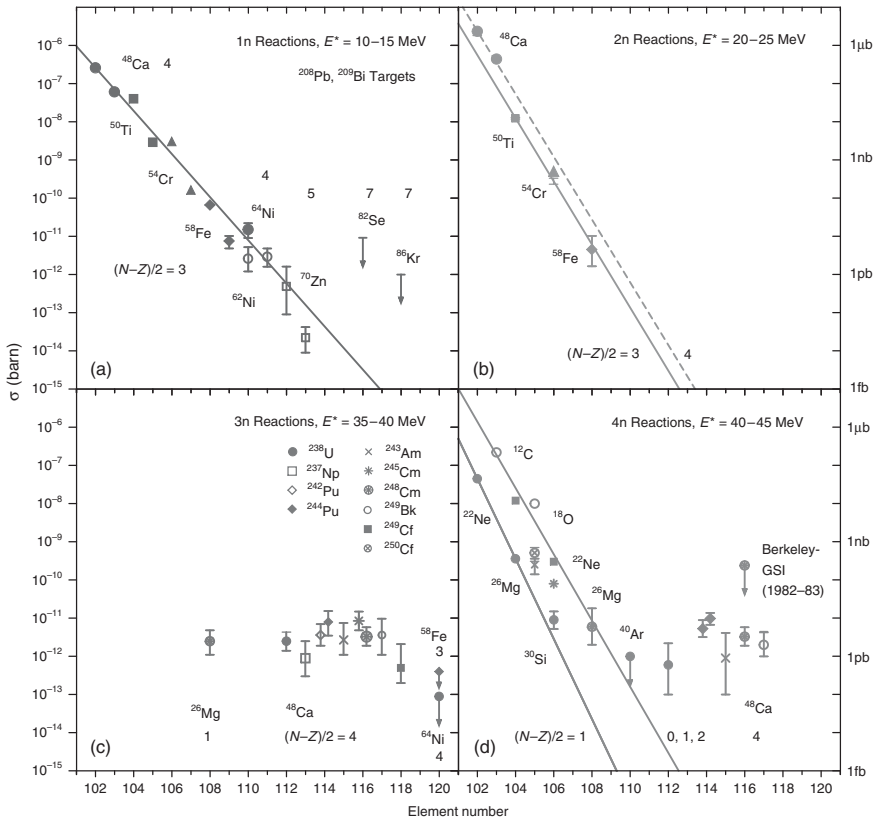


Figure 6.11 Measured cross sections for fusion reactions with ^{208}Pb and ^{209}Bi targets and evaporation of one neutron (a) and two neutrons (b) and for fusion reactions with actinide targets and evaporation of three (c) and four neutrons (d). Cross-section limits are marked by arrows. Values $(N - Z)/2$ characterize the neutron excess of the projectile. Straight lines are fits to data points.

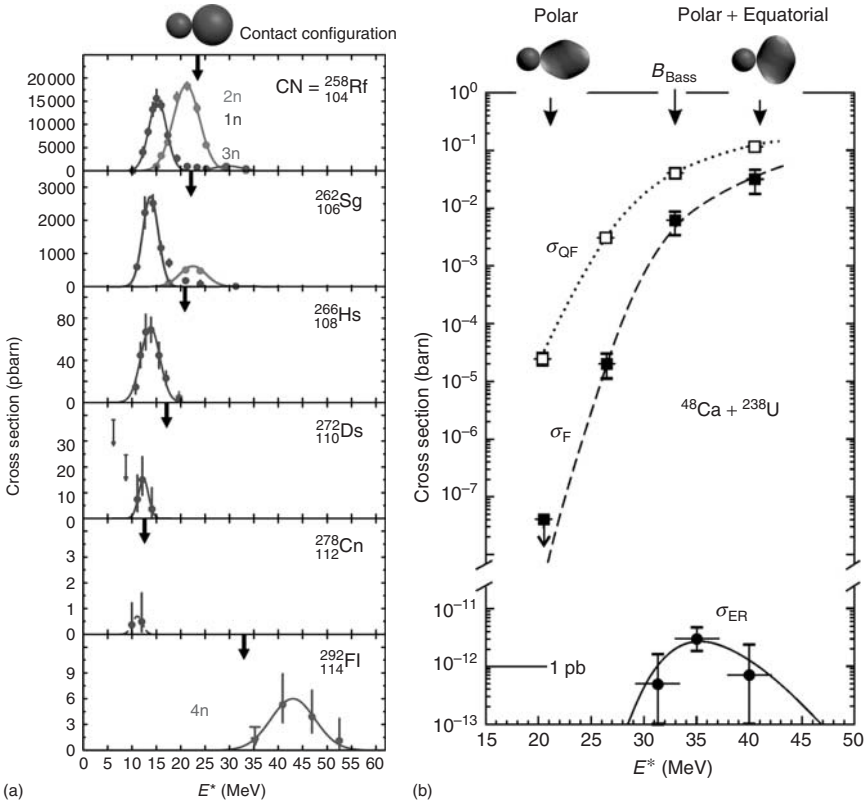


Figure 6.12 (a) Measured excitation functions of even elements from rutherfordium to copernicium produced in reactions with ^{208}Pb targets and beams from ^{50}Ti to ^{70}Zn. These data were measured in experiments at SHIP. For comparison, the excitation function for synthesis of flerovium ($Z = 114$) in the reaction ^{48}Ca + ^{244}Pu is plotted in the bottom panel. The arrows mark the energy at reaching a contact configuration using the model by Bass [22]. (b) Comparison of the cross sections as function of the excitation energy E^* for quasifission (σ_{QF}), compound fission (σ_F), and ERs (σ_{ER}) for the reaction ^{48}Ca + ^{238}U. (Source: Part (b) was taken from [23].)

observed (one event limit). In this case, statistical fluctuations are not considered.

Before starting another attempt to synthesize element 110, the experimental techniques were improved during 1987–1994. A factor of 10 in sensitivity was reached. A main factor of 2.2 was due to the new ECR ion source delivering higher beam intensities at a high charge state, which avoided stripping and intensity losses at an intermediate energy of $1.4 \times A$ MeV. The beam macropulse was extended from 4.5 to 5.5 ms, the SHIP transmission was increased by a factor of 1.5 by positioning the target closer to the first quadrupoles and, finally, the detector efficiency was increased by a factor of 1.4 using the newly available 16-strip position-sensitive detectors consisting of large area silicon wafers. The background suppression was improved by installing a fifth dipole magnet so that the detectors can be placed out of the straight beam line (see Figure 6.7).

The choice of the beam energy was still an open problem discussed very controversially. To overcome this uncertainty, the excitation functions for synthesis of at least two elements had to be measured. The choice was element 104 and element 108. The first case should be easy to measure owing to a cross section of about 10 000 pb expected for the maximum. Element 108 is the closest even element to 110, and knowledge of its excitation function would have the largest weight. However, the measured cross section was only 19 pb, and long enough beam times had to be requested.

The two measurements of the excitation functions were performed in June and October 1994. The results are included in Figure 6.12. The data revealed that the energy in the previous element 108 experiment was chosen too high. The new maximum value was 63 pb at $E^* = 13.4$ MeV. The need for an extra-push energy as predicted in several theoretical models, was excluded by the new measurement.

In the case of element 104, the 1n cross-sectional maximum was 10 000 pb at $E^* = 15.6$ MeV. An extrapolation of the energies of the cross-sectional maxima resulted in $E^* = 12.3$ MeV as optimum value for the synthesis of element 110. At this energy, the first decay chain of $^{269}110$ was measured in the reaction $^{62}Ni + ^{208}Pb$ in November 1994, already 1 day after the start of the irradiation. The measured cross section was 2.6 pb. This value is based on three measured decay chains. The second chain of the four ones published originally was retracted later.

Changing the beam from ^{62}Ni to ^{64}Ni increased the cross section. Nine decay chains of the isotope $^{271}110$ were measured compared to a total of three decay chains of $^{269}110$. A maximum cross section of 15 pb was measured at $E^* = 12.1$ MeV. This excitation energy is 6 MeV below the value used in the 1986 experiment.

The experiment with the ^{64}Ni beam was continued with irradiation of a ^{209}Bi target aiming to synthesize element 111 in December 1994. Three decay chains of $^{272}111$ were measured and a cross section of 3.5 pb was determined at $E^* = 12.5$ MeV.

Synthesis of element 112 was studied in 1996. In the reaction $^{70}Zn + ^{208}Pb \rightarrow ^{277}112 + 1n$ one chain was observed consisting of six consecutive α decays after implantation of an ER. The chain is shown in the left-hand side of Figure 6.13. A cross section of 0.5 pb was measured at $E^* = 10.1$ MeV. Also in this case, one of the originally published two decay chains was retracted later.

The data on the new elements 110–112, a total of 32 decay chains were measured at SHIP (not counted are the two retracted chains), was confirmed in follow-up experiments at SHIP itself and independently at RIKEN (Rikagaku Kenkyūjo, in English: The Institute of Physical and Chemical Research) near Tokyo in Japan [25–27] and at Lawrence Berkeley National Laboratory (LBNL) in California [28]. The names darmstadtium (Ds), roentgenium (Rg), and copernicium (Cn) were officially approved by IUPAC.

Cold fusion was also applied to search for elements 116 and 118. In both cases ^{208}Pb targets were irradiated. The beams were ^{82}Se and ^{86}Kr, respectively. The 116 experiment was performed at SHIP with the aim to search for *radiative capture*, the 0n channel. At five different excitation energies between 0 and 11 MeV, cross-sectional limits of about 5 pb were reached.

Subsequent to reports on a synthesis of element 118 at LBNL in 1999 [29], several laboratories tried to confirm these data, however, with negative result. Cross-sectional limits were reached of

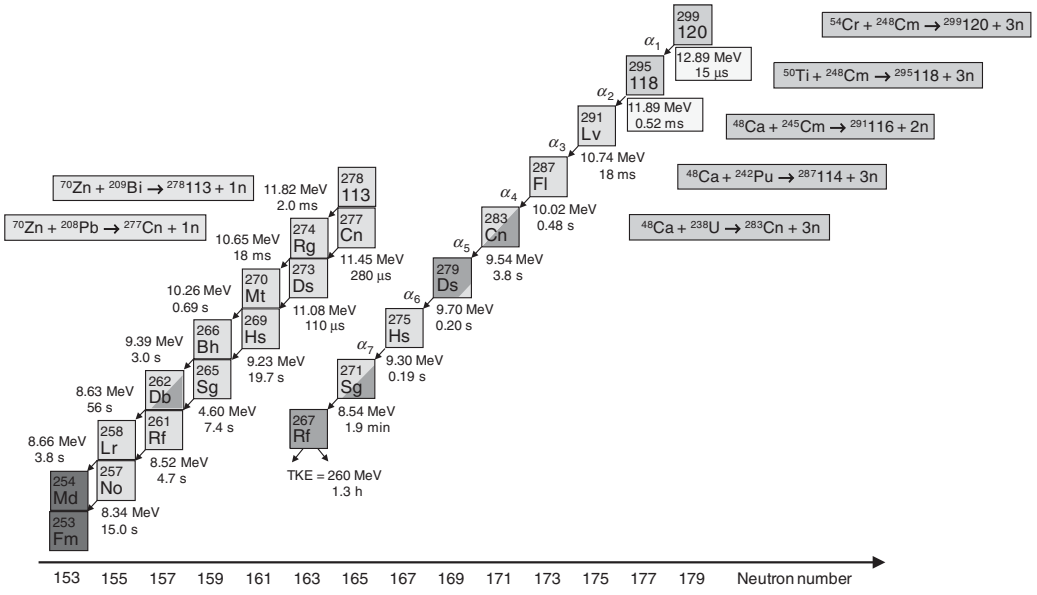


Figure 6.13 Decay chains measured in the cold fusion reactions $^{70}Zn + ^{208}Pb \rightarrow ^{277}112 + 1n$ and $^{70}Zn + ^{209}Bi \rightarrow ^{278}113 + 1n$ (in the left) and decay chains measured in hot fusion reactions in irradiations of ^{238}U, ^{242}Pu, and ^{245}Cm targets with ^{48}Ca ions (in the right). Alpha decay of the isotope $^{299}120$ of the not yet known element 120 produced in the reaction $^{54}Cr + ^{248}Cm$ would populate the new isotope $^{295}118$ of element 118 and subsequently the known decay chain of ^{291}Lv. Experimental decay data are given for the known nuclei; theoretical predictions in the case of $^{295}118$ and $^{299}120$ are given in [24].

about 1 pb, which was below the published value of 2.2 pb. Eventually, the first announcement was retracted in 2001 after additional experiments at LBNL and after a reanalysis of the data of the first experiment [30].

So far, 113 is the heaviest element produced by a cold fusion reaction. The reaction studied was $^{70}Zn + ^{209}Bi \rightarrow ^{279}113^*$. A first attempt to synthesize element 113 was made at SHIP in 1998 and continued in 2003. The net irradiation time was 82 days, a total beam dose of 1.5×10^{19} was reached, and an upper one event cross-sectional limit of 160 fb was deduced.

In the search for element 113 at RIKEN, the same reaction, $^{70}Zn + ^{209}Bi$, was used.

The irradiations started in 2003. With several intermissions, they were continued up to 2012. Three α-decay chains were measured, which were assigned to the isotope $^{278}113$ [31]. The first two chains, measured in 2004 and 2005, terminated after four subsequent α's by SF of the known isotope ^{262}Db. Also known was the last α emitter ^{266}Bh. New were the chain members ^{270}Mt and ^{274}Rg. The third chain, measured in 2012, terminated after six α decays at ^{254}Md. In this chain, the known α decays of ^{262}Db ($b_\alpha = 0.64$) and ^{258}Lr were observed, too. The chain is shown in Figure 6.13. The net irradiation time of the RIKEN element-113 experiment was 553 days, and the beam dose was 13.5×10^{19}. The cross section determined from all parts of the experiment is 22^{+20}_{-13} fb. This is the lowest value ever measured for a heavy-ion fusion cross section.

6.4.2 Nuclei Produced in Hot Fusion Reactions

Three advantages could be expected from the use of actinide targets, which, however, had to be proved experimentally: first, a smaller probability for reseparation in the entrance channel due to the higher asymmetry and thus less Coulomb repulsion between the reaction partners; second, a finite survival probability of the CN due to increasing fission barriers when the region of SHN is approached; and, thirdly, longer half-lives of the produced isotopes compared to the more neutron-deficient nuclei produced in cold fusion.

Exploratory studies were performed at the U400 cyclotron of Flerov Laboratory of Nuclear Reactions (FLNR) in Dubna during 1998–2003. The experiments were conducted at two different recoil separators, which had been built during the 1980s: The energy-dispersive electrostatic separator VASSILISSA and the Dubna gas-filled recoil separator (DGFRS). A specific characteristic of the DGFRS is the hydrogen gas used in the separator, which enables better suppression of projectile and target-like recoils at the focal plane than the filling with helium gas. Both separators were equipped with time-of-flight detectors and with an array of position-sensitive Si detectors in an arrangement similar to the one shown in Figure 6.7.

Targets of ^{238}U, ^{242}Pu, ^{244}Pu, ^{248}Cm, ^{245}Cm, and ^{249}Cf were irradiated with ^{48}Ca ions for the synthesis of the elements 112, 114, 116, and 118. The double magic isotope ^{48}Ca was used as projectile, because of its high number of neutrons, and because the formed CN have exceptional low excitation energy (see Figure 6.9). The observed isotopes and decay chains were assigned to neutron-evaporation channels of the produced CN. The duration of the beam times was several weeks up to about 4 months.

Eventually, in 2003, excitation functions were measured for the reactions with

^{244}Pu and ^{242}Pu. These results allowed for a correction of the previously made xn assignments, and a consistent interpretation was given of all so far observed decay chains from elements 114, 116, and 118. One of the measured decay chains, which were populated in three different reactions, is shown in the right-hand side of Figure 6.13. The nuclei produced in hot fusion are shown in Figure 6.10. An overview on the results obtained at FLNR up to 2007 is given in the review article [32].

An irradiation of ^{243}Am for the synthesis of the odd element 115 was performed at Dubna in 2003. In the same year, a search for element 113 was started at RIKEN using the cold fusion reaction ^{70}Zn + ^{209}Bi. Already in the first part from July to August, the Dubna experiment was successful.

The experiment was repeated in 2010–2011 [33]. At a total of four different beam energies, ER's of the 2n (1 chain), 3n (24 chains), and 4n (1 chain) channels were observed. A cross-sectional maximum of 8.5 pb was measured for the 3n channel. The decay chains of the produced odd–odd nucleus ^{288}115 ended by SF after α decay of ^{284}113, ^{280}Rg, ^{276}Mt, ^{272}Bh was detected.

The interesting question arises, which actually could be the fissioning nucleus at the end of these chains (see Figure 6.10). Taking into account theoretical and systematic values for Q_α and Q_{EC} energies and resulting half-lives as well as calculated SF half-lives and odd particle hindrance factors, it seems most likely that the odd–odd isotope ^{268}Db populated by the measured α decay of ^{272}Bh decays by EC with the measured half-life of 28 h, to ^{268}Rf, which decays by SF with a short half-life of ≈1 s.

The long half-life of ^{268}Db (16 h were measured in the 2003 experiment) provided an opportunity for chemical separation of this isotope. The experiment took place in 2004. Products from the ^{48}Ca + ^{243}Am reaction were implanted into a copper plate. The implanted isotopes were subjected to a chemical separation, which was specific for separation of groups 4 and 5 elements.

A total of 15 SF events were detected. A half-life of 32 h was measured in agreement within statistical uncertainties with the previous data obtained at DGFRS. The result was in agreement with the assignment of the parent isotope to the odd–odd nucleus ^{288}115.

Up to 2009, a number of reactions were performed aiming at consolidating the results on the exploration of the island of SHN. Among those was a direct production of element 113 isotopes by irradiation of a ^{237}Np target.

Special efforts were needed for the synthesis of element 117 with beams of ^{48}Ca, which were related to the production of a ^{249}Bk target. This nucleus has a half-life of only 320 days and has to be specifically produced in a high flux nuclear reactor. In 2009, such a target was provided for the Dubna group by the collaborators of Oak Ridge National Laboratory (ORNL) in Tennessee. During 140 days of irradiation, five decay chains were measured consisting of an implanted ER, three α decays, and then SF; and one chain consisting of an implanted ER, six α decays, and then SF. The chains were assigned to the isotopes ^{293}117 (4n channel) and ^{294}117 (3n), respectively [34]. The assignment to ^{293}117 was corroborated by the repetition of the element-115 experiment in 2010–2011, where the decay of the nucleus ^{289}115 produced in the 2n channel was fully in agreement with that of the daughter nucleus of ^{293}117 [33]. The 2009 data on element 117 were confirmed in a

second experiment performed in 2012. Five more decay chains of $^{293}117$ and two more of $^{294}117$ were measured [35]. An eighth decay chain was assigned to $^{294}118$ produced in a reaction with ^{249}Cf. This isotope is the β^- decay daughter of ^{249}Bk increasingly contaminating the target during long irradiation periods.

The nucleus with the longest half-life of 22 h at the end of the $^{294}117$ chain is ^{270}Db (see Figure 6.10). As in the case of ^{268}Db, it seems likely that this nucleus decays with a long half-life by EC populating the even–even nucleus ^{270}Rf, which then undergoes SF with a short half-life. Moreover, this isotope is well suited for further chemical investigations or for mass-measurements using ion traps.

The new decay chains of element 117 complete the region of known nuclei located in or close to the island of SHN. A consistent picture emerges concerning decay properties and production cross sections. The heaviest nuclei decay by α emission, and half-lives increase with neutron numbers (see Figures 6.4 and 6.6). Partial fission half-lives exhibit the pronounced influence of odd particle hindrance factors. Hot fusion cross sections decrease exponentially between nobelium and darmstadtium, where at neutron numbers between 166 and 172 the fission barriers traverse a minimum. The cross sections increase again when the island of SHN is approached and reach maximum values of 10 and 8.5 pb for synthesis of element 114 and 115 using targets of ^{244}Pu and ^{243}Am, respectively (see Figure 6.11c,d).

Hot fusion data measured at FLNR were confirmed independently in several subsequent experiments. Also in these experiments beams of ^{48}Ca were used. Targets of ^{238}U were irradiated at SHIP for synthesis of element 112 [36] and targets of ^{242}Pu in a chemistry experiment at FLNR itself [37]. The same target was used at LBNL, where, in addition to the known decay chains from $^{287}114$ and $^{286}114$, a new decay chain starting at $^{285}114$ was measured [38, 39]. Targets of ^{244}Pu were used at the GSI-TASCA (TransActinide Separator and Chemistry Apparatus) gas-filled separator for the synthesis of element 114 [40] and, finally, element 116 was confirmed in irradiations of ^{248}Cm at SHIP [41]. Priority of discovery of two of the six new elements was assigned to the FLNR experiments in 2012 and the names flerovium (Fl) proposed for element 114 and livermorium (Lv) proposed for element 116 were officially accepted by IUPAC.

After the experiment prove that cross sections and half-lives increase when the neutron shell $N = 184$ is approached, attempts were undertaken to synthesize element 120, where a subshell closure is expected and where the CN are located even closer to $N = 184$. At FLNR, the reaction $^{58}\mathrm{Fe} + {}^{244}\mathrm{Pu} \rightarrow {}^{302}120^*$ was investigated in 2007 [42] and at the GSI SHIP the reaction $^{64}\mathrm{Ni} + {}^{238}\mathrm{U} \rightarrow {}^{302}120^*$ in 2007–2008 [43]. No events originating from isotopes of element 120 were observed. Upper cross-sectional limits of 400 and 90 fb, respectively, were measured.

The search for element 120 was continued at SHIP in 2011 within a larger program based on the use of ^{248}Cm as target. During 33 days of beam-on-target, a cross-sectional limit of 560 fb was reached in a first part of the experiment. The decay chain expected from the reaction $^{54}\mathrm{Cr} + {}^{248}\mathrm{Cm} \rightarrow {}^{302}120^*$ is shown in the right-hand side of Figure 6.13. In the case of a 3n channel, known nuclei would be populated after α decay of the so far unknown daughter $^{295}118$. It is also planned to use a ^{51}V beam, aimed at searching for element 119 and at performing reaction studies using a variety of other beam isotopes. This

way it will be possible to fill the still existing gap between the isotopes produced in cold fusion and the more neutron-rich ones produced in hot fusion (see Figure 6.10).

6.5 Summary and Perspectives

The experimental work of the last three decades has shown that cross sections for the synthesis of the heaviest nuclei do not decrease continuously as it was measured up to the production of element 112 using a cold fusion reaction. Recent data on the synthesis of elements 113–118 using hot fusion break this trend when the region of spherical SHN is reached. Results originally obtained in Darmstadt and in Dubna were confirmed in independent experiments and with different methods, including the use of chemical, element-specific properties. Owing to the systematic study of different reactions for the production of SHN and the measurement of excitation functions, the present assignment of the measured decay chains to new relatively neutron-rich isotopes of SHN is highly reliable. The region of the predicted spherical SHN has finally been reached and the exploration of the island has started and can be performed even on a relatively high cross-sectional level.

The progress toward the exploration of the island of spherical SHN is difficult to predict. Hot fusion based on actinide targets and ^{48}Ca beams terminates at element 118, because targets beyond californium can be produced only with tremendous costs and efforts. How heavier beams such as ^{50}Ti and ^{54}Cr will affect the fusion cross section is the subject of experiments being presently performed or that are planned for the near future. However, these heavier beams are mandatory for exploration of the island of SHN into the north and northeast direction. Strong shell effects, if they exist at $Z = 120$ or 126, could positively influence the reaction cross sections.

Also worth studying is the transition from hot to cold fusion expected to occur with actinide targets and the most strongly bound isotopes of iron and nickel for synthesis of elements at $Z = 126$. Losses by reseparation in the entrance channel due to the higher value of $Z_p \times Z_t$ could be compensated by a smaller probability of fission of the CN.

No technical limitations exist for the exploration of the island of SHN in west and northwest direction. Sufficient neutron-deficient projectile isotopes are available. However, owing to Q-value effects, the excitation energy of the CN at the barrier will increase.

Most interesting but also most difficult will be the synthesis of more neutron-rich isotopes located in southeast direction of the island of SHN near the β-stability line and close to the neutron shell $N = 184$. There, α half-lives are long, and in the case of odd–odd nuclei, SF could be significantly hindered so that small amounts of such nuclei might be found even on earth or in meteorites. In the laboratory, reactions using radioactive beams and transfer reactions are options to be studied in the future.

Element 114 was discovered in a hot fusion reaction. However, another possibility could be the cold fusion reaction ^{76}Ge + ^{208}Pb. As predicted by the late W. Swiatecki using his fusion-by-diffusion model [44], the cross section should be considerably higher than that for the synthesis of element 113. This experiment is still waiting to be performed.

The region far beyond element 126 was already addressed theoretically 40 years

ago. For example, a region of relatively higher stability against SF was obtained for spherical nuclei at $Z = 164$ and $N = 318$ in [45]. The properties calculated for toroidal and spherical bubble nuclei were presented in [46]. Although speculative, the now highly advanced experimental technology should be used also to search for such really exotic phenomena in the accessible region of nuclei between $Z = 126$, $N = 184$ and $Z = 190$, $N = 297$ using beams of ions up to ^{238}U.

An opportunity for the continuation of experiments in the region of SHN at low cross sections affords, among others, further accelerator developments. At increased beam currents, values of tens of particle microamperes may become accessible, the cross-sectional level for the performance of experiments can be shifted down into the region of tens of femtobarns, and excitation functions can be measured on the level of tenths of picobarns. High currents, in turn, call for the development of new targets and separator improvements. Radioactive ion beams, not as intense as those with stable isotopes, will allow for approaching the closed neutron shell $N = 184$ already at lighter elements. The study of the fusion process using radioactive neutron-rich beams will be highly interesting.

The half-lives of spherical SHN are expected to be relatively long. On the basis of nuclear models, which are effective predictors of half-lives in the region of the heaviest elements, values from microseconds to years have been calculated for various isotopes. This wide range of half-lives encourages the application of a wide variety of experimental methods in the investigation of SHN, from the safe identification of short-lived isotopes by recoil-separation techniques to atomic physics experiments on trapped ions, and to the investigation of chemical properties of heavy elements using long-lived isotopes. Finally, on the basis of a broader knowledge of the properties of SHN and the mechanisms of their synthesis, it will become possible to answer reliably the questions, can SHN be produced in nature? and where the highest probability may exist for finding them?.

Glossary

Atomic masses (MeV/c^2): The mass $m_{at}(A, Z)$ of an atom with the mass number $A = Z + N$ is smaller than the masses of its components Z protons, Z electrons, and N neutrons. The difference is due to the binding energy given by $B_{at} = [m_{at}(A, Z) - Zm_H - Nm_n]c^2$, with c being equal to the velocity of light. Note that in bound systems the value of B is negative. In tables usually given is the atomic mass excess $\Delta m_{at} = m_{at}(A, Z)(\text{MeV}/c^2)/u - A(\text{MeV}/c^2)$ where u (unified mass unit) $= 931.4940023(7)$ MeV/c^2. Nuclear masses can be calculated from the atomic masses by using the formula $m_{nu}(A, Z) = m_{at}(A, Z) - Zm_e - B_e(Z)/c^2$.

Barn, symbol b: Unit used for expressing the cross-sectional area of nuclei. The cross section of the nucleus ^{120}Sn is approximately 1 b $= 10^{-24}$ cm^2, and that of ^{238}U is 1.7 b. The unit is also used to express the reaction cross section.

Bimodal fission: Two modes of spontaneous fission that can occur in one and the same nucleus with certain probability. The configuration path in the potential energy surface determines whether the nucleus fissions in a compact high-energy mode or in an elongated low-energy mode. Bimodal fission was observed for nuclei in the region of ^{258}Fm.

Branching ratio b_i: Heavy nuclei can decay by α or $\beta^{\pm}$ decay or spontaneous fission. The branching ratio is defined as number of nuclei decaying by one of these decay modes with respect to the total number of decays. The branching ratio is equal to the ratio of the partial decay constant (e.g., λ_α) to the overall decay constant $\lambda_{tot} = \lambda_\alpha + \lambda_\beta + \lambda_{SF}$, $b_\alpha = \lambda_\alpha/\lambda_{tot}$.

Closed shells: See shell model.

Compound nucleus (CN): Intermediate system in a fusion reaction consisting of the sum of all protons and neutrons of projectile and target. Cooling of the excited CN by emission of neutrons, protons, α particles, and γ-rays results in an evaporation residue (ER) in the ground state. CN in the region of heavy elements have a high probability of fission.

Cross section σ (barn): The cross section is a measure to quantify the probability of reaction channels.

Excitation energy E^* (MeV): The energy difference between the ground state of a nucleus to any particular excited state. The excitation energy of a compound nucleus produced in a fusion reaction is given by $E^* = Q + E_{cm}$, with Q being equal to the Q-value and E_{cm} being equal to the kinetic energy in the center-of-mass system.

Excitation function: The measurement of reaction cross sections as function of beam energy.

Fermi level (MeV): Supposed energy between the last occupied and the first empty level in a system of fermions such as protons or neutrons in a nucleus in the case of a sharp occupation of the energy levels.

Fission barrier: The binding energy of a nucleus as function of deformation shows a barrier-like shape, which can be further modulated by shell effects of the individual nucleus. The ground state of the nucleus is located at the deepest minimum inside the barrier. Fission of the nucleus follows along a path of minimum barrier height so that at the barrier maximum a geometrical saddle point exists. Fission occurs by quantum mechanical tunneling through this barrier.

Hindrance factor: Relative number quantizing the amount of delay in a usually not well-understood nuclear transition with respect to a well- or better-understood one. Hindrance factors are determined for spontaneous fission of odd-A or odd–odd nuclei relative to the neighboring even–even nuclei, or for α decay relative to a calculated value of same energy, but assuming a transition between states without change of angular momentum.

***K* isomer:** The projection of the angular momentum vector on the symmetry axis of a deformed nucleus has the value $K\hbar$. It is zero for even–even nuclei in the ground state. The breaking of a pair of nucleons in such a nucleus results in an isomeric state having a high value of $K\hbar$, if the nucleon pair occupied a level of high spin. K isomers also occur in odd nuclei and in the case of breaking of more than one pair. The heaviest nucleus for which a K isomer was observed is ^{270}Ds.

Liquid-drop model: The model considers the nucleus as a charged liquid drop. Many nuclear properties, for example, the overall trend of nuclear binding energies as function of proton and neutron numbers resulting in the valley of stability, are well described. Not described are local structures of the binding energy based on shell effects.

Lifetime τ (s): The lifetime is the inverse of the decay constant, $\tau = 1/\lambda$. It is related to the half-life through $\tau = T_{1/2}/\ln 2$.

Partial half-life $T_{1/2,i}$ (s): The partial half-life is related to the partial decay constant

λ_i through $T_{1/2,i} = \ln 2/\lambda_i$ (see branching ratio). In experiments, always the total half-life is measured, even if only one decay mode (e.g., α decay) is measured.

Q-value (MeV): Difference of masses between all nuclei before and all nuclei after the reaction. In the case of a fusion reaction, the Q-value is given by $Q = (m_{\text{projectile}} + m_{\text{target}} - m_{\text{CN}})c^2$.

Reaction channel: A nuclear reaction can result in a number of different final configurations called reaction channels. In the case of particle evaporation from a compound nucleus the evaporation channels can be xn, pxn, αxn where x = 1, 2, 3 and so on. Similarly a nucleus may decay in various decay channels (e.g., emission of α particles, $\beta^{\pm}$ decay, or spontaneous fission).

Saddle point: See fission barrier.

Shell model: The model describes the energy levels occupied by the nucleons protons and neutrons in the nuclear potential. This potential is formed by the attractive nuclear forces acting at the surface. Inside the potential, the energy levels are arranged according to their quantum numbers, resulting in energy degeneracy (shells) of levels having the same number of knots in the wave function (main quantum number) and the same total angular momentum. The known closed shells or magic numbers are obtained when the correct spin–orbit potential is used.

Specialization energy (MeV): Increase in the fission barrier due to the strong rise of high-spin single-particle levels at prolate deformation and low-spin levels at oblate deformation.

Spin–orbit potential (MeV): Term in the nuclear potential describing the coupling between spin and orbital angular momentum of single particles. For nucleons, this term results in higher binding energy when spin and orbital angular momentum are parallel and vice versa. This is different from the arrangement of electrons in the Coulomb potential of the nucleus where electrons are stronger bound when spin and orbital angular momentum are antiparallel.

References

1. Strutinsky, V.M. (1967) *Nucl. Phys. A*, **95**, 420.
2. Cwiok, S. and Sobiczweski, A. (1992) *Z. Phys. A*, **342**, 203.
3. Cwiok, S., Dobaczewski, J., Heenen, P.H., Magierski, P., and Nazarewicz, W. (1996) *Nucl. Phys. A*, **611**, 211.
4. Bender, M., Heenen, P.H., and Reinhard, P.G. (2003) *Rev. Mod. Phys.*, **75**, 121.
5. Smolanczuk, R., Skalski, J., and Sobiczewski, A. (1995) *Phys. Rev. C*, **52**, 1871.
6. Smolanczuk, R. and Sobiczewski, A. (1995) Shell effects in the properties of heavy and superheavy nuclei, in *Proceedings of the XVth Nuclear Physics Divisional Conference on Low Energy Nuclear Dynamics* St. Petersburg, Russia, World Scientific, Singapore, p. 313.
7. Möller, P., Nix, J.R., and Kratz, K.L. (1997) *At. Data Nucl. Data Tables*, **66**, 131.
8. Polikanov, S.M., Druin, V.A., Karnaukhov, V.A., Mikheev, V.L., Pleve, A.A., Skobelev, N.K., Subbotin, V.G., Ter-Akopian, G.M., and Fomichev, V.A. (1962) *Zh. Eksp. Teor. Fiz.*, **42**, 1016. (1962) *Sov. Phys. JETP*, **15**, 1016.
9. Balagna, J.P., Ford, G.P., Hoffman, D.C., and Knight, J.D. (1971) *Phys. Rev. Lett*, **26**, 145.
10. Hulet, E.K., Wild, J.F., Dougan, R.J., Lougheed, R.W., Landrum, J.H., Dougan, A.D., Schädel, M., Hahn, R.L., Baisden, P.A., Henderson, C.M. *et al.* (1986) *Phys. Rev. Lett.*, **56**, 313.
11. Armbruster, P., Eidens, J., Grueter, J.W., Lawin, H., Roeckl, E., and Sistemich, K. (1971) *Nucl. Instrum. Methods*, **91**, 499.
12. Münzenberg, G., Faust, W., Hofmann, S., Armbruster, P., Güttner, K., and Ewald, H. (1979) *Nucl. Instrum. Methods*, **161**, 65.

13. Hofmann, S., Faust, W., Münzenberg, G., Reisdorf, W., Armbruster, P., Güttner, K., and Ewald, H. (1979) *Z. Phys. A*, **291**, 53.
14. Hofmann, S. and Münzenberg, G. (2000) *Rev. Mod. Phys.*, **72**, 733.
15. Schmidt, K.H., Simon, R.S., Keller, J.G., Heßberger, F.P., Münzenberg, G., Quint, B., Clerc, H.G., Schwab, W., Gollerthan, U., and Sahm, C.C. (1986) *Phys. Lett. B*, **168**, 39.
16. Seaborg, G.T. (1994) *Modern Alchemy: Selected Papers of Glenn T. Seaborg*, World Scientific, Singapore.
17. Audi, G. and Wapstra, A.H. (1993) *Nucl. Phys. A*, **565**, 1.
18. Myers, W.D. and Swiatecki, W.J. (1996) *Nucl. Phys. A*, **601**, 141.
19. Oganessian, Y.T., Tretyakov, Y.P., Iljinov, A.S., Demin, A.G., Pleve, A.A., Tretyakova, S.P., Plotko, V.M., Ivanov, M.P., Danilov, N.A., Korotkin, Y.S., and Flerov, G.N. (1974) *JETP Lett.*, **20**, 265.
20. Münzenberg, G. (1988) *Rep. Prog. Phys.*, **51**, 57.
21. Hofmann, S. (1998) *Rep. Prog. Phys.*, **61**, 639.
22. Bass, R. (1974) *Nucl. Phys. A*, **231**, 45.
23. Oganessian, Y.T., Utyonkov, V.K., Lobanov, Y.V., Abdullin, F.S., Polyakov, A.N., Shirokovsky, I.V., Tsyganov, Y.S., Gulbekian, G.G., Bogomolov, S.L., Gikal, B.N. *et al.* (2004) *Phys. Rev. C*, **70**, 064609.
24. Sobiczewski, A. (2011) *Acta Phys. Pol.*, **42**, 1871.
25. Morita, K., Morimoto, K., Kaji, D., Haba, H., Ideguchi, E., Kanungo, R., Katori, K., Koura, H., Kudo, H., Ohnishi, T. *et al.* (2004) *Eur. Phys. J., A*, **21**, 257.
26. Morita, K., Morimoto, K., Kaji, D., Haba, H., Ideguchi, E., Peter, J.C., Kanungo, R., Katori, K., Koura, H., Kudo, H. *et al.* (2004) *J. Phys. Soc. Jpn.*, **73**, 1738.
27. Morita, K., Morimoto, K., Kaji, D., Akiyama, T., Goto, S., Haba, H., Ideguchi, E., Katori, K., Koura, H., Kudo, H. *et al.* (2007) *J. Phys. Soc. Jpn.*, **76**, 043201.
28. Folden, C.M. III,, Gregorich, K.E., Düllmann, C.E., Mahmud, H., Pang, G.K., Schwantes, J.M., Sudowe, R., Zielinski, P.M., Nitsche, H., and Hoffman, D.C. (2004) *Phys. Rev. Lett.*, **93**, 212702.
29. Ninov, V., Gregorich, K.E., Loveland, W., Ghiorso, A., Hoffman, D.C., Lee, D.M., Nitsche, H., Swiatecki, W.J., Kirbach, U.W., Laue, C.A. *et al.* (1999) *Phys. Rev. Lett.*, **83**, 1104.
30. Ninov, V., Gregorich, K.E., Loveland, W., Ghiorso, A., Hoffman, D.C., Lee, D.M., Nitsche, H., Swiatecki, W.J., Kirbach, U.W., Laue, C.A. *et al.* (2002) *Phys. Rev. Lett.*, **89**, 039901.
31. Morita, K., Morimoto, K., Kaji, D., Akiyama, T., Haba, H., Ozeki, K., Kudou, Y., Sumita, T., Wakabayashi, Y., Yoneda, A. *et al.* (2012) *J. Phys. Soc. Jpn.*, **81**, 103201.
32. Oganessian, Y.T. (2007) *J. Phys. G: Nucl. Part. Phys.*, **34**, R165.
33. Oganessian, Y.T., Abdullin, F.S., Dmitriev, S.N., Gostic, J.M., Hamilton, J.H., Henderson, R.A., Itkis, M.G., Moody, K.J., Polyakov, A.N., Ramayya, A.V. *et al.* (2012) *Phys. Rev. Lett.*, **108**, 022502.
34. Oganessian, Y.T., Abdullin, F.S., Bailey, P.D., Benker, D.E., Bennett, M.E., Dmitriev, S.N., Ezold, J.G., Hamilton, J.H., Henderson, R.A., Itkis, M.G. *et al.* (2010) *Phys. Rev. Lett.*, **104**, 142502.
35. Oganessian, Y.T., Abdullin, F.S., Alexander, C., Binder, J., Boll, R.A., Dmitriev, S.N., Ezold, J., Felker, K., Gostic, J.M., Grzywacz, R.K. *et al.* (2012) *Phys. Rev. Lett.*, **109**, 162501.
36. Hofmann, S., Ackermann, D., Antalic, S., Burkhard, H.G., Comas, V.F., Dressler, R., Gan, Z., Heinz, S., Heredia, J.A., Heßberger, F.P. *et al.* (2007) *Eur. Phys. J. A*, **32**, 251.
37. Eichler, R., Aksenov, N.V., Belozerov, A.V., Bozhikov, G.A., Chepigin, V.I., Dmitriev, S.N., Dressler, R., Gäggeler, H.W., Gorshkov, V.A., Haenssler, F. *et al.* (2007) *Nature*, **447**, 72.
38. Stavsetra, L., Gregorich, K.E., Dvorak, J., Ellison, P.A., Dragojevic, I., Garcia, M.A., and Nitsche, H. (2009) *Phys. Rev. Lett.*, **103**, 132502.
39. Ellison, P.A., Gregorich, K.E., Berryman, J.S., Bleuel, D.L., Clark, R.M., Dragojevic, I., Dvorak, J., Fallon, P., Fineman-Sotomayor, C., Gates, J.M. *et al.* (2010) *Phys. Rev. Lett.*, **105**, 182701.
40. Düllmann, C.E., Schädel, M., Yakushev, A., Türler, A., Eberhardt, K., Kratz, J.V., Ackermann, D., Andersson, L.-L., Block, M., Brüchle, W. *et al.* (2010) *Phys. Rev. Lett.*, **104**, 252701.
41. Hofmann, S., Heinz, S., Mann, R., Maurer, J., Khuyagbaatar, J., Ackermann, D., Antalic, S., Barth, W., Block, M., Burkhard, H.G. *et al.* (2012) *Eur. Phys. J. A*, **48**, 62.
42. Oganessian, Y.T., Utyonkov, V.K., Lobanov, Y.V., Abdullin, F.S., Polyakov, A.N.,

Sagaidak, R.N., Shirokovsky, I.V., Tsyganov, Y.S., Voinov, A.A., and Mezentsev, A.N. (2009) *Phys. Rev. C*, **79**, 024603.
43. Hofmann, S., Ackermann, D., Antalic, S., Comas, V.F., Heinz, S., Heredia, J.A., Heßberger, F.P., Khuyagbaatar, J., Kindler, B., Kojouharov, I. *et al.* (2009) Probing Shell Effects at Z=120 and N=184, GSI Scientific Report 2008, GSI Report 2009-1, GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt p. 131.
44. Swiatecki, W.J., Siwek-Wilczyńska, K., and Wilczyński, J. (2005) *Phys. Rev. C*, **71**, 014602.
45. Grumann, J., Mosel, U., Fink, B., and Greiner, W. (1969) *Z. Phys.*, **228**, 371.
46. Wong, C.Y. (1973) *Ann. Phys.*, **77**, 279.

Further Readings

Hofmann, S. (2002) *On Beyond Uranium – Journey to the End of the Periodic Table*, Taylor & Francis, London and New York, p. 216.
Seaborg, G.T. and Loveland, W.D. (1990) *The Elements Beyond Uranium*, John Wiley & Sons, Inc., New York, p. 359.

7
Nuclear γ-Spectroscopy and the γ-Spheres

Mark Riley and John Simpson

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

7.1 Introduction

γ-ray spectroscopic techniques play a vital role in our investigations into the properties and behavior of the unique strongly interacting aggregation of fermions that we call the atomic nucleus. These studies of the γ-ray emissions from excited nuclei reveal an extremely rich system that displays a wealth of static and dynamical facets. The number of nucleons is sufficient (<300) to allow correlations, but it is still finite. Thus, nuclei exhibit a variety of collective properties yet are simple enough to display both single-particle properties and a single-particle basis of the collective effects. The remarkable diversity of phenomena and symmetries exhibited by nuclei continues to surprise and fascinate scientists, as unexpected properties are continually revealed by new experimental investigations arising from the development of increasingly sensitive instrumentation and new accelerator developments.

Every major advance in γ-ray detector technology has resulted in the discovery of new phenomena, bringing significant fresh insight into the structure of nuclei. The different time periods or eras associated with different detector technical advances are presented in Figure 7.1. At the present time, we find ourselves at the transition from the ''γ-spheres'' or large 4π arrays of escape-suppressed spectrometers, such as Gammasphere and Euroball, to the beginning of the development of 4π Ge shell arrays, such as the Gamma-Ray Energy Tracking Array (GRETA) in the USA and the Advanced GAmma Tracking Array (AGATA) in Europe (see Ref. [1] and references therein). These latter systems will abandon physical suppression shields and instead employ the technique of γ-ray tracking in electrically segmented Ge crystals. As a first step toward the implementation of these 4π Ge arrays, the $\sim 1\pi$ systems, Gamma-Ray Energy Tracking In beam Nuclear Array (GRETINA) in the USA and the first phase of AGATA in Europe have been recently constructed and have just begun their initial physics campaigns.

This chapter summarizes significant recent developments in high-resolution γ-ray spectroscopy. A special emphasis is placed on the new revolutionary technology of γ-ray tracking. Descriptions of a number of current large γ-ray arrays in operation around the world are given, together with an outlook into the future of this exciting field. More detailed technical reviews of many of these recent developments in Ge-based detectors along with a variety of physics highlights may be found in Refs [2, 3]. Developments and highlights from

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

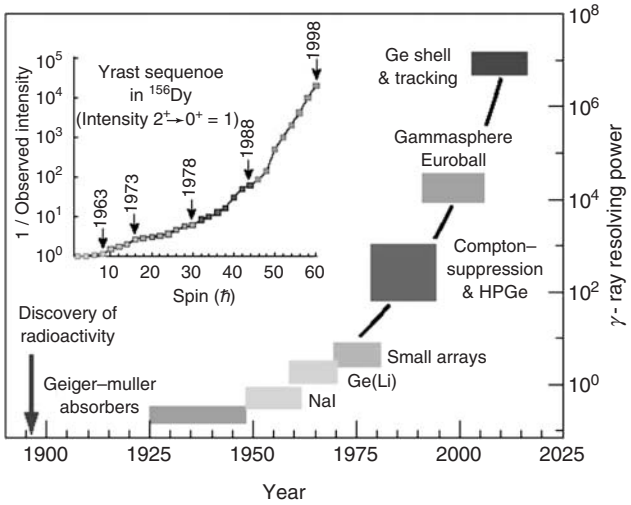


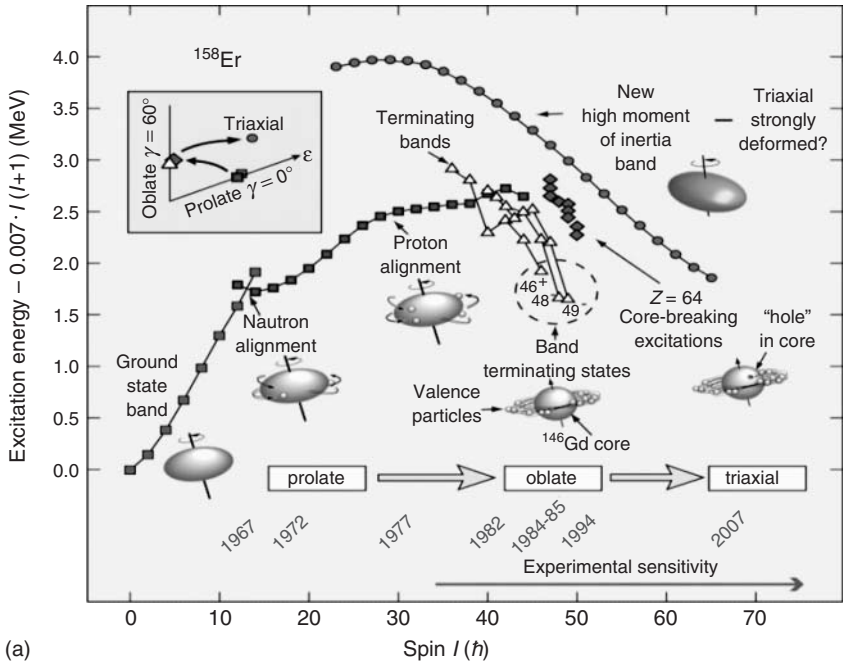
Figure 7.1 The evolution of γ-ray detector technology. The calculated γ-ray resolving power is a measure of the ability to observe faint emissions from rare and exotic nuclear states. This is illustrated in the upper left inset, which indicates the strong relationship between resolving power and the inverse of the experimental observational limit.

the previous two decades are covered in Refs [4, 5].

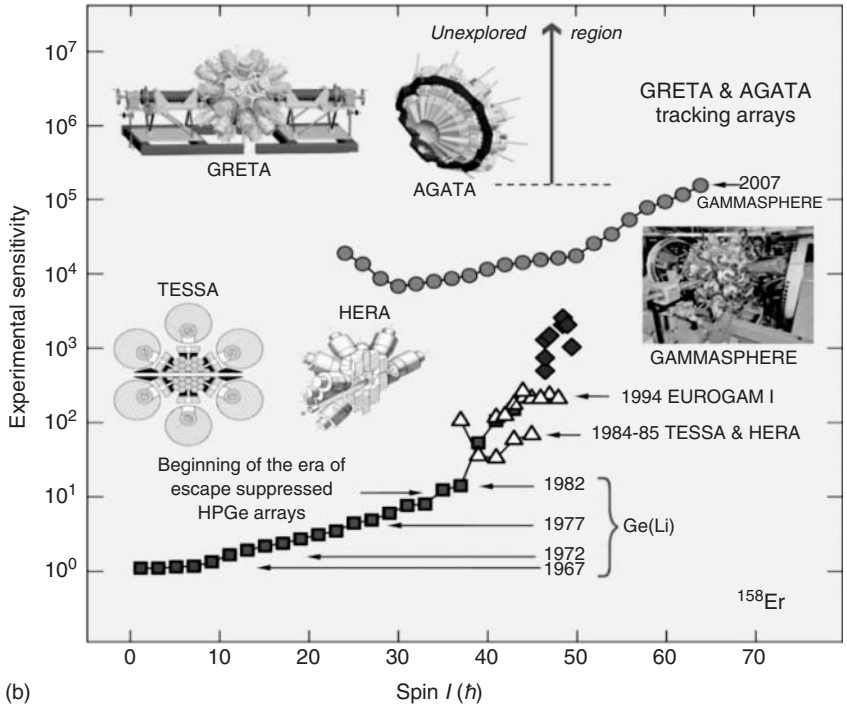
7.2 An Example: the Spectroscopy of ^{158}Er through the Decades

An excellent example of the evolution of γ-ray detector systems through the decades, which illustrates the ability to discover new physical phenomena, comes from the study of the yrast states (lowest energy state for a given angular momentum) in the rare-earth nucleus, ^{158}Er, as displayed in Figures 7.2 and 7.3 (see also Refs [6–9] and references therein). At very low spin values, the yrast states form a rotational band corresponding to a spinning collective structure with a low moment of inertia value of ∼30% of the rigid body value, which reveals that superfluid behavior or pairing correlations play a significant role. However, with increasing angular momentum, or spin, the Coriolis force induces the breaking or rotational alignment of specific high-j pairs of particles in a process known as

Figure 7.2 (a) The evolution of nuclear structure in ^{158}Er with excitation energy and angular momentum (spin). The inset illustrates the changing shape of ^{158}Er with increasing spin within the standard (ε, γ) deformation plane. The parameters ε and γ represent the eccentricity from sphericity and triaxiality, respectively. (b) The experimental sensitivity of detection (proportional to the inverse of the observed γ-ray intensity) is plotted as a function of spin showing the progression of γ-ray detector techniques with time that are associated with nuclear structure phenomena in ^{158}Er. "TESSA," "HERA," "EUROGAM I," "GAMMASPHERE," "GRETA," and "AGATA" are specific names of γ-ray detector arrays.


^{158}Er
Oblate $\gamma = 60°$
Triaxial
Prolate $\gamma = 0°$
ε
Terminating bands
New high moment of inertia band
Triaxial strongly deformed?
Proton alignment
$^{46+}_{48}$
49
Z = 64 Core-breaking excitations
"hole" in core
Nautron alignment
Ground state band
Band terminating states
Valence particles
^{146}Gd core
prolate
oblate
triaxial
1967
1972
1977
1982
1984-85
1994
2007
Experimental sensitivity
Excitation energy − 0.007 · I (I+1) (MeV)
Spin I (ħ)

(a)


Unexplored region
GRETA & AGATA tracking arrays
GRETA
AGATA
2007 GAMMASPHERE
GAMMASPHERE
TESSA
HERA
1994 EUROGAM I
1984-85 TESSA & HERA
Beginning of the era of escape suppressed HPGe arrays
1982
1977
1972
1967
Ge(Li)
^{158}Er
Experimental sensitivity
Spin I (ħ)

(b)

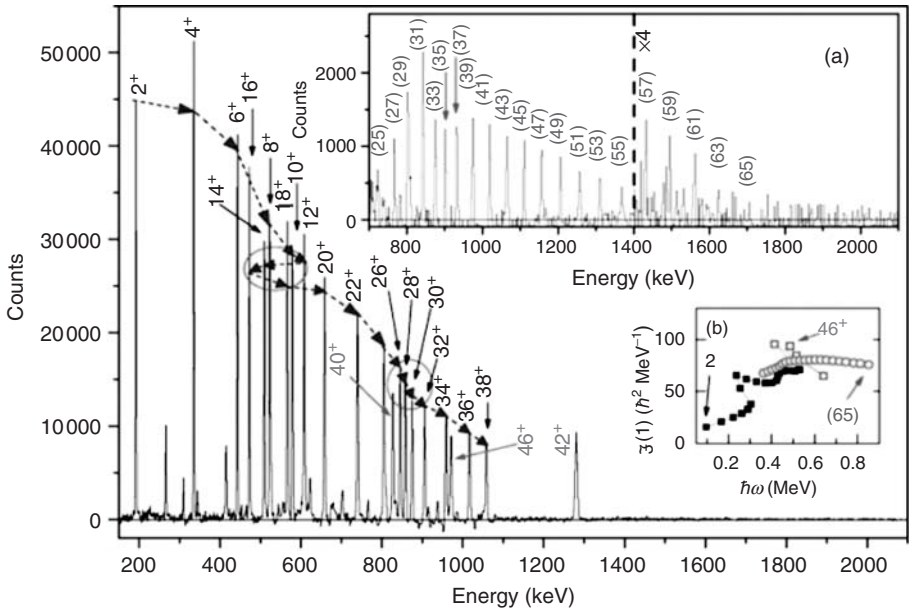


Figure 7.3 Spectrum of γ rays representative of the yrast band in ^{158}Er in coincidence with the 44^+ to 42^+ γ-ray transitions. Transitions are all labeled with the states (spin and parity) from which they decay. The observed neutron (energy ∼ 550 keV, spin $I \sim 14$) and proton (energy ∼ 850 keV, $I \sim 28$) rotational alignments are marked in circles. (a) Spectrum representative of the strongest collective band at ultrahigh spin observed in ^{158}Er (band 1). The portion of this spectrum above 1400 keV has been magnified by a factor of 4. The spins assigned are tentative. (b) Kinematic moments of inertia, $J^{(1)}$, as a function of rotational frequency $\hbar\omega$, for the yrast sequence (collective prolate states: filled squares; noncollective oblate states: open squares) and band 1 (open circles) in ^{158}Er.

backbending at spin $I \sim 14$ for a pair of $i_{13/2}$ neutrons and at $I \sim 28$ for a pair of $h_{11/2}$ protons. These rotational alignments cause large changes in the moment of inertia behavior. Another sudden change in the moment of inertia along the yrast line occurs at spin $I \sim 38$. These features were all observed using Ge(Li) detectors with (now) relatively low resolving powers. It was only with the development of escape-suppressed arrays that this latter anomaly in the moment of inertia could be delineated and explained. It is now understood as a dramatic shape transition from a collective prolate rotation to noncollective or weakly collective oblate configurations along the yrast line at spins 40–50. This transition manifests itself as energetically favored, fully aligned band termination states at 46^+, 48^-, and 49^- (Figure 7.2).

It has taken two decades and the development of large 4π escape-suppressed γ-ray detector arrays to discover what happens above "band termination" in ^{158}Er. Using the Gammasphere spectrometer, a return to collective rotation at the highest spins was observed when rotational structures of very low intensity ($\sim 10^{-4}$ of the respective channel intensity), very high moments of inertia, and large quadrupole moments were found [7, 8] (Figure 7.3). These structures extended discrete-line spectroscopy in this nucleus up to $I \sim 65$.

7.3 Modern High-Resolution γ-Ray Spectroscopy

The most important properties of a γ-ray detector array are (i) high efficiency in detecting incident γ rays, (ii) high resolution resulting in very narrow full-energy peaks, (iii) high ratio of full-energy to partial-energy events, (iv) high counting rates, and (v) high granularity to localize individual γ rays and reduce the probability of two γ-ray hits in one detector from the same event. For γ rays in the few megaelectronvolt range, by far the best combination of these properties is given by semiconductors made of high-purity germanium (Ge) crystals. The largest such crystals that can currently be produced commercially are cylinders about 10 cm in diameter and 10 cm long which, for about 28% of incident 1 MeV γ rays, produce a full-energy peak with a full width at half its maximum height of about 2 keV. To improve peak-to-background performance, escape-suppression techniques have been employed. To improve efficiency, a number of Ge crystals have been placed together in the same cryostat in, for example, clover (four) and cluster (seven) detectors.

7.3.1 The Escape-Suppression Principle

For a better ratio of full-energy to total recorded events (called the *peak-to-total* or *P/T ratio*), the Ge detectors are surrounded by a dense scintillator (bismuth germanate (BGO) being the most common), which detects γ rays that Compton scatter or escape out of the Ge crystal. This detection triggers the electronics system to reject or suppress the partial-energy pulse in the Ge detector (Figure 7.4 and Figure 7.5). This results in an improvement in the peak-to-total ratio for a 1.3 MeV γ-ray from about 0.28 for the bare crystal to about 0.6. This is an enormously important signal to background gain, without which high-fold coincidence measurements would not be practical. For example, for a typical situation in the Gammasphere spectrometer when six γ rays hit separate Ge detectors, the fraction of events with full-energy photo-peaks rises by a factor of about 100 with escape-suppression. For the increase in both efficiency and granularity, escape-suppressed detectors are assembled into arrays. The first such array, called the escape-suppressed spectrometer array (TESSA), was set up at the Niels Bohr Institute in 1980 and consisted of five Ge(Li) detectors surrounded by NaI(Tl) shields. Many arrays worldwide were based on this escape-suppressed technology, one of the latest being Gammasphere with ∼100 Ge crystals surrounded by BGO shields.

7.3.2 γ-Ray Tracking

This is a time of great opportunity in nuclear spectroscopy. The development of radioactive beam capabilities around the world is opening a new landscape for discovery, and the connections between new nuclear structure studies and astrophysics, neutrino physics, and physics beyond the standard model are stronger than ever. New detector technologies have evolved that can meet the challenges of the new generation of experiments. The new technology of "γ-ray tracking," which was first discussed in the early 1990s, is poised to revolutionize γ-ray spectroscopy in a way that large arrays of escape-suppressed γ detectors have done previously. Tracking arrays covering roughly 1π of solid angle have been constructed in the USA (GRETINA)

Figure 7.4 Schematic of a modern escape-suppressed spectrometer array.

and Europe (AGATA). Both devices are now engaged in physics campaigns. Hopefully the momentum in developing this technology to its full potential will continue toward full 4π GRETA and AGATA spectrometers, which will have unsurpassed sensitivity and discovery potential. These calorimeters will carry γ-ray spectroscopy into the next generation and will be needed to fully exploit the science opportunities at radioactive beam facilities as well as to greatly increase the reach of stable beam facilities. In addition, γ-ray tracking technology will have a wide range of important applications, such as in environmental monitoring, SPECT and PET medical imaging systems, and in homeland security situations.

This next major step in γ-ray spectroscopy is therefore to abandon the concept of a physical escape-suppression shield, which greatly reduces the overall possible efficiency, and move toward the goal of a 4π Ge ball utilizing the technique of γ-ray energy tracking in electrically segmented Ge crystals (Figure 7.6). With the new tracking technique, the position and energy of γ-ray interaction points are

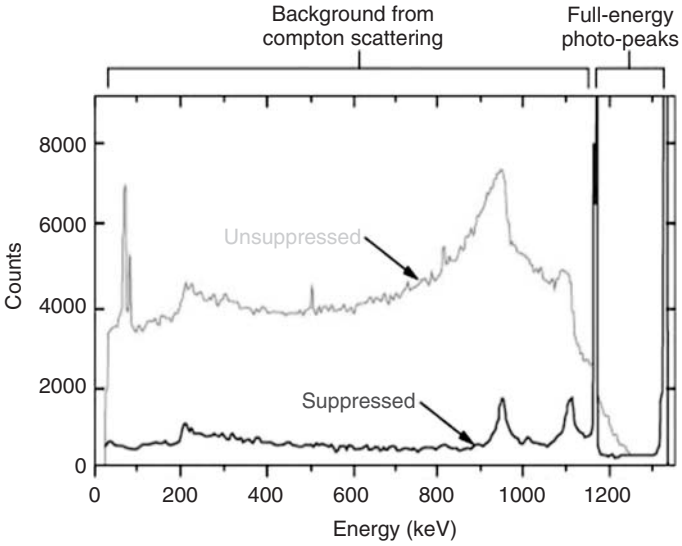


Figure 7.5 Unsuppressed and suppressed spectra for ^{60}Co illustrating the huge gains in peak-to-background performance from the use of escape-suppressed shielding techniques. Note that the full-energy photo-peaks have been normalized in the two spectra.

identified in the detector segments. As most γ rays interact more than once within the crystal, the energy–angle relationship of the Compton scattering formula is used to track the path of a given γ-ray. The full γ-ray energy is obtained by summing only the interactions belonging to that particular γ-ray. In this way, there are no vetoed Compton scatters in the suppression shields and scattered γ rays between crystals are recovered. Thus, the planned 4π GRETA and AGATA will have a high overall efficiency, ∼60% for a single 1 MeV γ-ray. Other key benefits of a tracking array include good peak-to-total ratio (∼85%), high counting rate (∼50 kHz) capability per crystal, excellent position resolution (∼2 mm), the ability to handle high multiplicities without summing, the ability to pick out low-multiplicity events hidden in a high background environment, and high sensitivity for linear polarization measurements. For many experiments, these 4π tracking arrays, as shown in Figure 7.1, will provide orders of magnitude improvement in resolving power over present arrays.

Major technological advances and a huge amount of R&D have been necessary in order to make tracking spectrometers possible. These include the fabrication of highly segmented Ge detectors, the production of fast digital electronics, efficient signal analysis and tracking algorithms, and improved computational power. In the following, we give a summary of these technical accomplishments.

7.3.2.1 Segmented Ge Detectors

A critical detector technology is the manufacture of segmented coaxial germanium detectors, which provide signals with sensitivity for locating interaction points in three dimensions. Both GRETINA and AGATA

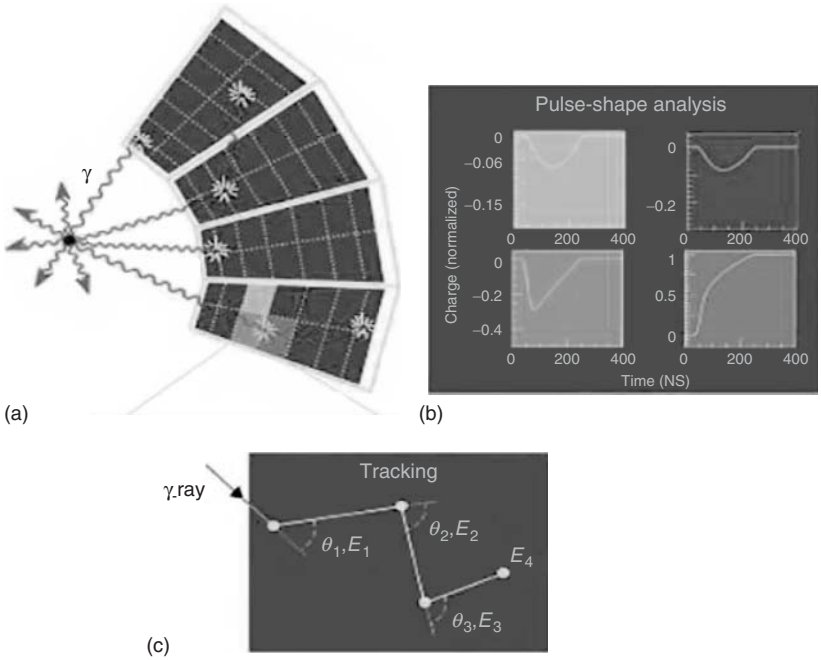


Figure 7.6 Principles of the γ-ray tracking technique. (a) A tracking array will consist of a closed shell of segmented Ge detectors. (b) Pulse-shape analysis of signals from segments containing the interaction(s), as well as analysis of transient signals in adjacent segments, allows the determination of the three-dimensional locations of the interactions and their energies. (c) Tracking algorithms, which are based on the underlying physical processes such as Compton scattering or pair production, are able to identify and separate γ rays and to determine the scattering sequence.

use coaxial Ge crystals that are 36-fold segmented. These crystals have a length of ∼9 cm and a diameter of ∼8 cm at the rear. They are shaped into tapered irregular hexagons for close packing into a spherical shell with high solid angle coverage. In AGATA the taper angle is 8°, and for GRETINA/GRETA it is 10°. These angles are a result of the efficient packing of the different numbers of detectors envisaged in the full 4π geometry. The geometries are based on tiling a sphere with the geodesic arrangement of irregular hexagons and regular pentagons; for AGATA the full array comprises 180 hexagons and 12 pentagons, and for GRETA 120 hexagons and 12 pentagons. These geometries result in three slightly different crystal shapes for AGATA and two for GRETA.

7.3.2.2 Digital Electronics

While the segmentation of the Ge crystal provides the initial position-of-interaction information in order to perform γ-ray tracking, the positions and energies of the γ-ray interactions in the Ge must be even more accurately determined from the signal waveforms. The position resolution is a key metric in the detector performance. It directly affects the efficiency, peak/total, and the tracking efficiency of the array, as well as the effective energy resolution of the array when the emitting nucleus is traveling with a high recoil velocity.

AGATA and GRETA require state of the art, purpose-built digital electronics, and an associated data acquisition system to process the signals from the Ge crystals. Signals from the 37 preamplifiers of each detector (36 segments and 1 central contact) are digitized. From the digitized pulse shapes, processing units derive and provide additional parameters such as the channel number, the energy, the raw data samples from the leading edge of the pulse (trace), and a timestamp (in 10 ns steps). The ''data-acquisition'' section of the system collates and distributes the events to a computer cluster to perform signal decomposition, tracking, and event reconstruction. These systems need to cope with a large numbers of channels (over 6000) and also with very high rates in each crystal (up to 50 kHz). The principle of the AGATA and GRETA systems is to sample these outputs with fast ADCs to preserve the full signal information so that accurate energies, times, and positions can be extracted using the signal decomposition algorithms. Interfaces to enable ''easy'' electronic coupling of ancillary detectors to the acquisition system have been developed.

7.3.2.3 Signal Decomposition

In order to perform γ-ray tracking, the positions and energies of the γ-ray interactions in the Ge (usually several Compton scattering events, followed by photoelectric absorption) must be accurately determined from the signal waveforms. The procedure must handle cases where two or more interactions occur within one of the detector segments. For each crystal, a simulation is performed on a grid of points whose spacing reflects the sensitivity of the detector, depending on the geometry, bias voltage, and crystal impurities. About 300 000 grid points per crystal are used, with spacing varying from 0.5 to 3 mm. Measured signals are then compared with linear combinations of these simulated signals, as there are typically multiple interaction points in a given crystal, with the best fit giving the location and charge (energy) of the interaction points.

The signal decomposition itself can be performed using a variety of algorithms, but to perform it in real time, count-rate requirements and computer costs impose stringent limits on how much CPU time can be spent per event. At present, the GRETINA decomposition code uses a two-step process that starts with an adaptive grid search for one and two interactions per segment followed by a sequential quadratic programming (non-linear least-squares) fit that allows multiple interactions in multiple segments within a crystal. On the current generation of 2 GHz processors, the GRETINA algorithm requires <16 ms per CPU core per crystal.

It has been shown experimentally that the pulse shape algorithms developed for GRETINA and AGATA can achieve an average position resolution of better than 2 mm (RMS), which is sufficient for efficient γ-ray tracking.

7.3.2.4 Tracking

The tracking process uses the energies and positions of the interaction points, produced by the signal decomposition, to determine the scattering sequence for a particular γ-ray. Algorithms have been developed to track events based on Compton scattering, pair-production, and photoelectric interactions. For events with multiple γ rays, these algorithms are able to resolve and identify interaction points belonging to a particular individual γ-ray. γ Rays that deposit only partial

energy in the detector are identified and rejected. Tracking efficiencies ranging from ∼100 to 50% respectively for γ-ray multiplicities from 1 to 25 have been calculated.

7.3.3 Auxiliary Detectors

Auxiliary detectors used in conjunction with large γ-ray arrays have played, and continue to play, a vital role because they augment tremendously the physics that can be achieved. These systems, which detect emitted light charged particles, heavy recoils, emitted neutrons, or plunger devices for precise timing measurements, are all designed to physically combine with the large γ-ray arrays. The latter are therefore different to other systems such as recoil separators, which also play a critical role in expanding the physics discovery potential of the large γ-ray arrays. With γ-ray tracking arrays, auxiliary detectors and separators will perhaps play an even more important role. This is because in order to take advantage of the considerably smaller opening angle (∼2° in θ and ϕ for GRETINA/AGATA compared with 7° for Gammasphere) and to make good use of this angular resolution for Doppler corrections, the angle of the recoiling nucleus must now also be determined with comparable angular resolution.

Thus, a new generation of auxiliary detectors, often based on previously successful systems used with Gammasphere and Euroball, are being constructed with higher segmentation and higher counting rate capabilities. The number of auxiliary detector systems is large, and thus the reader is referred to the respective references and web pages of the γ-ray systems discussed in the following section.

7.4 Large High-Resolution γ-Ray Detector Arrays

As illustrated in Figure 7.1, we are now in a transitional period moving from the era of large 4π escape-suppressed Ge arrays, illustrated by Gammasphere and Euroball, into the new era of Ge tracking arrays GRETINA/GRETA and AGATA. These major systems developed respectively by the collective γ-ray communities in the USA and Europe are summarized below along with a number of other significant large high-resolution γ-ray arrays in current use around the globe. For further detailed information on these systems, see Ref. [3] and references therein, along with the web pages listed in the following sections.

7.4.1 GRETINA/GRETA

GRETINA consists of seven modules, each containing four 36-fold segmented germanium detectors. It allows for localization of individual interaction points to an accuracy of 1–2 mm. These interactions are then tracked to reconstruct the energies of individual γ rays and reject events with incomplete energy collection. The excellent position resolution allows more accurate Doppler-shift corrections, making GRETINA an ideal detector for measurements of γ rays emitted in flight.

GRETINA was completed in April 2011. Following a year of testing and commissioning at LBNL by a collaboration of scientists from 12 US Laboratories and Universities, GRETINA was successfully installed at the National Superconducting Cyclotron Laboratory (NSCL), and began its first science campaign in Summer 2012 (Figure 7.7). Following the fast-beam

Figure 7.7 GRETINA at the NSCL.

campaign at NSCL, GRETINA is scheduled to move to Argonne National Laboratory in mid-2013 for its next science campaign.

GRETINA is the first phase of GRETA and it is hoped to complete GRETA for ''day 1'' experiments at the future Facility for Rare Isotope Beams (FRIB) at Michigan State University, which is due for completion at the end of the decade. Prior to FRIB operations, GRETA will enable forefront experiments at existing stable and radioactive beam facilities with its scientific reach increasing significantly as more detectors are added to GRETINA. The completion of GRETA requires 23 additional detector modules together with the associated electronics, computing, and mechanical support.

For further information about the GRETINA/GRETA project, see *http://www.physics.fsu.edu/GRETINA.org/*.

7.4.2 AGATA

The AGATA collaboration includes more than 40 institutes in 13 countries in Europe, with the aim of developing and constructing a 4π tracking spectrometer. The first phase of the project consists of 15 detector capsules, with all the electronics and data acquisition, pulse shape and tracking algorithms, and associated spectrometer infrastructure fully operational. In AGATA, the Ge crystals are packed together in groups of three in a common cryostat.

The AGATA collaboration has agreed upon an MoU that defines the planning, funding, construction, and operation of the device, as it is built up from a 15-detector system toward the full 4π device. As with GRETINA, AGATA will be sited at different laboratories, taking full advantage of the different beams and facilities available, in order to maximize the range of science that can be addressed.

A series of commissioning experiments, with sources and stable beams, have been performed using the AGATA at Legnaro National Laboratory (LNL) in Italy (Figure 7.8). An initial physics campaign began at Legnaro in early 2010, and the array is now at GSI in Germany for its second physics campaign, where it is being used at the end of the fragment recoil separator. In late 2013, it will be moved to the Grand Accelerateur National d'Ions Lourds (GANIL) laboratory in France to use the wide range of radioactive ions from the coupled cyclotrons and Systeme de Production d'Ions Radioactifs et d'Acceleration en Ligne, SPIRAL I/II.

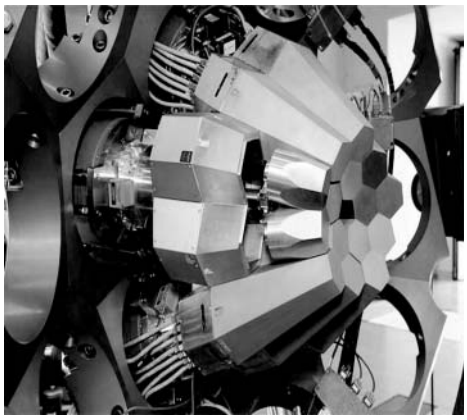


Figure 7.8 The AGATA demonstrator at Legnaro National Laboratory.

During these first three physics campaigns, the number of modules will increase toward 60 systems, which constitutes one-third of the full AGATA system.

For further information about the AGATA project, see *http://www-win.gsi.de/agata/*.

7.4.3 Gammasphere

Gammasphere was built by a consortium of US scientists from national laboratories and many universities. The project was coordinated by scientists at Lawrence Berkeley National Laboratory, where the device was first assembled. Gammasphere consists of a spherical shell of up to 110 large-volume, high-purity germanium detectors, each about the size of a coffee cup, and enclosed in a BGO escape-suppression shield (Figure 7.4). It was commissioned in December 1995, although experiments using parts of the system, the so-called early implementation phase, began in mid-1993. Gammasphere moved to Argonne National Laboratory in 1997 for a 2-year period before returning to LBNL in 2001 (Figure 7.9). Then, in 2003, it returned to Argonne again where it has remained. The array has a total photopeak efficiency of ∼9% at 1.3 MeV. As with other large Ge arrays, a suite of powerful auxiliary detectors has been specially developed to enhance the resolving abilities of Gammasphere.

Even now, after nearly two decades of discovery, Gammasphere is still an extraordinarily powerful array and for many experiments, such as high multiplicity fusion evaporation reactions, it is the best in the world. In addition, it has just undergone a major upgrade, which significantly improves its performance, by allowing an increase rate in the Ge detectors and a higher overall data throughput rate of ∼3–4. This has been achieved by replacing the aging analog electronics with a digital pulse processing data acquisition systems utilizing the GRETINA digitizers and trigger modules. Thus, Gammasphere, aided by synergistic developments of the GRETINA project, will continue to perform cutting-edge science for many years to come.

For further information, see *http://www.phy.anl.gov/gammasphere/index.html*, *http://nucalf.physics.fsu.edu/∼riley/gamma/* and *http://en.wikipedia.org/wiki/Gammasphere*.

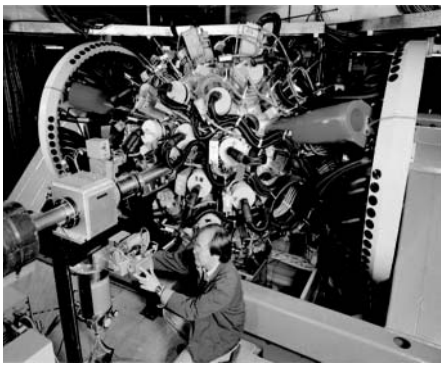


Figure 7.9 Gammasphere was initially at Lawrence Berkeley National Laboratory and is now at Argonne National Laboratory.

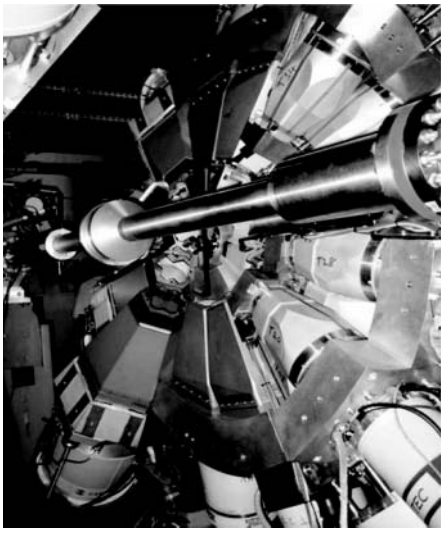


Figure 7.10 The EUROBALL detector.

7.4.4
Euroball, JUROGAM I and II, CLARA, RISING, GASP, and GALILEO

The European equivalent to Gammasphere was Euroball. This large project involved scientists from Denmark, France, Germany, Italy, Sweden, and the UK. In the development of EUROBALL, significant advances in detectors, electronics, and data acquisition were made (Figure 7.10). In particular, these included the revolutionary clover and cluster Ge designs, which were designed to increase the granularity of the array and thus minimize the double hit probability (Figure 7.11). In addition, the effective Ge efficiency was increased by adding the energies of γ rays that scattered between Ge crystals. Another useful benefit of the clover detector is its sensitivity to the linear polarization of γ rays, which comes from the different intensities of events Compton scattered parallel or perpendicular to the reaction plane. The cluster detector combined seven Ge detectors into a single cryostat and surrounded by a BGO escape-suppression shield of a hexagonal shape. For optimum long-term reliability, another technical innovation was that the Ge detectors

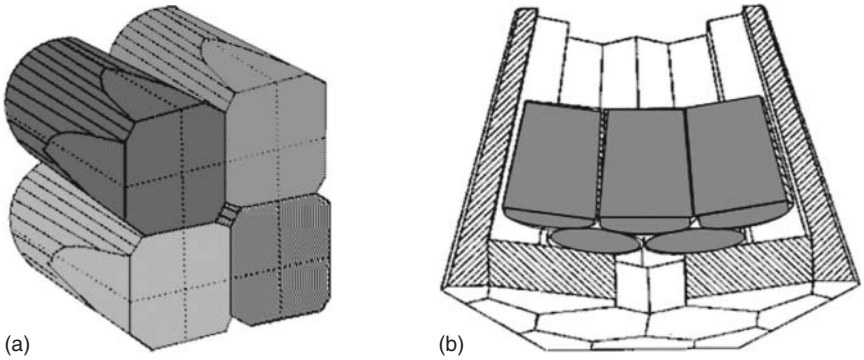


Figure 7.11 (a) Clover and (b) cluster Ge detector designs.

were hermetically sealed or encapsulated. Indeed, both GRETINA/GRETA and AGATA use encapsulated Ge detectors.

The Euroball array consisted of 30 large-volume germanium detectors (similar to Gammasphere), 26 clover germanium detectors, and 15 cluster germanium detectors each with an associated escape-suppression shield. This gave up to 239 individual Ge elements. The array had a total photopeak efficiency of ∼10% at 1.3 MeV.

The first phase of Euroball began in early 1997 at LNL, Italy. Then, in November 1998, Euroball was moved to the Institut de Recherches Subatomiques, Strasbourg, France, and upgraded to include an inner BGO array consisting of 370 BGO detectors. The ball and the Ge detectors measured the total number of γ rays emitted together with their total summed energy. Euroball IV commenced operations in June 1999 and ceased operations at the end of 2002. For further details on Euroball, see *http://eballwww.in2p3.fr/EB/*.

However, the detectors used in Euroball continue to perform world-class science and have been moved to a number of newer initiatives. These projects include rare isotope spectroscopic investigation at GSI (RISING), Jurogam I and now Jurogam II at Jyväskylä, and CLARA-PRISMA at Legnaro.

JUROGAM is an array of escape-suppressed HPGe detectors for use at the target position of the RITU gas-filled recoil separator at the Accelerator Laboratory of the University of Jyvaskyla (Figure 7.12). A variety of different detectors types have been used in assembling this array and include Eurogam clover, Eurogam Phase I, and GASP Ge detectors provided by the Euroball collaboration and the UK–France detector loan pool. The support structure and frame of the array are from the earlier implementations of Eurogam 1, which was situated at Daresbury Laboratory in the UK, and Eurogam II at IReS Strasbourg in France. The total peak efficiency for Jurogam 1 (43 escape-suppressed single Ge detectors) is 4.2%, and is 6% for Jurogam 11 (15 single and 24 clover escape-suppressed spectrometers), for 1.3 MeV γ rays.

For further information, see *https://www.jyu.fi/fysiikka/en/research/accelerator/nucspec/jurogam*.

The RISING array is built from Euroball cluster Ge detectors. This array has been operated in several configurations. A ''fast beam'' setup, as shown in Figure 7.13, is specifically designed to

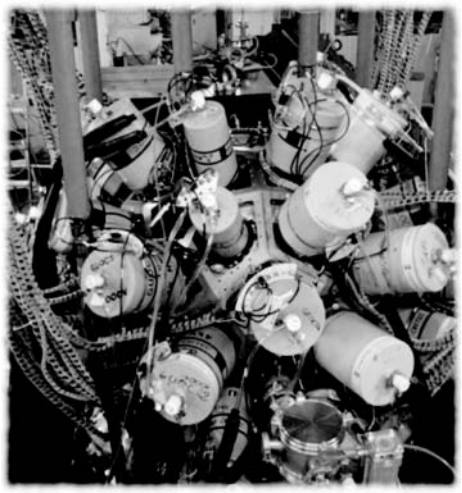


Figure 7.12 JUROGAM I at Jyvaskyla.

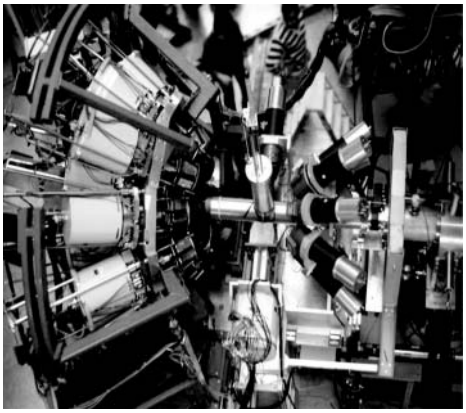


Figure 7.13 RISING at GSI.

maximize the efficiency following reactions with relativistic heavy ion beams. The Miniball triple cluster detectors were also added to this fast beam setup. A "stopped beam" setup was used with the seven clusters (105 Ge detectors) compactly arranged around a passive or active target to study the decay properties of exotic species selected by the FRS. A setup for *g*-factor measurements was also used.

For more information on RISING, see *http://www-aix.gsi.de/~wolle/EB_at_GSI/main.html*.

The CLARA array is composed of 25 clover detectors placed at the target position of the magnetic spectrometer PRISMA at Legnaro (Figure 7.14).

For further information, see *http://euroball.lnl.infn.it/EBcontents/EBweb_pdf/EBRW_10.pdf*.

Another European 4π array first developed in the mid-90's was GASP, which consisted of 40 escape-suppressed Ge detectors. This array ran successfully at LNL in Italy for many years. A new initiative using these detectors is the GALILEO array also at LNL. The GALILEO system

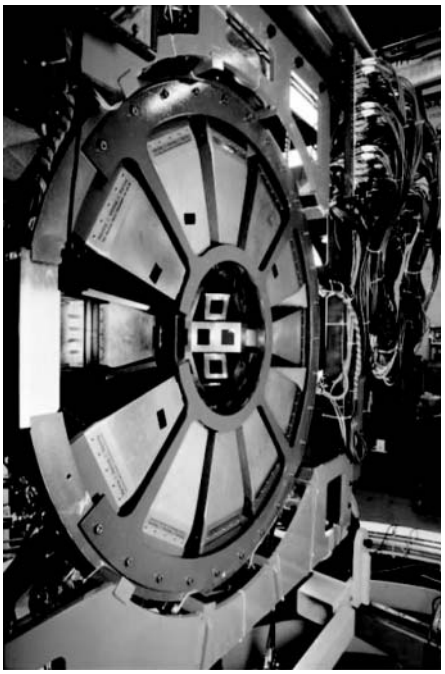


Figure 7.14 The CLARA-PRISMA array at Legnaro.

will combine the GASP tapered detectors and European Gammapool cluster detectors. The new array will take the original Euroball seven-element cluster detectors and put them into triple cluster detector modules. The latter will be placed at 90° with respect to the beam axis, and the GASP detectors will cover the forward and backward angles symmetrically. A large suite of auxiliary detector systems will be coupled to GALILEO.

7.4.5 MINIBALL

The MINIBALL array is designed for low-γ-ray multiplicity reactions using radioactive ion beams at the REX-ISOLDE facility at CERN (Figure 7.15). It consists of sixfold segmented, encapsulated Ge detectors. The high segmentation allows useful Doppler correction of γ rays emitted by fast-moving nuclei. The 40 Ge detectors are

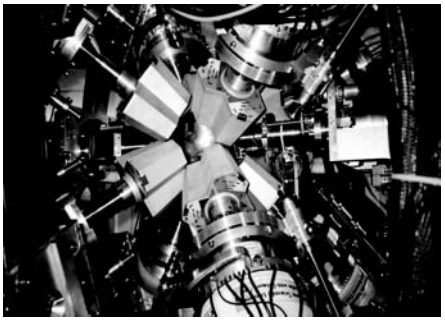


Figure 7.15 MINIBALL at REX-ISOLDE/HIE-ISOLDE.

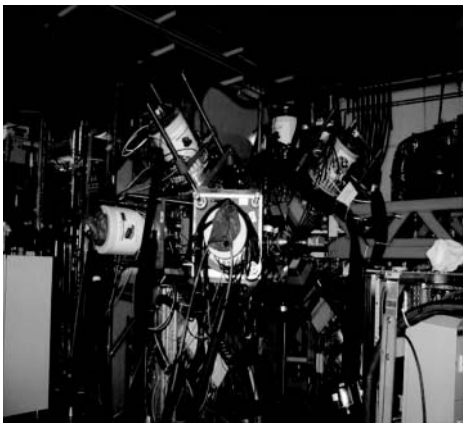


Figure 7.16 EXOGAM at GANIL.

arranged in eight cryostats with three detectors each and four with four detectors each. This arrangement provides optimal 4π Ge detector coverage, making an absolute efficiency of 14% for MINIBALL at 1.3 MeV.

For further information, see *http://cds.cern.ch/record/725972*.

7.4.6 EXOGAM

The EXOGAM spectrometer is a compact, flexible, high-efficiency, highly segmented, array of 16 clover-type γ-ray detectors. This array was designed to optimize spectroscopic investigations using the exotic beams from the SPIRAL facility at the GANIL along with a suite of auxiliary devices and the high efficiency recoil spectrometer VAMOS (Figure 7.16). For further information, see *http://www.ganil-spiral2.eu/science-us/detection-system/detectors*.

7.4.7 TIGRESS and GRIFFIN

The TRIUMF-ISAC gamma-ray escape-suppressed spectrometer or TIGRESS consists of up to sixteen 32-fold segmented Ge clover-type detectors and an escape-suppression shield for use with exotic beams provided by the ISAC-II superconducting linear accelerator at TRIUMF (Figure 7.17). Similarly to EXOGAM, this

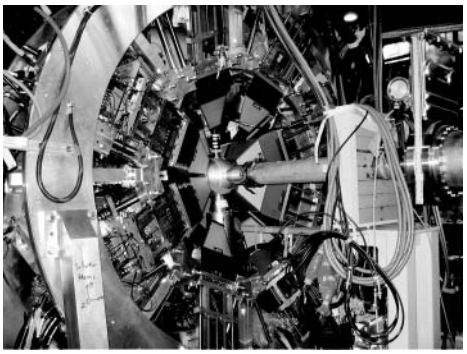


Figure 7.17 The TIGRESS array at TRIUMF.

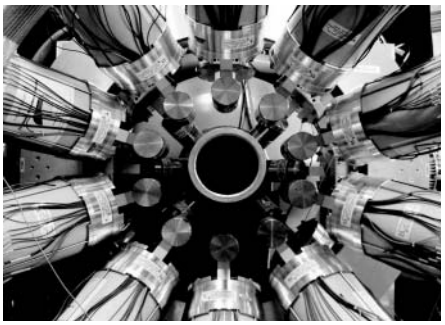


Figure 7.18 The SeGA array at the National Superconducting Cyclotron Laboratory.

array can be transformed into two modes; (i) high peak-to-total mode with the suppression shields surrounding the Ge clover modules and (ii) high-efficiency mode in which the Ge clover modules (without interclover suppressors) are pushed forward to only 11 cm from the target. A suite of auxiliary systems are available for charged particle and neutron detection. A customized fully digital electronics data acquisition system is employed. For more information on TIGRESS, see *http://www.physics.uoguelph.ca/Nucweb/tigress.html*.

A new and exciting initiative at TRIUMF is the gamma-ray infrastructure for fundamental investigations of nuclei (GRIFFIN) array for decay spectroscopy with low-energy radioactive ion beams from the ISAC accelerator. GRIFFIN will consist of 16 large clover Ge detectors arranged in a close packed geometry to optimize efficiency. Its total photopeak efficiency is ∼24% at 1 MeV. This means that GRIFFIN will be ∼17 times more efficient than the 8π spectrometer it will replace, and thus ∼300 times more efficient for $\gamma-\gamma$ coincidences. GRIFFIN is designed to be compatible with existing auxiliary detector systems of the 8π spectrometer. The system will use customized digital electronics. GRIFFIN will be at its early implementation phase in 2014 and is expected to be fully operational in 2015. For further information, see *http://www.physics.uoguelph.ca/Nucweb/griffin.html*.

7.4.8 SeGA

The Segmented germanium array (SeGA) is specially designed to perform high-resolution γ-ray spectroscopy of fast exotic beams from the NSCL at Michigan State University (Figure 7.18). The array consists of 18 Ge detectors (without escape-suppression) each electronically divided into 32 segments. The segmentation, as with the other discussed systems, allows the γ-ray hit information to be localized, which helps enormously in reducing the uncertainty in the Doppler correction due to the finite opening angle of the detector. The detectors can be placed at various distances to optimize between efficiency and resolution tradeoffs. The standard setup gives a total photo-peak efficiency of ∼3% at 1.3 MeV. For further information on SeGA, see *http://www.nscl.msu.edu/tech/devices/gammarayspectrometer*.

7.4.9 CLARION

The CLARION Ge array consists of 11 segmented clover detectors and is situated on the recoil mass separator beam line

Figure 7.19 The CLARION array at HRIBF at Oak Ridge National Laboratory.

of the Holfield Radioactive Isotope Beam Facility (HRIBF) at the Oak Ridge National Laboratory (Figure 7.19). The clovers can be placed a several distances away from the target, but at ~22 cm the array has a total photopeak efficiency of about ~2.4% at 1.33 MeV. For further information, see *http://www.phy.ornl.gov/hribf/research/equipment/clarion/*.

7.4.10 INGA

A large collaboration of Indian research institutions and universities has led to the building of the Indian National Gamma Array (INGA). It is composed of up to 24 escape-suppressed clover detectors and uses a full digital electronics data acquisition system. INGA has a total photo-peak efficiency of ~5% at 1.3 MeV. It is currently in use at the TIFR-BARC Pelletron-LINAC facility in Mumbai, but will rotate between the other accelerator facilities within India; see *http://www.tifr.res.in/~nsg/* for further information (Figure 7.20).

7.4.11 AFRODITE

The AFRODITE array is situated at iThemba LABS in South Africa and is

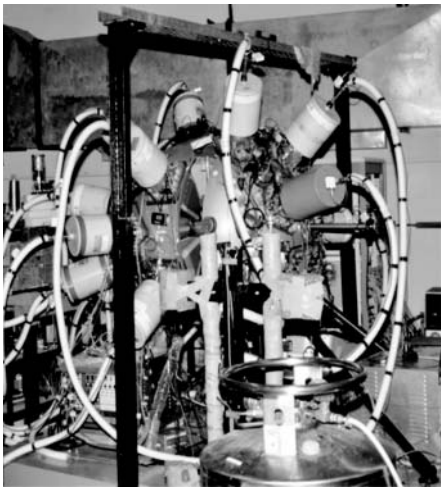


Figure 7.20 The INGA array at the TIFR-BARC Pelletron-LINAC facility.

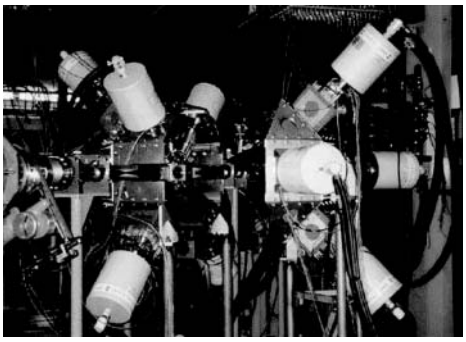


Figure 7.21 The AFRODITE array at iThemba-LABS.

Figure 7.22 The EURICA array at RIKEN.

composed of eight clover detectors of Euroball design and eight LEPS detectors (Figure 7.21). The latter provide exceptional efficiency for the detection of low-energy photons (<100 keV). A digital electronics system upgrade has recently been implemented. A number of auxiliary detectors are available; for further information, see *http://www.tlabs.ac.za/nucafrodite.htm*.

7.4.12 EURICA

The EURICA array at RIKEN (European Riken Cluster Array) in Japan currently comprises 12 cluster Ge detectors from the Euroball/RISING setup (Figure 7.22). An additional 21 $LaBr_3$ detectors, from the proposed future FATIMA array, will be used to complement the 12 cluster detectors with "fast-timing" coincidence capability from mid-2013. EURICA began operation in June 2012 and is scheduled to run decay and isomer spectroscopy experiments at RIKEN until the end of June 2014 using the projectile fragmentation and projectile fission secondary radioactive beams from the RIBF facility, see *http://www.riken.jp/engn/r-world/info/release/press/2012/120326_4/index.html*.

7.4.13
Other Ge-Based Arrays

A number of smaller or previous generation γ-ray arrays continue to perform first-class physics at many laboratories and institutions around the globe. It is not possible within this work to discuss all of these systems in detail, but examples include: the CAESAR array at the Australian National University, the FSU Gamma-Array at Florida State University, the STARS-LIBERACE system at Lawrence Berkeley National Laboratory, the GEANIE array at Los Alamos National Laboratory, the ex-Yrastball array from Yale University, which is now part of a mobile US clover pool of detectors, the OSIRIS and EAGLE arrays in Warsaw, the 8π array at TRIUMF, ORGAM at IPN Orsay, GABRIELA at Dubna, and in Japan, GRAPE, HyperBall 2, and GEMINI II.

7.5
Future Outlook

High-resolution γ-ray spectroscopy is one of the most powerful tools to study the structure of atomic nuclei and has seen many major advances in recent decades. It is a field that continues to reveal fascinating and often surprising new aspects of nuclei at the limits of isospin, excitation energy, angular momentum, temperature, and charge. The development of the next-generation γ-ray tracking arrays and their huge gains in sensitivity or resolving power, combined with new accelerator developments, assure a most exciting future for nuclear structure physics. One should also remember that the instrumentation and technical advances driven by this work, and the knowledge gained by those involved, are important in a wide range of applications. These advances impact areas such as medical imaging systems, homeland security, space exploration, and environmental monitoring.

References

1. Lee, I.Y. and Simpson, J. (2010) AGATA and GRETA: the future of gamma-ray spectroscopy. *Nucl. Phys. News Int.*, **20**, 23.
2. Lee, I.Y., Deleplanque, M.A., and Vetter, K. (2003) Developments in large gamma-ray detector arrays. *Rep. Prog. Phys.*, **66**, 1095.
3. Eberth, J. and Simpson, J. (2008) From Ge(Li) detectors to gamma-ray tracking arrays 50 years of gamma spectroscopy with germanium detectors. *Prog. Part. Nucl. Phys.*, **60**, 283.
4. Sharpey-Schafer, J.F. and Simpson, J. (1988) Escape suppressed spectrometer arrays: a revolution in γ-ray spectroscopy. *Prog. Part. Nucl. Phys.*, **21**, 293–400.
5. Nolan, P.J., Beck, F.A., and Fossan, D.B. (1994) Large arrays of escape-suppressed gamma-ray detectors. *Annu. Rev. Nucl. Part. Sci.*, **44**, 561–607.
6. Simpson, J. *et al.* (1994) Multiple band terminations in ^{158}Er. *Phys. Lett. B*, **327**, 187.
7. Paul, E.S. *et al.* (2007) The return of collective rotation in ^{157}Er and ^{158}Er at ultra-high spin. *Phys. Rev. Lett.*, **98**, 012501.
8. Wang, X. *et al.* (2011) Quadrupole moments of collective structures at ultrahigh spin in ^{157}Er and ^{158}Er: a challenge for understanding triaxiality in nuclei. *Phys. Lett. B*, **702**, 127.
9. Wang, X., Riley, M.A., Simpson, J., and Paul, E.S. (2013) Evolution of nuclear structure in erbium-158, 2013. McGraw-Hill Yearbook of Science & Technology, 119–123.

8
γ Optics and Nuclear Photonics

Dietrich Habs

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

8.1 Introduction

Until a few months ago, it was generally assumed that no refractive lenses for γ-rays exist and that they always go straight through matter. We showed in a recent publication in the *Physical Review Letters* [1] that this paradigm is no longer true. A similar situation occurred for X-rays, where C. W. Röntgen, after his discovery of X-rays in 1895, thought that X-rays were electromagnetic waves and searched for an index of refraction, but he could not find any deviation from 1 within his accuracies, and thus for many years, it was assumed that there is no refractive optics for X-rays also. However, in 1996, Snigirev *et al*. [2] showed that X-ray lenses could be built, using the rather small deviations from 1 of the index of refraction for X-rays.

We want to show in this section that it is very useful to compare the similarities and differences between X-rays and γ-rays more closely. In the recent years, the diagnostics with brilliant X-ray beams flourished in many fields of science [3], and we expect a similar development for γ beams. Our starting point will be to explain the new field of γ optics. It allows for a totally new high-resolution nuclear spectroscopy, but, even more interesting, many new facets of nuclear applications open up, which we call *nuclear photonics*.

The first ingredience was the unprecedented increase in brilliance of X-ray beams with time, which is shown in Figure 8.1. It was realized in the 1970s that synchrotron radiation from electron storage rings with insertion devices, such as wigglers and undulators, delivered much more brilliant X-ray beams compared to the standard X-ray tubes or those with rotating, cooled anodes. Then dedicated synchrotron light sources were built, such as the large storage rings ESRF (Grenoble, France), APS (Chicago, United States), and SPRING8 (Harima Science Park City, Hyogo Prefecture, Japan). Now the even more brilliant Linac Coherent Light Source (LCLS) with a free-electron laser (FEL) has started operation at Stanford, and in 2015, the European facility XFEL in Hamburg will follow up. Owing to relativistic effects, the range of X-rays has been limited to below 120 keV, and owing to strong absorption effects, the range has been limited to the lower energies to above 6 keV [4].

Similarly, we see a strong increase in the spectral intensity for γ beams in Figure 8.2. The Dynamitron [5] was a bremsstrahlung facility, where the electrons emitted a broad spectrum when being stopped in a cooled target. Then, the Compton backscattering sources, using brilliant monochromatic

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

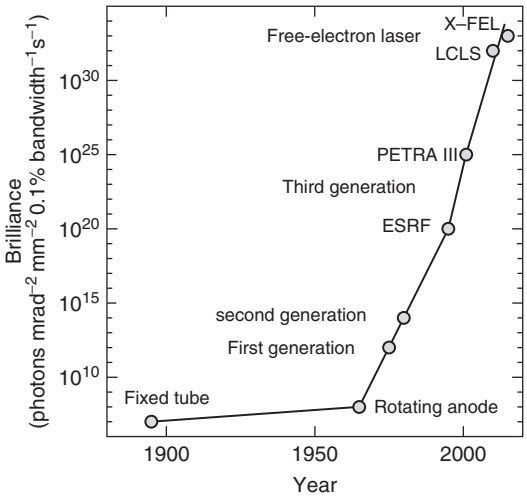


Figure 8.1 Brilliance of X-ray beams as a function of time.

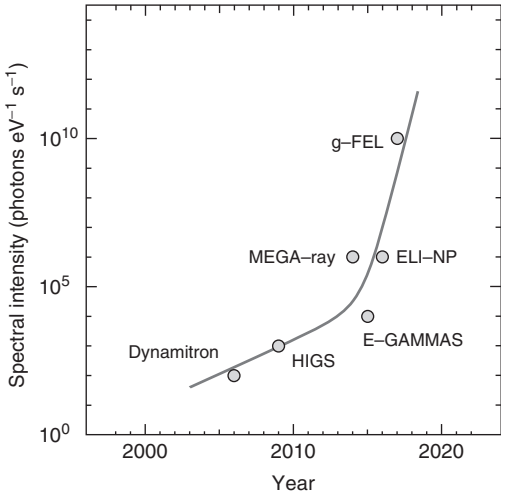


Figure 8.2 Spectral intensity of γ beams as a function of time. Following bremsstrahlung beams such as the Dynamitron [5], the different types of Compton backscattering sources are being developed. γ FELs may provide a large increase in spectral intensity.

electron beams with backscattering of a laser beam, were used to reach up to spectral intensities, where within a line width of 1 eV, the number of γ quanta per second is shown. At present, the HIGS facility [6] delivers worldwide the most brilliant γ beams. In Figure 8.2, also the values for the planned facilities E-GAMMAS [7], MEGa-Ray [8], and ELI-NP [9] are shown. Investigations are underway to realize, similar to the X-ray FEL, a much more brilliant γ-ray FEL inside the strongly focusing plasma channel after laser acceleration of electrons to gigaelectronvolt energies [10]. For γ beams, we characterize the beam by the spectral intensity and not by the brilliance, because for γ beams, there is frequently a strong correlation between the emission angle and the γ energy, as a larger recoil may be transferred to the electron. Thus, a narrow energy bandwidth leads to a small angular spread, and spectral intensity is the more adequate quantity to characterize

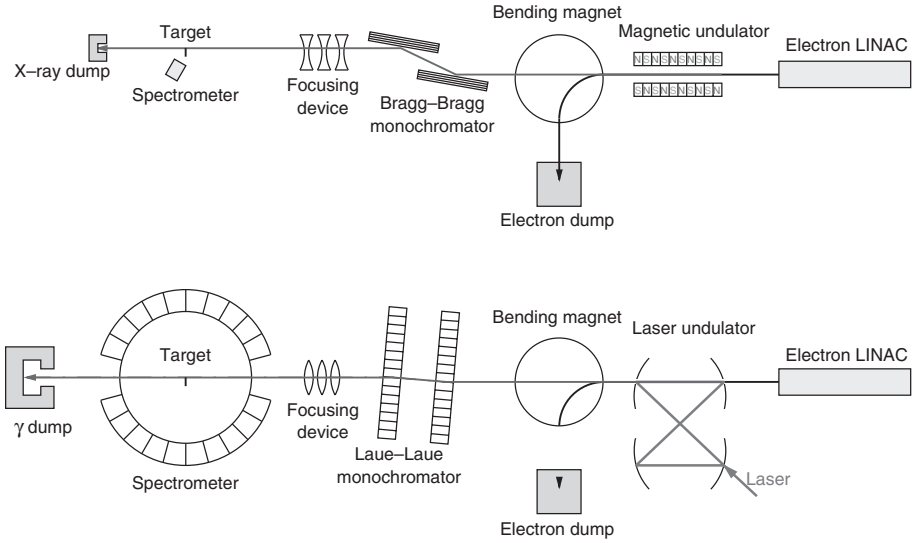


Figure 8.3 Comparison of the setups of a typical X-ray facility and a γ-beam facility.

a γ beam. Here we are mainly interested in a spectral energy of 0.5–20 MeV, where a strong Delbrück scattering should prevail.

Let us now compare two typical setups of an X-ray facility and a γ-beam facility (Figure 8.3). Both contain similar components, but frequently, they are realized in a different way. While magnetic undulators are used for the production of X-rays, we use laser undulators with four orders of magnitude smaller wavelengths for γ-rays. Both need very brilliant electron beams, which can either be produced by classical accelerators or by laser accelerators. The produced short-wavelength beams are then monochromatized and focused onto the target. Again different double-crystal monochromatization [11] is used: double Bragg crystals (crystal planes parallel to the surface plane) for X-rays and double Laue crystals (crystal planes normal to the surface plane) for γ-rays. As the refractive materials have an index of refraction smaller than 1 for X-rays, we require concave lenses for focusing. On the other hand, the materials for refractive optics of γ-rays have an index of refraction larger than 1, and hence, we use convex lenses for focusing. X-rays have a short range in detectors so that very small pixelated detectors can be used, whereas the stopping of γ-rays requires much larger detector arrays.

For X-rays, the second important step is the matching of the produced brilliant X-ray beams by X-ray optics to the experiments. Perfect double- crystal Bragg spectrometers using the Bragg reflection are used to monochromatize the beams. Nowadays, stacks of many concave X-ray lenses are used to focus X-ray beams down to micrometer or even nanometer range [3]. X-rays became a very sensitive probe to study the structure of matter, including the structure of DNA or much more complex protein molecules. With the focused beams, one can perform microscopy and tomography in physics, chemistry, geoscience, environmental science, biomedicine, and so on.

Similarly, all the components to match the γ beam to the experiment can be carried over, but frequently, little details are

different. In our γ optics, we use again perfect double-crystal spectrometers from silicon or germanium for monochromatization, but we use crystals cut in the Laue geometry [11]. First we expected that the deviation of the index of refraction from 1 decreases with $1/E_\gamma^2$, where E_γ is the γ-ray energy. This index would become that close to 1, that an extremely large number of γ lenses would be needed to achieve a reasonable focusing length, but then absorption in the lens material would become that strong, that it was assumed that no refractive lenses exist for γ beams. In contrast to this long prevailing paradigm, we recently discovered in a detailed measurement of the index of refraction for γ-rays [1] that contrary to theoretical estimates a much stronger deviation of the index of refraction from 1 occurred, even resulting in an index larger than 1, due to new strong-field interactions in the Delbrück scattering. Thus, convex lenses can be used to focus γ beams. While lens materials of low atomic numbers, Z, are preferred for X-ray lenses, the Delbrück effect increases strongly with Z. We will soon perform measurements with high-Z gold lenses and expect reasonable focal lengths. In our γ-ray experiments, we will not address atoms, but nuclei and their isotopes. γ-Rays can be used for diagnostics in very different fields of science or applications. Because γ-rays are much more penetrating than X-rays, we can noninvasively study isotopic distributions deep inside massive objects. As the relative nuclear level widths are much narrower than those for atomic levels, highly sensitive methods can be developed. Thus, totally new fields of nuclear spectroscopy and extremely rich fields of applications will open up.

As γ beams with very high spectral intensity are essential for all applications of γ beams, we want to explain the possible γ FEL more closely, where laser-driven electron bunches should develop ordered structures in the strong plasma-focusing fields. We plan to use these ordered beams for the coherent Compton backscattering of laser light to obtain a γ FEL.

In Ref. [10], we describe this new scheme in detail, aiming to reach spectral intensities of about 10^{10} $e^{-1}V^{-1}s^{-1}$ by producing an electron bunch with an internal crystalline structure, which can be used as a train of microbunches for a γ FEL. In Figure 8.4, a first laser is used to accelerate electrons to energies of 1 GeV by the so-called bubble acceleration mechanism [12–14]. Within the strong focusing fields of the plasma channel of several 100 GV m^{-1}, the relativistic electrons

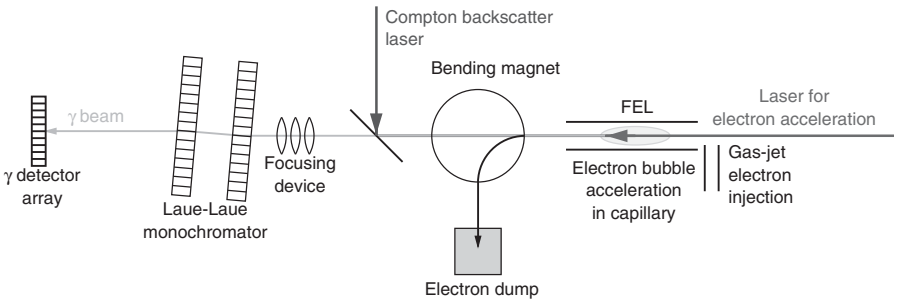


Figure 8.4 Using a laser-accelerated electron beam, which acquires an elongated crystal structure, via the coherent Compton backscattering at the right wavelength, a γ FEL can be realized.

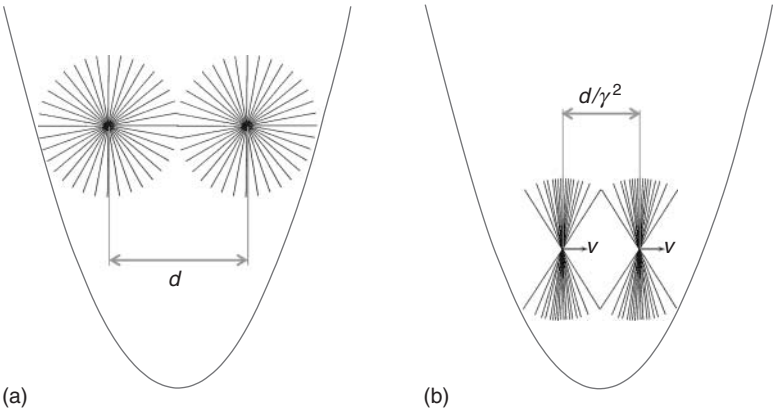


Figure 8.5 (a) If electrons at rest ($\gamma = 1$) are focused in a harmonic potential, the electric field distributions are spherical, while (b) for relativistic electrons ($\gamma \gg 1$), the field distributions are flattened and linear stringlike crystalline structures are favored, which then can be used to realize a γ FEL.

with a typical γ of 2000 reorganize themselves, because the field distributions of the electrons are relativistically flattened to pancake-like shapes, as shown in Figure 8.5. The electrons exhibit a smaller spacing by γ^2 along the beam direction, compared to a transverse arrangement of electrons. Approaching the center of the focusing fields of the laser-plasma bubble, a strong cooling of the electron bunch occurs because of the fast dipole transitions. If the wavelength of the Compton backscatter laser and the energy of the electron bunch are properly matched, an FEL-like coherent reflection from the crystal lattice occurs, and a much more intense γ beam is obtained.

Certainly, both components, γ beams with the highest spectral intensities and γ optics to match these beams in an optimum way to the experiment, are essential to open up new areas of nuclear photonics. After describing the new Grenoble measurements of the index of refraction, we explain the new high-field quantum electrodynamics (QED) concept based on the ultralarge nuclear electrical fields. Then we discuss the new toolbox of γ-ray optics. This opens up new worlds of nuclear spectroscopy and nuclear photonics.

8.2 The Grenoble Measurement of the Index of Refraction of γ-Rays

As described in Ref. [1], we performed the first measurements of the index of refraction $n = 1 + \delta$ for γ-rays up to energies of 2 MeV for silicon. Surprisingly, we observed a change in the sign of δ from negative values at lower energies to a positive δ above 0.7 MeV, which can only be explained by Delbrück scattering, that is, the scattering of γ quanta by the extremely strong electric field of a nucleus. An effect of the atomic electrons (such as the Rayleigh or the Compton effect) can be ruled out, because it would result in negative δ. Nuclear resonances of the stable Si nuclei ^{28,29,30}Si can be ruled out, since they have levels with very

narrow widths and excitation energies, far away from the used γ energies of 1.165 and 1.951 MeV. Thus, we regard the Delbrück scattering as the only explanation for this broad new structure in the 1–3 MeV range. Furthermore, we found that the forward-scattering amplitude at, for example, 1 MeV, was about a factor of 10^5 larger than that predicted by all the first-order Born approximation Delbrück calculations [15, 16]. It represents the first measurement of a scattering amplitude at exactly 0°, while all the former Delbrück measurements were performed at larger angles outside the primary beam. Although one knows that the static electric fields at the nucleus are typically thousandfold stronger than the so-called Schwinger field of $1.3 \times 10^{18}\mathrm{Vm}^{-1}$, and since it is also well known that in the regime of the Schwinger field perturbative QED breaks down, one has to perform much more difficult nonperturbative QED calculations. In this way, our measurement represents very likely the first accurate measurement of a high-field nonperturbative QED effect.

We performed the measurement at the ILL high-flux reactor in Grenoble. After neutron capture, we obtained very intense, monochromatic γ beams. We injected the beams into the double-crystal spectrometer GAMS5 [11], which allowed us to perform uniquely precise measurements of the index of refraction. Using the Bragg reflection, the first crystal of the spectrometer produced γ beam with an extremely small opening angle of about 100 nrad. This beam was then sent to a silicon prism, as shown in Figure 8.6. The upper part of the γ beam

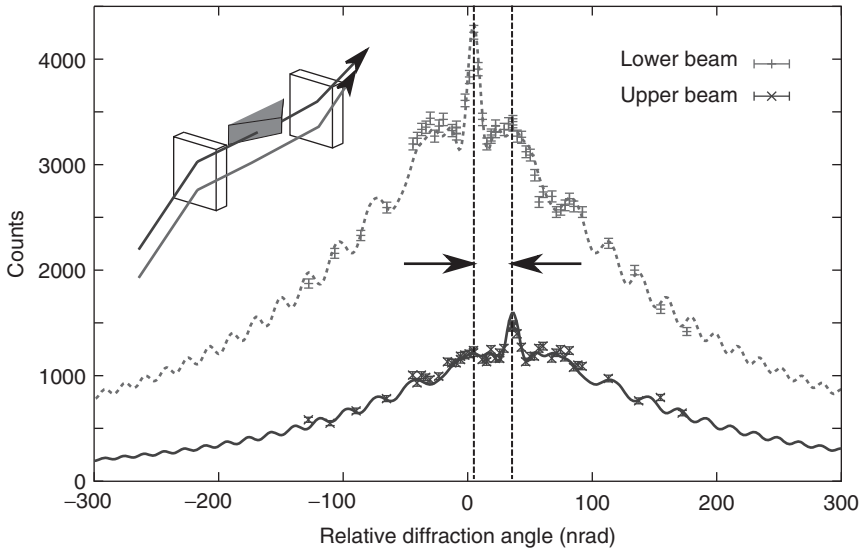


Figure 8.6 Illustration of the measurement principle. The γ beam is coming from the ILL high-flux reactor. A silicon wedge, placed between the two crystals of the GAMS5 spectrometer, deflects only the upper part of the beam, while the lower beam propagates in air. The two lineshapes were taken during preparation for illustration purposes with the 184 keV line of ^{168}Er with a long acquisition time. The upper beam shows a reduced intensity compared to the lower beam because of the absorption in the wedge. The spectra show "pendel solution"-like intensity oscillations typical of a double-crystal spectrometer [11]. One clearly sees the angular deflection by the wedge.

was deflected by the silicon prism, while the lower part of the beam was used as a reference beam. This technique to measure accurately an index of refraction with a double-crystal spectrometer and a prism is well established for X-rays [17] and is extended to γ-rays by us. In the spectra of Figure 8.6, the quantum mechanical interference pattern, called *pendel solutions* [18], is clearly seen. It is due to the standing γ waves between the crystal planes, similar to Fabry–Perot interferences. By comparing the two beams, one through the prism and one below the prism, one can determine accurately the deflection angle from which one can calculate the δ of the index of refraction $n = 1 + \delta$ of the prism. In Figure 8.7, we show the measured δ as a function of γ energy. For lower energies, we see the well-known negative value of δ. Above 0.7 MeV, we observe a change in the sign of δ, which can only be explained by the Delbrück scattering. However, even more important is the fact that the forward-scattering amplitude is about a factor of 10^5 larger than that predicted by perturbative QED. It is interesting to calculate also the real forward-scattering length $L_D(E_\gamma)$ via

$$\delta(E_\gamma) = \frac{\lambda_\gamma^2}{2\pi} \cdot N_c \cdot L_D(E_\gamma) \tag{8.1}$$

where N_c is the number of nuclear scattering centers per volume and λ_γ is the wavelength of the γ quanta. To obtain a constant $\delta(E_\gamma)$, it requires a scattering amplitude $L_D(E_\gamma)$ rising with E_γ^2 or by a factor 2.8 between 1.17 and 1.95 MeV. The measured δ values at 1.165 MeV with 1.48×10^{-9} and at 1.951 MeV with 1.11×10^{-9} correspond to an increase in the Delbrück scattering length L_D by a reduced factor of only 2.2. It has the highest efficiency and points to the Delbrück scattering. While the first-order Delbrück scattering leads to a very small efficiency, the next higher order in $(\alpha Z)^4$ results in a much larger cross section. Finally, the scattering length will be limited by the screening length of the nuclear electric field. The Delbrück scattering length should then show a strong

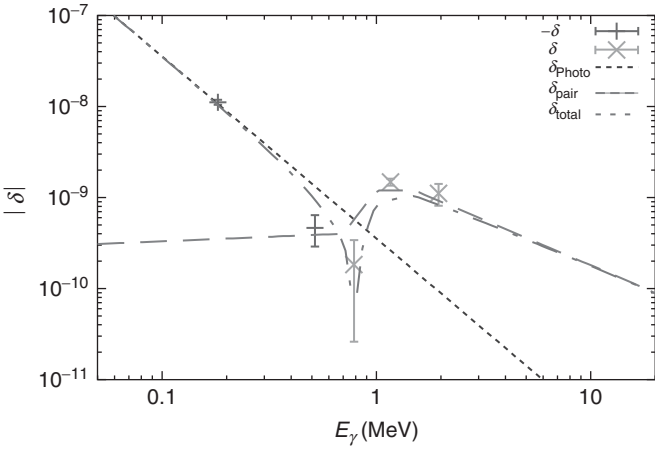


Figure 8.7 Newly measured index of refraction $|\delta|$ for γ energies up to 2 MeV. The dotted curve shows the negative δ from the virtual photo effect (the Rayleigh scattering), which is confirmed by the measured values for lower γ energies. For the positive δ of virtual pair creation (the Delbrück scattering), we used a shape from our dispersion relation calculations. The interference between the two δ contributions is also shown.

decrease for higher γ energies, because in the gigaelectronvolt range, the theoretical values of first-order perturbation theory have been confirmed experimentally [19]. Hence, we expect that there is some energy region beyond 10 MeV, where δ becomes that small, that γ optics due to absorption in the lenses will no longer be feasible.

In the present measurement, for the first time at zero degree scattering angle, the Delbrück scattering length L_D was determined, whereas all the previous measurements had been performed at much larger angles to avoid the large background from the noninteracting primary γ beam [16, 20]. Only for the zero degree scattering amplitude the Kramers–Kronig dispersion relations allow for a calculation of the real scattering amplitude from the imaginary part, the absorption cross sections. The dispersion relations are a powerful tool in nonperturbative relativistic quantum field theory [4, 21, 22]. We find that higher order contributions, such as the fourth or sixth order, give dominant contributions. An extrapolation from the zero degree scattering amplitude to larger scattering angles is difficult, because it depends on coherence properties of the γ beam with destructive interferences and on the screening of the Coulomb potential [23].

We are preparing a new experiment, where we replace the double Laue–Laue monochromator and the prism from perfect silicon by perfect germanium crystals at the crystal laboratory ESRF (Grenoble). It this way, we can study the Z-dependence of the Delbrück scattering from silicon ($Z = 14$) to germanium ($Z = 32$). In an accurate measurement with the gamma lines after neutron capture of ^{35}Cl, we will study for these perfect crystals and wedge for Si and Ge the zero degree scattering amplitude in a comparison with and without the wedge at the same time at ILL (Grenoble). We expect at least a Z^3 increase in the Delbrück scattering, and thus with the second Z, we will learn this scaling law.

The unexpected large δ values can be explained by nonperturbative high-field QED. Thus the perturbative first-order Born calculation shown in Figure 8.8a with the Feynman box diagram has to be replaced by the fully dressed electron–positron loop in Figure 8.8b. Presently, several theoretical groups are trying to perform such highly sophisticated, nonperturbative QED calculations. With dispersion relations, it could be shown [1] that the next higher orders in (αZ) are much larger than the first order, contradicting the idea of a perturbative expansion.

The new process of the nonperturbative Delbrück scattering results in a large forward-scattering amplitude for γ-ray energies, which is important not only for refractive optics but also for the Bragg reflection:

$$n \cdot \lambda_\gamma = 2 \cdot d \cdot \sin\theta \tag{8.2}$$

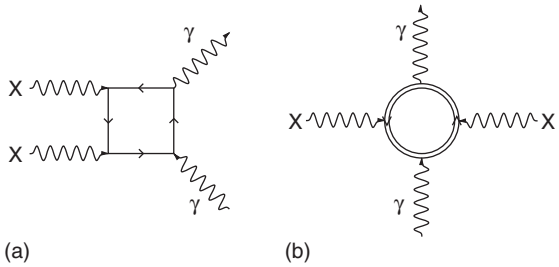


Figure 8.8 The Delbrück scattering of a γ from the Coulomb field: (a) first-order Feynman box diagram and (b) the interaction via a dressed electron–positron loop for high-field QED.

with the lattice spacing d and the reflecting angle θ. Here, both the Rayleigh scattering and the Delbrück scattering contribute, but the larger Delbrück scattering amplitude results in a correspondingly larger Darwin's width, thus opening possibilities for more efficient monochromatization. As the recoil momentum is taken up by the crystal lattice and the heavy nucleus for the Delbrück scattering and not by the electron like in Rayleigh or Compton scattering, a coherent scattering is realized at higher γ energies and for larger scattering angles.

8.3 Nonperturbative High-Field Quantum Electro Dynamics (QED)

Now we will outline the surprising new picture of nonperturbative QED in the ultrahigh electrical fields of the nucleus, where the vacuum is strongly polarized into a virtual plasma of e^+ e^- pairs. This polarization is even enhanced by the probing γ fields. As the scattering of light by light is extremely weak, the dominant cross section proceeds via pair creation. While the Delbrück scattering shown in Figure 8.9 has been studied theoretically in detail for the two cases where the γ energy $E_\gamma \ll mc^2$ [24] and $E_\gamma \gg mc^2$ [16, 25, 26, 27] for higher order contributions in higher orders of $(Z\alpha)$, the so-called Coulomb corrections, this could not be performed for the case of interest here, where the γ energy E_γ is comparable to mc^2. Furthermore, in the two cases, γ energy $E_\gamma \ll mc^2$ and $E_\gamma \gg mc^2$, a quasiclassical approximation supposing large transferred angular momenta $l \gg (\alpha A)$ could be assumed. However, here, with $E_\gamma \simeq mc^2$, the transferred angular momentum $l \ll 1$ and a perturbative approach is invalid. Thus, the resonant, really interesting case of the Delbrück scattering with $E_\gamma \simeq mc^2$ has to be studied, while until now only the far-off resonance cases with γ energy $E_\gamma \ll mc^2$ and $E_\gamma \gg mc^2$ were studied.

The left side of Figure 8.9 shows the first-order Born description of the Delbrück scattering, while the right-hand side shows the fully dressed electron–positron loop in all orders or nonperturbatively. After we had observed this strong enhancement of the Delbrück scattering process, it became clear that similar processes, shown in Figure 8.9, such as splitting of a γ-ray into two γ photons in the strong Coulomb field (photon splitting) or pair creation close to the threshold will show similar nonperturbative QED enhancements. Thus, it will become possible to study this new field of QED in detail, providing a more general understanding of these polarization effects in the strong nuclear electric fields. For photon splitting, one field line from the Coulomb field is replaced by one de-exciting γ line. While for a truly constant electric field the box diagram of photon splitting vanishes [28] and only the hexagonal diagram with three interactions with the field gives the first contribution, for nonconstant electrical fields (which are of interest here), the box diagram is the first important contribution. For pair creation, the opening of the electron loop into an electron–positron pair occurs. While the box diagram on the left side describes the processes in the first-order perturbation Feynman theory, the double-circled loop on the right side represents the electron–positron loop fully dressed by the strong nuclear electric Coulomb field. These three processes are the dominant nonperturbative high-field QED processes using the ultrahigh electric nuclear fields.

The reference field for high-field QED is the Schwinger field $E_s = m^2/e =$

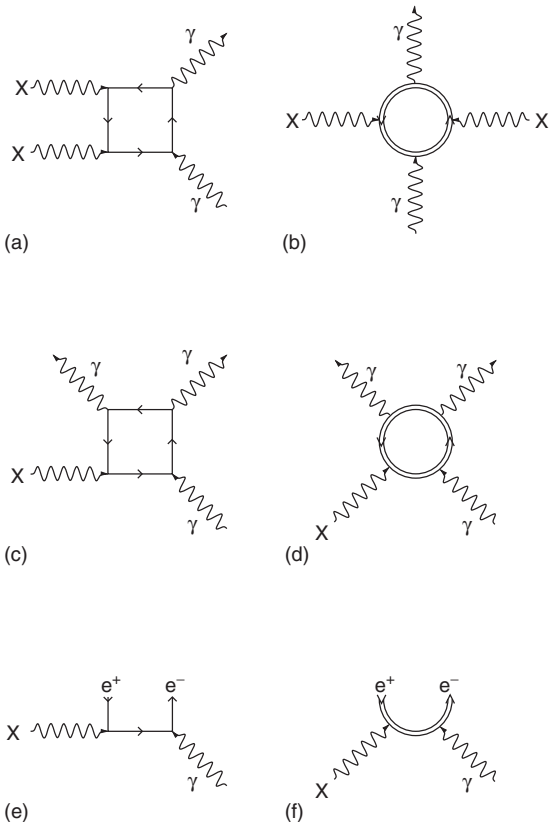


Figure 8.9 The γ (a) Delbrück scattering in perturbative QED, (b) Delbruck scattering in nonperturbative QED, (c) photon splitting in perturbative QED, (d) photon splitting in nonperturbative QED, (e) pair creation close to threshold in perturbative QED, and (f) pair creation close to threshold in nonperturbative high-field QED.

1.3×10^{18}V m^{-1}, where a virtual electron–positron pair is materialized in the strong electric field E over its Compton wavelength of about 400 fm. For a constant electric field E, one obtains the Sauter–Schwinger effect [29–31], where the number of pairs, $P_{e^+e^-}$, per unit time and unit volume is given by

$$P_{e^+e^-} \sim \frac{e^2E^2}{4\pi^3} \exp\left[-\pi\frac{m^2}{eE}\right] \qquad (8.3)$$

with $\hbar = c = \varepsilon_0 = 1$. This is a pure nonperturbative result, as $P_{e^+e^-}$ has no perturbative expansion in the coupling strength e. Thus, even the Feynman diagrams of an arbitrarily large order cannot describe this physics. This is a pure theoretical prediction that lacks experimental verification. Because we probe our high electric fields with γ-rays, it is also important to study the predictions for a constant electric field assisted in pair creation by a γ quantum of energy E_γ. For this pair production per unit volume $P_{e^+e^-}$, we find [32]:

$$P_{e^+e^-} \sim \frac{(eEE_\gamma)^2}{4\pi^3m^2} \exp\left(-\frac{8}{3}\frac{m^3}{|e|E\cdot E_\gamma}\right) \qquad (8.4)$$

The electric field E is increased by the relativistic boost factor E_γ/m. For very large E_γ values, the exponential tunneling factor gets close to 1, if the original electric field is already close to the Schwinger field E_s. Then we get back to a situation that can be treated with perturbative QED. Again,

these theoretical predictions have not been verified experimentally.

However, in our case, we are treating very strong Coulomb fields Ze/r^2, which have a strong radial dependence and not a constant electric field. Furthermore, these nuclear Coulomb fields have a typical field strength at the nuclear radius that exceeds the Schwinger field by about 10^3. Still the equations for a constant electric field can give useful expectations on what will happen in the strong nuclear electric field. If we increase the γ energy, the outer regions of the nuclear electric field will be boosted even closer to the Schwinger field, and nonperturbative QED becomes more important. However, for γ energies in the gigaelectronvolt range, the effective electric fields become that large, that the tunneling factor becomes close to 1; the Delbrück scattering and photon splitting can be described again in a quasiclassical perturbation theory in agreement with experimental findings [19, 20]. For the electron, the dressed propagator in such a strong Coulomb field has been calculated [26, 33], but until now, these multidimensional integral representations could not be transformed into useful approximations for our γ energy interactions in the few megaelectronvolt range.

Thus, a semiclassical rough picture may be helpful: squaring the field strength in the range of the nucleus provides the electrical energy density, which now even drops off with $1/r^4$ and thus is even more confined to the nuclear radius. This energy density must be correlated with the number of virtual e^+e^- pairs $N_{e^+e^-}$. This value $N_{e^+e^-}$ should scale with Z^3, as the nuclear radius scales with the mass number A via $A^{1/3}$ and A is roughly proportional to the nuclear charge number Z. Thus, we have a new scale length in our theory: the diameter D of the nucleus, with $D \ll \lambda_\gamma$. Also for the impact parameter ρ of the γ-ray, we have $\rho \leq D$. If one assumes that the virtual e^+e^- pair in a box with diameter $D \approx 10$–20 fm has an energy of about 10 MeV, for a nucleus with $Z = 14$, we obtain $N_{e^+e^-} \approx 10^3$ by dividing the total field energy by the energy of a single pair. It is interesting to note that the total electrical field energy represents a significant part of the total energy of the protons. From $N_{e^+e^-}$, we can get an estimate of the forward Delbrück scattering amplitude by treating it similar to the Rayleigh scattering, multiplying the classical electron radius r_0 with the number of leptons 2×10^3. If one compares this value with the scattering amplitude from the first-order perturbation theory of the Delbrück scattering [15] of $\sim 0.1(Z{\cdot}\alpha)^2 * r_0 \sim 10^{-3} r_0$, one obtains an enhancement factor of 10^6, which is in the right order of magnitude compared to our experimental finding of 4×10^5. Owing to the smallness of the nuclear radius, we can replace the virtual e^+e^- plasma by a close to point-charge distribution with $N_{e^+e^-}$ pairs. Once the Delbrück scattering length L_D becomes larger than the Thomas–Fermi screening length $l_{TF} = 1/(m\alpha(Z)^{1/3})$, the influence of the high electrical field is strongly reduced and limits L_D.

8.3.1 Photon Splitting

As shown in Figure 8.10, we want to measure the photon splitting of a γ beam by splitting, for example, a 1.5 MeV γ beam in a lead foil. While the unsplit photons will produce a large single- and double-escape peaks of 511 keV quanta from the pair annihilation in the Germanium detector, for the split γ quanta, which have an energy below the pair production threshold, no single- and double-escape

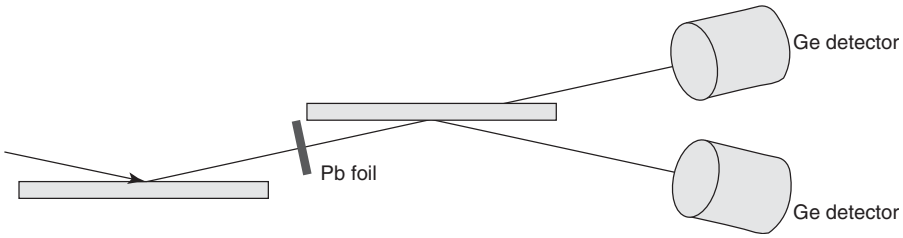


Figure 8.10 Experimental setup for photon splitting using a γ beam.

peaks should be observed. The photon splitting cross section is certainly reduced compared to the Delbrück scattering by a factor of α^2 due to the production of the second emitted photon, and also due to the reduced phase-space factor when emitting two photons. We expect cross sections that are somewhat smaller than the Compton cross section and should be easy to measure in detail. In later measurements, the entanglement of the two photons may also be of interest. For the Delbrück scattering and photon splitting, we have no threshold, but similar to the real pair production, here also the virtual pair production due to nonperturbative QED effects will be reduced for higher E_γ energies. Certainly, these expectations have to be verified by ongoing complex nonperturbative QED calculations (Ruhl, H. (2012) Private communication.).

8.3.2
Pair Creation Close to Threshold

The third important nonperturbative QED process is the pair creation close to threshold. Here, we expect strong deviations from the Bethe–Heitler formula because the Schwinger-like tunneling will give strong contributions. The first measurements show stronger deviations close to threshold [34], but we are now extending these measurements to a larger value of $Z = 47$ available for $LaBr_3$ detectors, and we will also measure closer to threshold.

Establishing the new high-field QED will also be helpful in understanding the nonperturbative QED much better. It will also be helpful to select the optimum material for the new γ optics.

8.3.3
Dispersion Relations

Dispersion relations relate the real amplitude $A_{\mathrm{rf}}(E_\gamma)$ and the imaginary amplitude $A_{\mathrm{if}}(E_\gamma)$:

$$A_{\mathrm{rf}}(E_\gamma) = \frac{E_\gamma^2}{\pi(4\pi^2\hbar c)^2} \lim_{\varepsilon\to 0+} \times \int_0^\infty \frac{A_{\mathrm{if}}(E)dE}{E(E^2-(E_\gamma+i\varepsilon)^2)} \quad (8.5)$$

Via the optical theorem [21, 22], the imaginary part of the forward-scattering amplitude is related to the total absorption cross section, $\sigma_{\mathrm{abs}}(E_\gamma) = 2{\cdot}\lambda{\cdot}A_{\mathrm{if}}(E_\gamma)$, or $A_{\mathrm{if}}(E) = (E/4\pi\hbar c)\cdot\sigma_{\mathrm{abs}}(E)$. We thus obtain

$$A_{\mathrm{rf}}(E_\gamma) = \frac{E_\gamma^2}{2\pi^2\hbar c} \lim_{\varepsilon\to 0+} \times \int_0^\infty \frac{\sigma_{\mathrm{abs}}(E)\mathrm{d}E}{E^2-(E_\gamma+i\varepsilon)^2} \quad (8.6)$$

This relation connects the total absorption cross section $\sigma_{\mathrm{abs}}(E)$ with the forward coherent scattering cross section

$$\frac{\mathrm{d}\sigma_{\mathrm{sca}}}{\mathrm{d}\Omega}(\mathrm{forward}) = |A_{\mathrm{rf}}(E_\gamma)|^2 \quad (8.7)$$

$A_{rf}(E_\gamma)$ is related to the real part of the index of refraction $\delta(E_\gamma)$.

As we measure a forward-scattering amplitude, which is a factor of 10^5 larger than that predicted by the first-order Born approximation, the dispersion relations require correspondingly larger absorption processes. One has to realize that a large constant cross section (like the pair creation process for higher energies) does not give a significant contribution to the real scattering amplitude, because the contribution below the pole is canceled by a similarly strong contribution above the pole. All the three newly discussed high-field processes will contribute with new absorptive cross sections, which have been difficult to observe until now. In comparison to the Compton scattering due to the $A \times 2000$ smaller recoil energy, the inelastic Delbrück scattering is difficult to observe. It is difficult to predict the Delbrück scattering amplitude at larger angles from the measured scattering at 0° [23]. However, here a bootstrap effect will occur – the larger the zero degree amplitude, probably, the larger the inelastic Delbrück scattering will be. The photon splitting is always an absorptive process, but it will be at least a factor of α^2 weaker due to the emission of the second γ-ray and also the phase-space factor leads to a reduction. Thus it will be a factor of about 10^5 smaller than the Delbrück cross section and will have a similar energy dependence like the Delbrück scattering. In competition with pair creation, it would have been difficult to detect the photon splitting in the process of stopping γ quanta. Also the enhanced pair creation close to threshold leads to an enhanced absorption with a different sloping cross section, resulting in changed predictions of the real forward-scattering amplitude via the dispersion relations, as shown in Figure 8.9.

8.4 The New γ Optics

The new large Delbrück scattering amplitudes have two applications: refractive optics and the Bragg refraction for monochromatization.

8.4.1 Refractive Optics

In our rough estimates of the Z scaling of the index of refraction via the Delbrück scattering, we expected a Z^3 dependence and estimated for gold, with $Z = 79$, a much larger index of refraction of about $\delta = 10^{-5}$, where the screening length will probably limit the Delbrück scattering length. In September 2012, we will perform the first experiment at the synchrotron light source ESRF with gold lenses (Figure 8.11) for γ energies of 0.5–0.8 MeV. If our expectation of a rather large index of refraction is fulfilled, a new area of γ optics will open up. Knowing the curvature radius of our convex gold lenses of 200 μm, we can calculate the deviation of the index of refraction δ from the measured focal length. Gold was chosen as the lens material because of its high atomic number Z, while still being easy to shape. The forward scattering will be limited by the screening length with

$$r_{thomasfermi} = \frac{1}{m\alpha \cdot Z^{1/3}} \tag{8.8}$$

Once this optics with convex lenses has been established, other elements such as prisms or γ waveguides will be explored.

8.4.2 Monochromators

As Delbrück scattering corresponds to Rayleigh scattering from electrons of the Dirac sea, the angular distribution of the

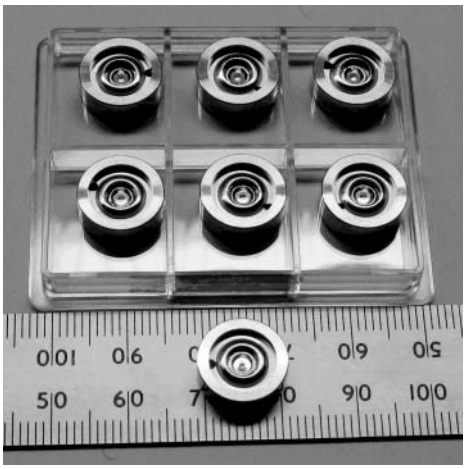


Figure 8.11 Convex parabolic γ lenses made from gold with 1 mm diameter for focusing γ-rays with an expected focal length of 10 m per lens for $\delta = 10^{-5}$. Only the innermost bulge of 1 mm represents the lens.

scattered γ's should be similar under small forward angles. Our aim is to develop the perfect crystals of germanium ($Z = 32$) in addition to those of silicon ($Z = 14$). Even though this technology was developed many years ago [35], the growing process and fabrication of perfect germanium crystals had to be reinvented. We will perform new perfect germanium development at the crystal laboratory in the ESRF (Grenoble). The theory of perfect crystal diffraction is called the *dynamic theory of diffraction*. It is studied in detailed monographs [36, 37, 38]. The advantage of the Delbrück scattering is that is has much larger forward-scattering amplitude and thus much larger angular Darwin width, thus reaching much higher efficiencies and much more flexibility in adjusting to a required bandwidth of the γ beam. If we want to have a high efficiency at a very narrow bandwidth, we can select the same bandwidth from different angular regions as described in Ref. [39]. This possibility of building much more efficient monochromators with double-crystal spectrometers is an essential feature of the new γ optics.

8.5 New Nuclear Spectroscopy

A typical present day spectrum of nuclear electromagnetic excitations is shown in Figure 8.12 with E1 and M1 excitations. Here we find the so-called scissors mode (SM) [40, 41], the two-phonon mode, where an octupole E3 excitation is coupled to an E2 quadrupole mode resulting in the quadrupole–octupole (QOC) mode [42], which decays by E1 transitions. At higher excitation energies close to the neutron binding energy, we see the pygmy dipole resonance (PDR) mode [43, 44] with oscillations of the neutron skin against the nuclear core. Then, at even higher excitation energies, one finds the giant dipole resonance (GDR) [45], where all neutrons oscillate against all protons.

This picture can now be drastically refined with the new high-resolution γ beams because we can now excite individual levels with higher multipolarity, while presently these excitations are buried under the much stronger E1 and M1 excitations. Thus, we expect the much richer spectrum shown in Figure 8.13 with E2, E3, and E4 excitations. Again

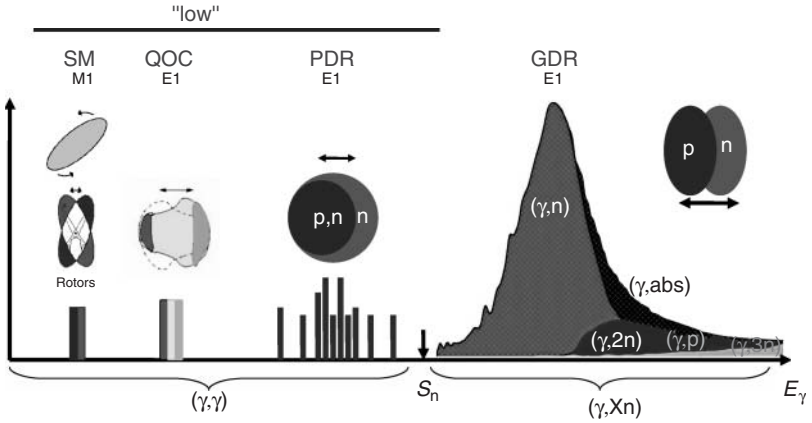


Figure 8.12 Nuclear M1 and E1 excitations. SM, scissors mode; QOC, quadrupole–octupole mode; PDR, pygmy dipole resonance; GDR, giant dipole resonance.

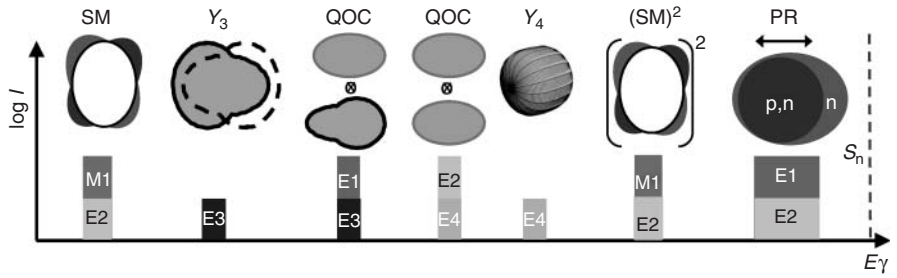


Figure 8.13 New nuclear spectroscopy including higher multipole modes E2, E3, and E4 becoming visible with intense, narrow bandwidth excitation.

we have shown the excitations roughly at their expected excitation energies. Thus, we will resolve the Y_3 octupole excitations. We expect not only the QOC two-phonon excitation but also the quadrupole excitation coupled to itself, that is, the two-phonon quadrupole mode (QQC). With an E4 excitation, we expect to see the hexadecapole oscillation Y_4. We also expect the double excitation of two SMs, the $(SM)^2$ mode. The PDR mode not only exhibits E1 excitations but also E2 excitations. Thus, a much more detailed, high-resolution spectroscopy becomes possible. As we can resolve all levels and determine their spins, a complete spectroscopy for higher spins and different parities becomes possible. This is of high interest to follow the transition from regular nuclear motion at low energies to chaotic motion at higher excitation energies [46].

Another interesting field is the study of parity violation in nuclei. Because the parity nonconservation (PNC) amplitude is described in perturbation theory by $\langle|\mathrm{PNC}|\rangle/(E_\pi - E_{-\pi})$ [47] and varies inversely proportional to the energy difference $E_\pi - E_{-\pi}$, small admixtures can be measured much more sensitively for small spacings of levels of opposite parity.

Thus, with better monochromatized beams, we want to study systematically, for example, 1^+ and 1^- doublets, at higher excitation energies and for nuclei with larger mass number A. Compared to neutron capture experiments [48– 50], where large parity violation amplitudes with $\langle|\mathrm{PNC}|\rangle \approx 10^{-3}$eV were observed, we can now study the states much more systematically below the neutron binding energy. Several different parity violating amplitudes such as πNN or ρNN can be studied [51–53]. Owing to the small bandwidth, the excitation will be followed by rather simple de-excitation cascades, which can be studied, for example, by an array of large $LaBr_3$ detectors.

8.6 Nuclear Photonics

In the discussion of possible γ-beam applications, we first consider the cross section of photonuclear reactions and the corresponding width Γ_γ. The cross section for a compound nuclear resonance populated by photoexcitation at the resonance energy E_r in the region below the neutron separation energy S_n is given by the Breit–Wigner formula [54]

$$\sigma(E_\gamma) = \left(\frac{\lambda_\gamma^2}{4\pi}\right) \cdot g \cdot \frac{\Gamma_\gamma \Gamma_2}{(E_\gamma - E_r)^2 + \frac{1}{4}(\Gamma)^2} \tag{8.9}$$

where $g = (2I_a + 1)/(2I_b + 1)$ is a spin factor for the spin of the target and the beam and $\lambda_\gamma = (h{\cdot}c)/E_\gamma$ represents the wavelength of the γ-rays with energy E_γ. The resonance is excited with the width Γ_γ and decays via the width Γ_2. The resonance has a total width $\Gamma = \Gamma_\gamma + \Gamma_2$. Figure 8.12 shows the dominant E1 and M1 nuclear excitation widths Γ_γ.

The typical maximum E1 width for a level up to the neutron separation energy S_n is about 10 milli-Weisskopf units. It corresponds for 2 MeV photons to about 0.1 eV. On the other hand, the Doppler broadening of a γ transition at room temperature $kT = 1/40$ eV for a nucleus with mass number $A = 200$ and a γ energy $E_\gamma = 2$ MeV is

$$\Delta E_\gamma = E_\gamma \sqrt{\frac{(2kT)}{m_p c^2 A}} \approx 1 \text{ eV} \tag{8.10}$$

It shows that we find a typical relative line width of 10^{-6} for a nucleus with mass number $A = 50$, or 1 eV at 1 MeV and 10 eV at 10 MeV, respectively. Thus, with a γ beam of 10^{-6} bandwidth, we can selectively address individual nuclear levels within their Doppler width. Because the maximum nuclear cross sections $\lambda_\gamma^2/4\pi$, or $1/4\pi$ are typically 1 kbarn for 1,MeV and thus much larger than atomic cross sections of 1–10 barn for the Compton scattering and pair creation, we can strongly reduce the atomic background. At the same time, the de-excitation cascades of individual levels are much less complex than when exciting a multitude of levels.

8.6.1 Nuclear Diagnostics

Now, let us consider the M1 and E1 excitations of Figure 8.12 as strong excitations for potential applications. The low-lying M1 excitation can be employed via nuclear resonance fluorescence (NRF), using focusing lenses for tomography and microscopy with micrometer resolution [3]. In Figure 8.14, we show M1 excitations for actinides, which can be used in nuclear fuel assay or radioactive waste monitoring [55, 56]. On the basis of the M1 transition in ^{7}Li (478 keV), we can, for example, perform a high-resolution tomography of growing defects in ^{7}Li batteries. In medical physics, we can study the distribution of ^{7}Li in the

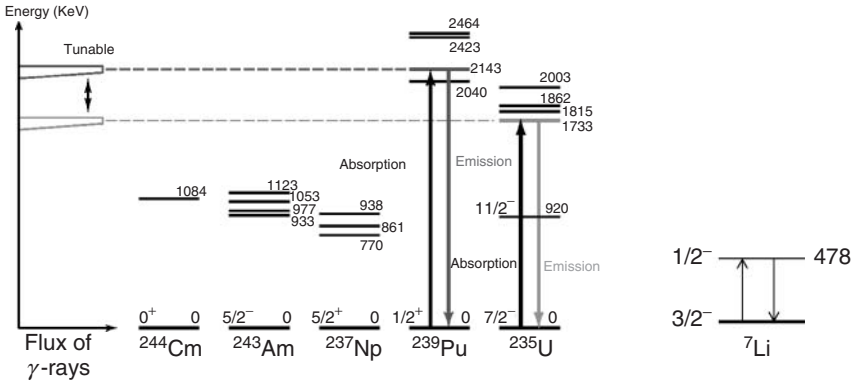


Figure 8.14 M1 transitions used in nuclear photonics.

brain during treatment of manic depressive alternating insanity [57, 58].

8.6.2 Production of Longer-Lived Spin Isomers via Doorway States of Collective Cascades

Neutron capture can be used to populate many low-spin cascades. By gating on high-multiplicity cascades, ''inverted'' collective rotational bands can be identified on top of isomers, which start out from low-spin states and feed the isomer.

Sequences of collective cascades, as shown in Figure 8.15, may require less highly excited doorway states to reach certain isomers.

Owing to their internal structure, the higher QOC E1 excitations frequently decay by quadrupole transitions to higher-spin states. We want to use this feature to populate high-spin isomers such as ^{195m}Pt as the first example to produce medical radioisotopes for diagnostics and therapy [59]. We also want to look for the 31 year 16^+ isomer in ^{178}Hf by starting with $7/2^+$ of ^{177}Hf in thermal neutron capture. Cascades with five or more transitions should show this feeding.

8.6.3 Selective Population of Shape Isomers

We envisage to perform a detailed spectroscopy of fission isomers in the second and third minima of the nuclear potential energy after populating these states in γ reactions with highly monochromatic γ beams. With novel, four to five orders of magnitude more intense (γ) beams as, for example, available from the upgraded HIγS facility (HIγS2), we can select much more narrow (γ) beams. After a first rough monochromatization, we will use the broader width of a Ge double-crystal monochromator, reaching an energy bandwidth of about 10^{-4} at 4 MeV. Thus, we obtain a factor of 100 better resolution and we can resolve 1^- and 2^+ states and their fission fragment angular distributions, using, for example, thick GEM (gas electron multiplier) detectors. In case of the (superdeformed) second minimum, we should see regular $0^+, 1^-$, and 2^+ bands, while for states in the (hyperdeformed) third minimum with large β_2, β_3 deformations, the 1^- excitations belong to separate bands. Thus, in the first step, we would like to study them with our newly developed THICK-GEM detectors.

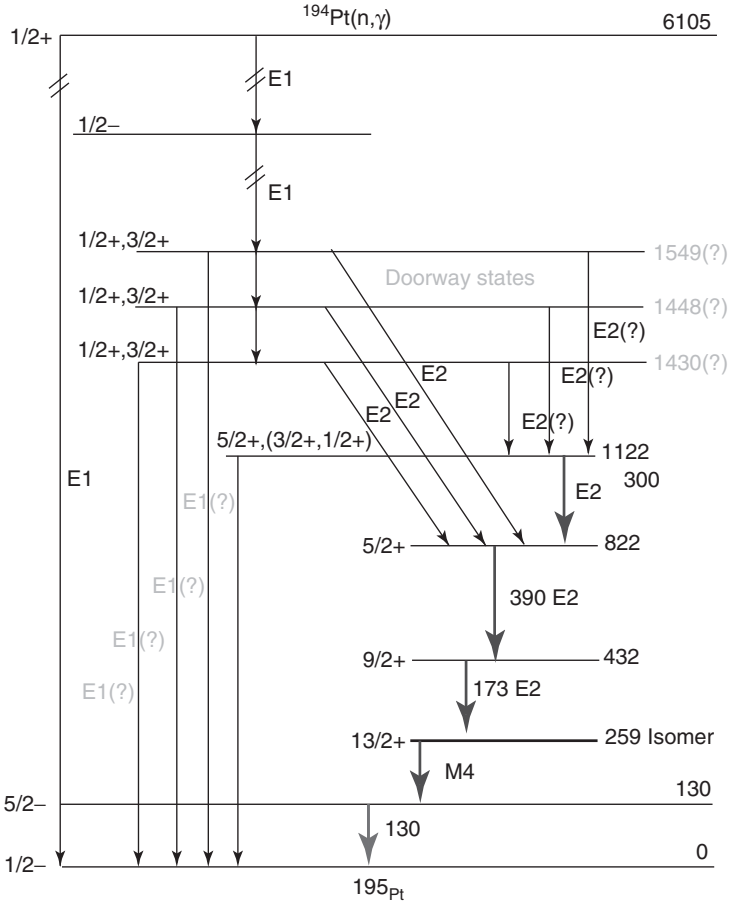


Figure 8.15 Inverted rotation band for ^{195}Pt observed in the ^{194}Pt(n,γ) excitation.

8.6.4 Secondary Beams

It has been reported that cold, brilliant microneutron beams [60] and intense, moderated positron beams [61] can be produced by releasing neutrons and positrons from special targets with focused γ beams. Compared to these former publications, the quality of the secondary beams can now be improved by the new γ optics, starting with much smaller target spots. For the positron production, the much stronger rise of the cross section for positron production close to threshold due to high-field QED effects will be explored. Then the foil thickness and foil diameters and distances between foils can be reduced significantly, resulting in higher positron yields.

8.6.4.1 Brilliant Neutron Beams

From the nuclear shell model, we know that weakly bound $4s_{1/2}$ or $3s_{1/2}$ neutrons are expected for the mass number $A = 140 - 180$ or 40–60. We expect that these states are isomers in the picosecond to microsecond range, because a neutron

with, for example, 1 eV binding energy will tunnel out by $\sim 10^4$ fm, exhibiting very little overlap with the nuclear core. Similar halo isomers have been studied in detail for neutron-rich light nuclei like ^{11}Be and ^{11}Li [62, 63], but now we want to study halo nuclei in stable nuclei at high excitation energies. By waiting for a short time, compared to the lifetime of the isomer, the recoil energy from the production of the isomer is thermalized. Now, in the second step, very low energy neutrons are released via a second, low-energy intense photon beam (see Figure 8.16). In this way, one obtains directed neutrons with about $(100\ \text{mrad})^2$ opening angle and a bandwidth better than 0.1%.

Thus pulsed, polarized microneutron beams of high intensity are obtained without moderation. Presently, the best thermal neutron beams at reactors have a brilliance of $10^2 \text{n}(\text{mm}^2\ \text{mrad}^2\ 0.1\%\ \text{BW s})^{-1}$, while we expect with the γ FEL and suitable γ optics a neutron beam brilliance of $10^{16} \text{n}(\text{mm}^2\ \text{mrad}^2\ 0.1\%\ \text{BW s})^{-1}$. For a coherent photon field of the second laser, a much more directed neutron emission may occur because of the collective alignment of the charged core nucleus with respect to the halo neutron, resulting in even 10^4 times higher brilliance. This would represent a major jump in neutron beam properties similar to the transition from a lamp to a laser. Furthermore, these beams can be pulsed with time structure down to picoseconds, much shorter than that achievable at reactors with chopper wheels in the millisecond range. An essential feature is the much smaller neutron beam source diameters in the micrometer to nanometer range, enabling the study of very small sized targets: (i) looking at very small targets with special functional properties such as colossal magnetoresistance, (ii) very small

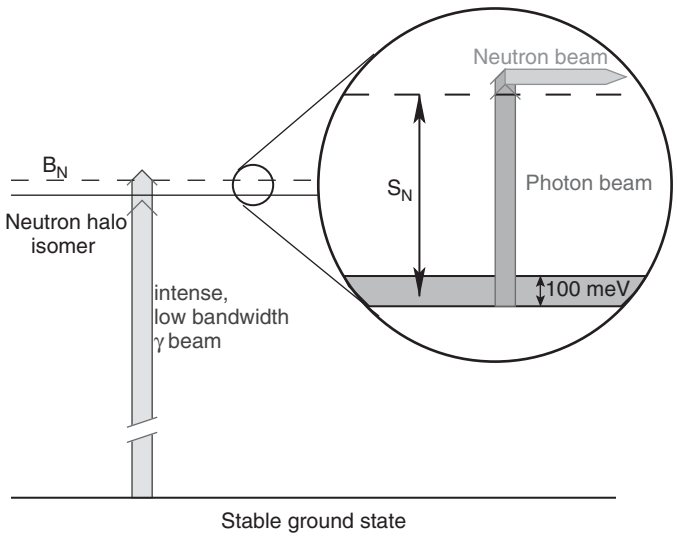


Figure 8.16 Schematic picture of the neutron production. With focused, monochromatized, intense γ beams, neutron halo isomers are produced below the neutron binding energy B_N at a separation energy S_N. In the second step, a photon beam of much lower energy is used to release the neutrons.

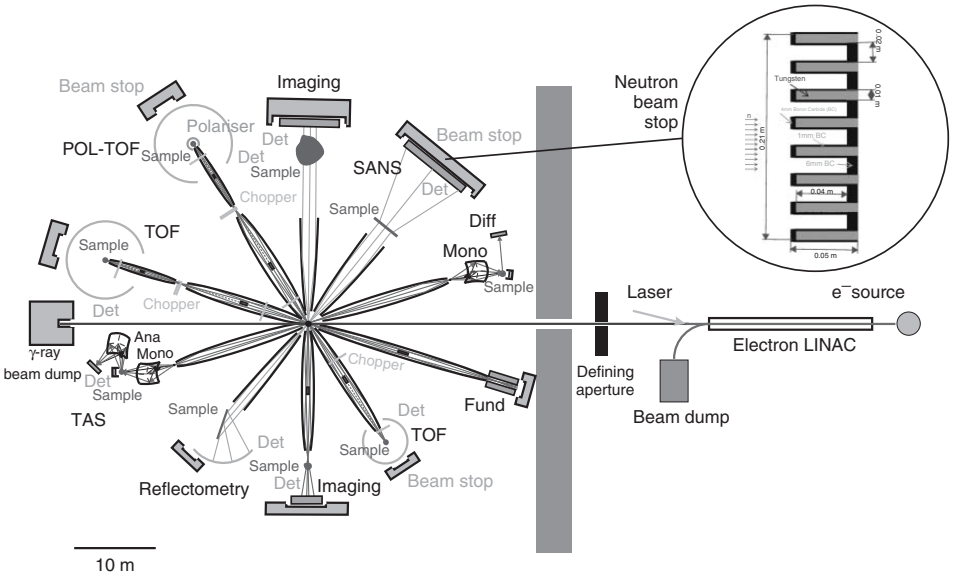


Figure 8.17 Experimental setup for the new brilliant microneutron source. Components of possible experiments are shown (TOF, time of flight; TAS, triple axis spectrometer; SANS, small-angle neutron scattering camera; Diff, diffractometer; Fund, fundamental or nuclear physics). Neutron beams with low and high divergence are extracted with parabolic (light gray beams) and elliptic (dark gray beams) neutron guides.

targets in high pressure cells to study phase transitions for geoscience, or (iii) biological samples. Here, the different neutron beam results in a magnification that helps overcome the poor spatial resolution of neutron detectors.

In Figure 8.17, we show the layout of a possible neutron facility, where we would always feed two opposite beamlines, depending on the direction of the electric field of the second laser, which should have the same requirements on the timing and energy structure. Again neutron beam guides will match the experiment to the beam. The parabolic or elliptic guides will have only one neutron reflection with small reflection losses. The direct line of sight may be blocked by a beam stopper, reducing, for example, γ quanta from the converter target. In Ref. [60], the efficiencies of different neutron setups are compared to present setups at reactors or spallation sources, and they mostly turn out to be much superior. Thus, many new possibilities open up in neutron physics. As this neutron source requires much less investment compared to a reactor or a spallation source and produces much less radioactivity and radioactive waste, the efforts for radioprotection are strongly reduced. Thus one could dream of facilities at various universities, serving a much larger community.

8.6.4.2 Brilliant Positron Beams

Weaker positron beams are generated using intense β^+ emitters such as ^{22}Na and thin W or solid Ne moderators, resulting in about $5 \times 10^5 - 5 \times 10^6$ moderated positrons per second. At present, the NEutron induced POsitron source MUniCh (NEPOMUC) at the Munich FRM-2 reactor delivers the highest flux of 9×10^8 moderated positrons per second [64]. We want to compare its setup shown in Figure 8.18 with the possibilities of positron production with the new brilliant γ beams. At NEPOMUC, about $10^{16}\gamma$ per second from

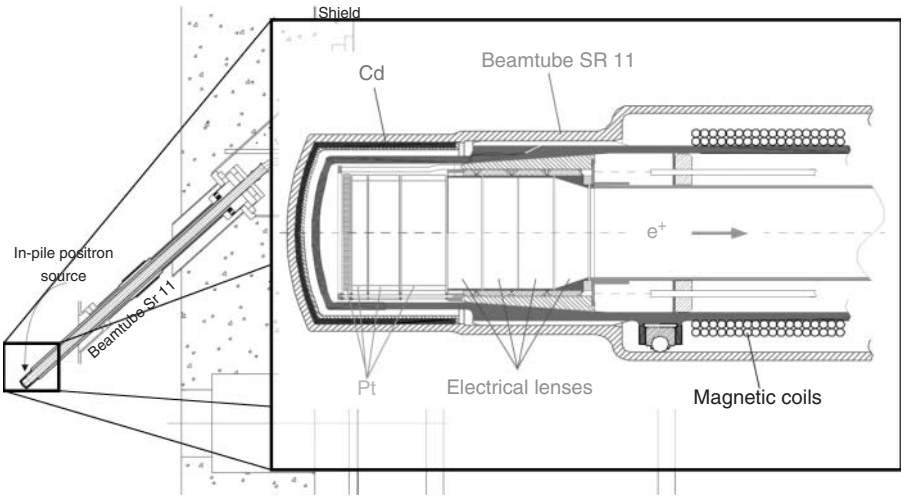


Figure 8.18 NEPOMUC positron source at the Munich reactor FRM-2 showing an enlarged cut through the in-pile source, with the Cd neutron converter, the Pt γ converter and positron moderator, the acceleration electrodes, and the solenoid coils for guiding the e^+ beam.

a Cd neutron converter with γ energies above 1.5 MeV produces positrons in the high-Z Pt converter with a broad spectrum peaking at about 800 keV. These positrons are moderated in the Pt foils to low energies, diffusing to the surface from a depth of about 50 nm into vacuum. Owing to the negative positron work function, they leave the surface with typical energies of 2 eV normal to the surface and an energy spread of about 0.5 eV. Here, typical moderation efficiencies for W and Pt foils are $10^{-3}-10^{-4}$, mainly due to the depth of 50 nm from where positrons can diffuse out to the surface. Subsequently, the positrons are extracted with voltages up to 1 kV from the moderator, resulting in a brilliant positron beam guided by a solenoidal magnetic field. With the new brilliant direct γ beams, a more efficient production from high Z-materials is possible [61], although we start with a less intense flux of $10^{13}\gamma$ per second. This γ beam is reduced to a small diameter of a few micrometer and focused on a stack of Pt converter foils, as shown in Figure 8.19. By choosing an optimum γ energy, rather than a broad spectrum of capture γ lines, we can work with much lower primary positron energies, perhaps even making use of the much improved yields close to threshold by high-field QED effects. Thus, we can use much thinner converter foils with much larger conversion efficiencies due to the better extraction by diffusion. The operation of the NEPOMUC source showed that any changes in the source close to the reactor core are time consuming. With the new γ beams, we have easy access to the converter/moderator and may additionally deposit layers of solid Ne at 7 K onto the Pt foils, thus improving the moderator efficiency [65]. As we now can realize a much smaller γ-beam diameter with much smaller foil width b and probably will have much smaller foil thicknesses d_W, the total length L will be strongly reduced, while at the same time the conversion and moderation efficiency can be strongly improved. Here, a detailed optimization has to be performed, but compared to Ref. 61, significantly improved yields are expected in the range of $10^{10}\ s^{-1}$. Furthermore, we get spin-polarized positron beams with a picosecond pulsing.

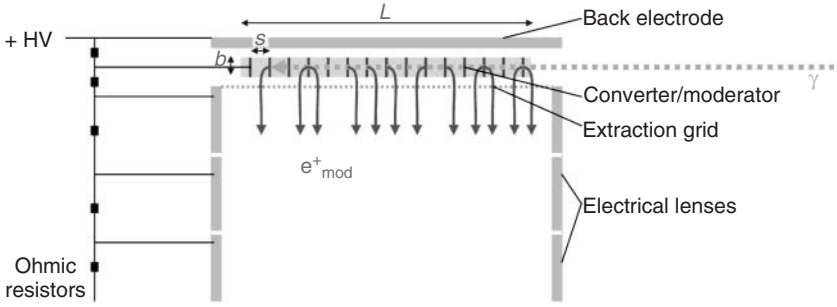


Figure 8.19 Experimental setup for a γ-beam-driven source of moderated positrons. The Pt foils are used as a converter/moderator. The thickness d_W of the W foils is determined by the average e^+ energy. The diameter b of the foils determines the spacing s between the foils with b : s = 3 : 1 for efficient extraction. The total length $L \approx N(s + d_W)$ is determined by the number N of converter foils.

There are many applications of intense positron sources [66] foreseeable in material sciences and solid-state physics, using spatially resolved defect spectroscopy with CBDS (coincidence Doppler-broadening spectroscopy) or AMOC (age-momentum correlation), where the positron lifetime and Doppler shift are detected simultaneously. Another technique is PAES (positron-induced Auger electron spectroscopy) for spatially resolved surface analysis. With PLEPS (pulsed low-energy positron system), depth- and position-dependent investigations of microstructures, for example, plastic deformations in the vicinity of microcracks, can be performed. Also, fundamental studies of few-body leptonic systems, such as Ps^-, or larger systems, such as $(Ps_2)^-$, become possible.

Thus, the new positron source will be comparable or superior to the present world-leading facilities, addressing many new fields in fundamental and applied sciences.

8.7 Conclusions and Outlook

A γ beam with a much larger spectral intensity becomes very essential, because many more applications become feasible. In this context, we presently explore the concept of a γ FEL. We envisage to use the strong focusing channel in the laser-plasma ''bubble acceleration'' of electrons, where the disk-like deformation of the Coulomb fields causes much closer linear chains and much smaller distances. Here, the required temperatures for crystallization become much larger for the same plasma parameter Γ

$$\Gamma = \frac{e^2}{R \cdot kT} \tag{8.11}$$

R becomes very small due to the strong focusing, the Lorentz contraction, and the transversely flat Coulomb fields of the individual relativistic electrons. Thus, for a cold electron injection into the bubble, there is a good chance to produce microbunched electron beams. Once beams with larger spectral intensity become available, one could send such a beam through a monochromator and refractive lenses and use it as the seed beam for the second γ-FEL stage. In this way, the first stage would act like an SASE (self-amplified stimulated emission) FEL and the second stage still should result in much larger spectral densities.

Another open question is up to which γ energies these studies can be extended. Can we reach up to γ energies of 10 or 100 MeV? Furthermore, where do the strong nonlinear processes stop?

It is clear that a totally new energy range of the Delbrück scattering will open up the energy range between 0.5 MeV and probably up to 20 MeV, supplementing the atomic physics of X-rays by the nuclear physics of γ-rays. It has to be seen how many new applications can be established and how large the field of secondary cold neutron beams and cold positron beams is going to be.

8.8 Key Topics of γ Optics

8.8.1 The Delbrück Scattering and Nonperturbative QED

The theories of the last century used for the Delbrück scattering of the second-order perturbation theory of γ quanta from the high electric field of the nucleus. Today, it is clear that the electric field of the nucleus

is about 10^3 times the Schwinger field of $E_s = m^2/e = 1.3 \times 10^{18}$V m^{-1} and we have to use nonperturbative QED even for the Schwinger field to describe these processes with much larger cross sections. At about 1 MeV, strong virtual pair creation is essential. There are several other high-field processes, such as photon splitting with nonperturbative QED, and we will establish this new nonperturbative field of physics. This then opens up the new fields of γ optics in the megaelectronvolt range.

8.8.2
Efficient γ Monochromators and the Delbrück Scattering

Again the Delbrück scattering from the high electric field of the nucleus results in much larger scattering amplitudes for higher γ energies, for example, above about 1 MeV, and thus leads to much more efficient γ monochromators. Until now, only the Rayleigh scattering of the outer atomic electron was used. Because we now can built about 10^7 times more brilliant γ beams, by the Compton backscattering of brilliant photon beams from brilliant electron beams, the germanium or silicon monochromators gain much more importance in the new γ optics.

8.8.3
Brilliant γ Beams and (γ,γ') Reactions

With the 10^7 times more brilliant γ beams, we can excite with E1 or M1 transitions with high-intensity-dedicated nuclear levels very selectively. Here we can use very special levels for NRF to detect low levels of these isotopes. This may be ^{7}Li or actinides to control via these isotopes, for example, radioactive waste management or green energy. With the much more narrow bandwidth, we can separate parity doublets by several orders better and afterward we can separate E1 and M1 decay easily by their different angular distributions. Owing to the much larger intensity, we can excite very selectively certain levels, which, for example, decay very selectively by ''inverted'' cascades to isomers that we want to use for cancer treatment or in diagnostics. With the narrowband nuclear excitation, at the same time, the atomic processes of the Compton scattering or pair creation are strongly reduced.

References

1. Habs, D. *et al.* (2012) *Phys. Rev. Lett.*, **108**, 184802.
2. Snigirev, A. *et al.* (1996) *Nature*, **384**, 49.
3. Schroer, C.G. *et al.* (2010) *Acta Phys. Polon.*, **A 117**, 357.
4. Als-Nielsen, J., McMorrow, D. *et al.* (2010) *Elements of Modern X-Ray Physics*, 2nd edn, John Wiley & Sons, Inc., New York.
5. Beil, H. and Bergere, R. (1980) *Monochromatic and Identifiable Photons Used in Photonuclear Research*, CEA-N-2144, C.E.N. Saclay, France.
6. Weller, H.R. *et al.* (2009) *Prog. Part. Nucl. Phys.*, **62**, 257.
7. Serafini, L. *et al.* (2012) A European Proposal for the Compton Gamma-ray Source of ELI-NP, IPAC2012, TUOBB01, New Orleans, May 22.
8. Barty, C. (LLNL), ELI-NP-meeting, *http://www.eli-np.rp /gamma-beam-meeting-august-presentation.php* (accessed 14 March 2013).
9. Barty, C. (2010) Development of MEGa-Ray technology at LLNL, *http:/www.eli-np.ro/excecutive committee-meeting-april12-13.php* (accessed 14 March 2013).
10. Habs, D. *et al.*, arXiv:1206.0940v1[physics.acc-ph].
11. Kessler, E.G. *et al.* (2001) *Nucl. Instrum Methods*, **A 457**, 187.
12. Leemans, A.P. *et al.* (2006) *Nat. Phys.*, **2**, 696.
13. Schmidt, K. *et al.* (2010) *Phys. Rev. ST AB*, **13**, 091301.
14. Malka, V. *et al.* (2012) *Phys. Plasma.*, **19**, 055501.

15. Falkenstein, H. *et al.* (1992) *Atom. Data Nucl. Data Tabl.*, **50**, 1.
16. Milstein, A. *et al.* (1994) *Phys. Rep.*, **243**, 183.
17. Deutsch, M., Hart, M. *et al.* (1984) *Phys. Rev. B*, **30**, 2, 640.
18. Petrascheck, D., Rauch, H. *et al.* (1984) *Acta Crystallogr.*, **A 40**, 445.
19. Akhmadaliev, S.Z. *et al.* (1998) *Phys. Rev.*, **C58**, 2844.
20. Akhmadaliev, S.Z. *et al.* (2002) *Phys. Rev. Lett.*, **89**, 601802.
21. Bjorken, J.D., Drell, S.D. *et al.* (1965) *Relativistic Quantum Fields*, McGraw-Hill Book Company, New York.
22. Jackson, J.D. *et al.* (1975) *Classical Electrodynamic*, John Wiley & Sons, Inc., New York.
23. Toll, J.S. *et al.* (1952) The Dispersion Relation for Light and the Applications involving Electron Pairs, Princeton.
24. Kirilin, G.G., Terekhov, I.S. *et al.* (2008) *Phys. Rev.*, **A 77**, 032118.
25. Cheng, H. *et al.* (1969) *Phys. Rev.*, **182**, 1873; (1970) *Phys. Rev.*, **D2**, 2444; (1992) *Phys. Rev*, **D5**, 3077.
26. Milstein, A., Strakhovenko, V.M. *et al.* (1982) *Phys. Lett.*, **A90**, 447.
27. Milstein, A. *et al.* (1983) *Phys. Lett.*, **A95**, 135.
28. Adler, S.L. *et al.* (1971) *Ann. Phys.*, **67**, 599.
29. Sauer, F. (1931) *Z. Phys.*, **69**, 742; (1932) **73**, 547.
30. Heisenberg, W., Euler, H. *et al.* (1936) *Z. Phys.*, **98**, 714.
31. Schwinger, J. *et al.* (1951) *Phys. Rev.*, **82**, 664.
32. Narozhnyi, N.B. *et al.* (1968) *Sov. Phys. JETP*, **27**, 360.
33. Scherdin, A. *et al.* (1992) *Phys. Rev.*, **D 45**, 2982.
34. Jentschel, M. *et al.* (2011) *Rapid Commun. Phys. Rev. C*.
35. Dewey, M.S. *et al.* (1994) *Phys. Rev.*, **B50**, 2800.
36. Authier, A. (2008) *Dynamic Theory of X-ray Diffraction*, IUCR Monographs on Crystallography 11, Oxford University Press.
37. Zachariasen, W.H. *et al.* (1967) *Theory of X-ray diffraction in crystals*, John Wiley & Sons, Inc., Dover.
38. Matsushita, T. and Hashizume, H. (1983) X-ray monochromators, in *Handbook on Synchrotron Radiation*, Vol. 1b (ed. E.E. Koch), North Holland.
39. Habs, D. *et al. Nucl. Photon.*, arXiv:1201:4466v1[acc-ph].
40. Bohle, D. *et al.* (1986) *Nucl. Phys.*, **A458**, 205.
41. Pietralla, N. *et al.* (1998) *Phys. Rev.*, **C 58**, 184.
42. Herzberg, R.D. *et al.* (1999) *Phys. Rev.*, **C60**, 051307.
43. Enders, J. *et al.* (2004) *Nucl. Phys.*, **A 741**, 3.
44. Volz, U. *et al.* (2006) *Nucl. Phys.*, **779**, 1.
45. Bohr, A., Mottelson, B.R. *et al.* (1975) *Nuclear Structure*, Benjamin INC, London.
46. Weidenmüller, H.A., Mitchell, G.E. *et al.* (2009) *Rev. Mod. Phys.*, **81**, 539.
47. Titov, A.I. *et al.* (2006) *J. Phys. G Nucl. Part. Phys.*, **32**, 1997.
48. Alfimenkov, V.P. *et al.* (1983) *Nucl. Phys.*, **A 398**, 93.
49. Alfimenko, V.P. *et al.* (1984) *JETP Lett.*, **39**, 416.
50. Sushkov, O.P., Flambaum, V.V. *et al.* (1980) *JETP*, **32**, 377.
51. Adelberger, B.G., Haxton, W.C. *et al.* (1985) *Annu. Rev. Nucl. Part. Sci.*, **35**, 501.
52. Desplanques, B. *et al.* (1998) *Phys. Rep.*, **297**, 1.
53. Mitchell, G.E. *et al.* (2010) *Rev. Mod. Phys.*, **82**, 2845.
54. Segre, E. (1977) *Nuclei and Particles: An Introduction to Nuclear and Subnuclear Physics*, 2nd edn. Benjamin, Reading, MA.
55. Hayakawa, T. *et al.* (2010) *NIM*, **A621**, 695.
56. Pruet, J. *et al.* (1006) *J. Appl. Phys.*, **99**, 123102.
57. WHO (2004) The global burden od disorder, update; Center for disease control and prevention *www,cdc.gov/injury/wisqars* (accessed 14 March 2013).
58. Lichtinger, J. *et al.* Detection of lithium traces in organic samples, FRM-II annual report 2009, 29, Munich.
59. Habs, D. *et al.* (2011) *Appl. Phys. B*, **103**, 501.
60. Habs, D. *et al.* (2011) *Appl. Phys. B*, **103**, 485.
61. Hugenschmidt, C. *et al.* (2012) *Appl. Phys. B*, **106**, 241.
62. Hansen, P.G. *et al.* (1995) *Annu. Rev. Nucl. Part. Sci.*, **45**, 691.
63. Tanihata, I. *et al.* (1996) *J. Phys. G. Nucl.Part. Phys.*, **22**, 158.
64. Hugenschmidt, C. *et al.* (2008) *Nucl. Instrum. Methods*, **593**, 616.
65. Mills, A.P., Gullikson, E.M. *et al.* (1986) *Appl. Phys. Lett.*, **49**, 1121.
66. Hugenschmidt, C. (2010) Positron sources and positron beams, *International School of Physics E. Fermi, Course CLXXIV: Physics with Many Positrons*, IOS Press, Amsterdam.

9
The Proton

Allen Caldwell

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

9.1 Introduction

The proton is one of the most common particles in the universe, the nucleus of hydrogen, and, together with the neutron, one of the two building blocks for all heavier nuclei. It has charge $+1$, intrinsic angular momentum (or spin) $\hbar/2$, is about 1 fm across, and weighs 1.67×10^{-27} kg. The proton quantum numbers are determined by the valence quarks (two up quarks with charge $+2/3$ and one down quark with charge $-1/3$). So we know the external features of the proton well. However, when we look at the proton more closely, we find an amazingly complex structure. At a fine resolution scale, one sees that the proton is filled with a dense soup of virtual quarks, antiquarks, and gluons. How this dense soup comes about, and how all these particles act together to give the proton its mass, angular momentum, size, and so on, are open research topics. The proton is far from understood at a fundamental level, and is today a laboratory for forefront research in particle and nuclear physics.

In this chapter, we describe the current knowledge concerning the proton. The understanding of the proton has gone through several revolutions in the last 100 years, and we start with a short review of this history. The central part of the text then describes the proton in terms of its constituents and the current state of knowledge on this structure is outlined. As will become clear, although we now have garnered a wealth of information on this structure, there are still many gaps in our knowledge. The last part of this chapter then gives an alternative description of the proton, where it is explained that what we call the proton depends on the rest frame, and the interpretation of what we observe – the result of scattering experiments involving the proton, say – can be very different. Very different pictures arise depending on the frame of reference one chooses to describe the physics.

9.2 The Proton as Building Block of Matter

9.2.1 Early History

The history of the proton can be said to have started in 1911, when Ernest Rutherford demonstrated that the data collected by Hans Geiger and Ernest Marsden in 1909 on the scattering of α-particles (which consist of two protons and two neutrons)

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

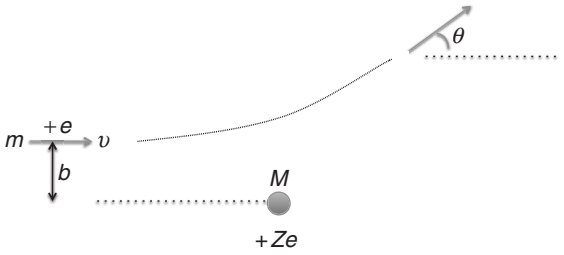


Figure 9.1 Diagram showing the quantities used to discuss Rutherford scattering. A light particle of mass m, electric charge $+e$, and speed v is incident on a heavy target of mass M and charge $+Ze$. The impact parameter, or distance of closest approach if the incident particle would proceed undeflected, is b, and the scattering angle is θ.

in a thin gold foil were consistent with atoms having a compact nucleus [1]. The idea of using scattering distributions to determine the structure of the target is at the heart of the rest of this chapter, so we start by describing the kinematics of such as scattering process.

Imagine a particle of charge $+e$ scattering on a fixed scattering center with charge $+Ze$, where e is the magnitude of the charge of the electron. The trajectory is described in Figure 9.1. Applying conservation of angular momentum as well as conservation of energy, Rutherford showed that the scattering angle θ is related to the impact parameter, b, as

$$b = \cot(\theta/2)\frac{Z\alpha\hbar c}{mv^2} \tag{9.1}$$

This equation has been written in the modern language using the Planck constant, h, divided by 2π; that is, $\hbar = \frac{h}{2\pi}$ and the speed of light c, as well as the fine-structure constant, $\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \approx 1/137$. The incident particle was assumed nonrelativistic, with mass m and speed v. The product $\hbar c$ appears frequently in particle physics expressions and has a value memorized by all practitioners

$$\hbar c = 197 \text{ MeV fm with}$$
$$1\,\text{MeV} = 10^6 \text{ eV and } 1 \text{ fm} = 10^{-15}\,\text{m}$$

For a particle with incident mass $m = 1\,\text{GeV}/c^2$ and speed $v = 0.1c$ scattering on a charge center with $Z = 100$, we have

$$\cot(\theta/2) \approx \frac{b}{14\,\text{fm}}$$

This formula, derived by Rutherford, shows that large scattering angles are possible if the impact parameter to the nuclear charge is small enough, which requires that the charge Ze is compact. The data of Geiger and Marsden indeed showed the presence of large scattering angles and proved that the nuclear charge was contained within a small volume. This discovery led to the modern picture of atoms consisting of a nucleus with positive charge, surrounded by orbiting electrons.

The proton is now understood to be one of the three particles making up ordinary matter. Atomic nuclei are made of neutrons and protons, and atoms contain an equal number of protons and electrons. The basic properties of atoms and molecules are determined by the number of protons in the atom.

9.2.2 Proton Structure

In a scattering process, we typically do not know the impact parameter of the incoming particle on the scattering center, and we can only calculate the probability for a scatter into a given solid angle. The Rutherford scattering formula in this case takes the form

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \tag{9.2}$$

where σ is the "cross section" for the process and Ω is the solid angle (two-dimensional angle subtended by an object). Here, E is the energy of the incident particle. The name *cross section* is appropriate in the sense that it gives an effective area seen by the beam particles. The differential cross section, $d\sigma/d\Omega$, is the quantity that is extracted from scattering experiments by counting the number of particles scattered into different solid angles.

The Rutherford scattering cross section must be modified in case the target mass cannot be taken to be infinite. Furthermore, if the scattering particles carry intrinsic angular momentum (spin), then additional modifications of the angular distribution must be accounted for. Finally, if the target is not represented as a pointlike charge, but is rather an extended charge distribution, the full result for, for example, electrons scattering on a proton, can be written as

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 \cos^2(\theta/2)}{4E^2 \sin^4(\theta/2)} \frac{E'}{E} \cdot F(Q^2) \tag{9.3}$$

where $F(Q^2)$ is a "form factor" that takes into account the fact that the target particle has an extended charge distribution. The quantity E' is the energy of the outgoing electron, and Q^2 is the negative of the square of the four-momentum transfered in the reaction (discussed in more detail below). If the target is pointlike, then $F(Q^2) = 1$. The extraction of $F(Q^2)$ from scattering experiments allows to infer the structure of the target.

The quantity Q can be understood intuitively. It is related, via the Heisenberg uncertainty relation, to the transverse distance scale, r, at which the proton is probed:

$$r \sim \frac{\hbar c}{Q} \approx \frac{0.2}{Q} \text{ fm}$$

with Q expressed in giga electronvolts. Large values of Q correspond to small transverse distance scales. There is a limit to how large Q can be in a given experiment – it cannot exceed the center-of-mass energy of the probe and target. Rutherford and his colleagues were limited to the energies available from radioactive decays, which are in the few mega electronvolt range. With the advent of particle accelerators in the 1950s, a new era ensued where ever higher center-of-mass energies could be reached. The resulting distance scale that could be probed is shown graphically in Figure 9.2.

9.2.3 Discovery of Quarks and Gluons

The advent of particle accelerators provided access to much higher energies in a controlled laboratory environment and allowed for a great number of discoveries. One of the first was the observation that the form factor of the proton, $F(Q^2)$ in Eq. (9.3), was not equal to one. In a famous 1956 paper [2], Robert McAllister and Robert Hofstadter reported on the study of elastic scattering of electrons from protons

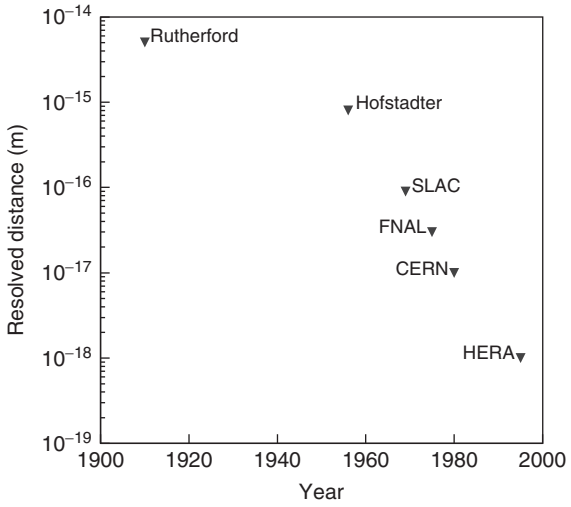


Figure 9.2 The development of the resolution to which the structure of the proton could be probed as a function of time. Note that the dates are just indicative – the measurements spanned a period of many years. The experiments mentioned are discussed in the text.

and α-particles. They used the 187-MeV electron beam newly commissioned at the Stanford Linear Accelerator Center (SLAC) and a gaseous hydrogen target, and studied the angular distribution of the scattered electrons. $F(Q^2)$ is related to the electric charge density distribution in the proton, $\rho(\vec{r})$, via a Fourier transform:

$$F(q^2) = \int \mathrm{d}^3 r e^{i\vec{q}\cdot\vec{r}} \rho(\vec{r}) \text{ with } q^2 = -Q^2$$

and the measurement of F therefore allows the extraction of a charge distribution of the proton. The conclusion from the measurements was that the proton has a radius of about 0.7 fm. It was clear at this point that the proton is an extended object and the study of its internal structure was launched.

9.2.3.1 Proton Size

As protons do not have sharp edges, the size of the proton is a matter of convention. One definition is given by the root-mean-square radius of the electric charge density:

$$r_{\mathrm{EM}} = \sqrt{\int r^2 \rho(\vec{r}) d^3 r}$$

The values found for this quantity are around 0.9 fm [3], but there is some discrepancy as different results are found when atomic electrons are replaced by muons [4]. The reasons for this difference are currently the matter of intense debate.

A different definition of the size of the proton results from studying the proton–proton scattering cross section. In contrast to scattering of electrons off protons, which is mediated by photons, scattering reactions involving only protons can be mediated by gluons. The typical proton–proton scattering cross sections are around 40 mb (milli Barns, where 1 b is $10^{-28}\mathrm{m}^2$), that is,

$$\pi r_{\mathrm{QCD}}^2 = 40{\cdot}10^{-31} \ \mathrm{m}^2$$

or $r_{\mathrm{QCD}} \approx 1$ fm. The scattering cross section for two protons increases with energy, indicating that the size of the proton is an energy-dependent quantity. The behavior of the scattering cross section with energy is directly related to the fundamental structure of the proton, and is discussed in more detail in later sections of this chapter.

9.2.3.2 Proton Substructure

As the energy of the electrons used in the scattering process was increased, it was seen that the proton could be excited to higher mass states (resonances) much like atoms can be excited. The deexcitation of the proton yielded other particles in the detectors, such as π mesons and neutrons. In 1969, a group of researchers from SLAC and from MIT published a paper in which it became clear that the proton was composed of pointlike particles [5], that is, their measurements showed that the total cross section for scattering on a proton could be decomposed into a sum of cross sections for scattering on pointlike objects. These objects were identified with the partons of Feynman's model [6], where the proton was considered to be a collection of particles each containing a fraction of the proton momentum. Feynman's parton model is described below. Bjorken [7] had earlier predicted the behavior expected for the cross section under the assumption that protons were composed of pointlike quarks (quarks had been hypothesized by Gell-Mann [8] and Zweig [9] to explain the complicated set of excited states seen for protons and other particles), and the results of the experiment were in line with Bjorken's calculations. At this point, the picture of the proton as composed of pointlike quarks became the accepted model.

A series of experiments ensued at SLAC, at the Fermi National Accelerator Laboratory (FNAL) outside Chicago, and at the Centre Europeen pour la Recherche Nucleaire (CERN) near Geneva, Switzerland. In addition to using electrons to probe the proton, neutrinos and muons were also used. In parallel, the theory of quantum chromodynamics (QCD) was established as the best theory for modeling the strong interactions, such as those between quarks in the proton. QCD predicted that there should be neutral force carriers, named *gluons*, present in the proton, and in fact the FNAL and CERN experiments demonstrated that the quarks only carried roughly 50% of the proton momentum (for high-energy protons). The rest was assumed to come from gluons. Direct confirmation of the existence of gluons was found at DESY in the late 1970s [10]. With this, the basic picture of the proton as composed of three valence quarks, two up- and one down-type, determining the quantum numbers of the proton, and a number of gluons, and quark–antiquark pairs, was established.

9.2.4 Spin Crisis

In 1988, the EMC Collaboration performing scattering experiments at CERN reported that the proton intrinsic angular momentum (spin) could not be accounted for by simply adding the spins of the quarks and antiquarks [11]. This came as a major surprise, and was dubbed the ''spin crisis.'' The measurements were later confirmed and extended by a series of other experiments, and it now looks as if the gluons do not carry a substantial fraction of the spin either. If the spin of the proton is not accounted for by intrinsic angular momentum of the quarks and gluons, then the orbital angular momentum of the partons must account for a large part of the spin of the proton. The simplest model of the proton as three quarks, two up quarks and one down quark, each of spin $\hbar/2$ adding up to a net spin of $\hbar/2$ for the proton may work as a crude model when the proton is not looked at too closely, but is clearly not adequate once the proton is probed at small distance scales. The spin crisis is still with us, and many new experiments have been proposed to try to untangle the different

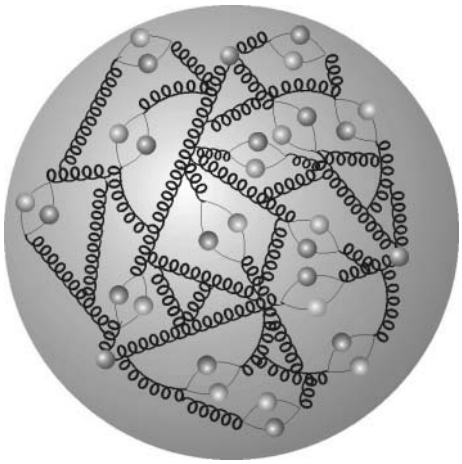


Figure 9.3 An artists conception of the proton viewed with a high-resolution probe. The spheres represent quarks and antiquarks, while the spirals represent gluons.

components of the proton intrinsic angular momentum.

9.2.5 The Proton Gets Complicated

The surprises were not over. With the advent of the hadron electron ring accelerator (HERA) collider, which started taking data in 1992, it was possible to probe proton structure at considerably smaller distance scales. Given the history of finding new levels of substructure at smaller distance scales, it was natural to ask whether the quarks themselves had some substructure. A form factor similar to the one in Eq. (9.3) was evaluated for quarks, but no signs of quark substructure were found. However, what was found was a very large density of quarks, antiquarks, and gluons, each carrying a tiny fraction of the proton momentum, and it was seen that the number of small momentum partons increased as the resolution of the probe improved [12], that is, objects that looked like single quarks or antiquarks at a coarser resolution were revealed to be composed of many quarks, antiquarks, and gluons. An artist's rendition of the current understanding of the proton, viewed with a high-resolution probe, is shown in Figure 9.3.

HERA presented the particle physics community with another surprise [12] – that it is possible to scatter an electron from the proton at large angles, produce a number of new particles, and still have the proton survive the scattering process intact. These reactions

$$\mathrm{e} + \mathrm{p} \rightarrow \mathrm{e} + \mathrm{p} + X \tag{9.4}$$

where e represents the incoming electron, p the proton, and X some particles produced in the scattering process, can only be understood as the scattering off a new type of component in the proton, something with no net color (the charge of the strong force, as described in Section 9.2.6). This color-neutral combination goes under the name "pomeron," after the Russian physicist Isaac Pomeranchuk, and it remains mysterious despite its invention more than 30 years ago.

The proton, viewed at distance scales much smaller than 1 fm, is clearly

a complicated object, consisting of a large collection of quarks, antiquarks, and gluons but also color-neutral objects like pomerons and perhaps other objects. The theoretical description of the proton at these small distance scales is a work in progress.

9.2.6 Proton Mass

The proton has a mass of $1.67262178 \times 10^{-27}$ kg with an uncertainty of about $0.00000007 \cdot 10^{-27}$ kg [13]; this is very well known experimentally. However, theoretically, the value is not well understood. The proton is composed of components that are either very light (the quarks) or massless (the gluons). The bulk of the proton mass arises from the energy of interaction between the partonic constituents and should, in principle, be calculable starting from the basic equations of QCD. However, these calculations are extremely challenging and can only be performed numerically on supercomputers using a lattice to discretize space–time. Current calculations are able to reproduce the proton mass to within a small percentage [14].

9.2.7 Stability of the Proton

As far as is known today, the proton is a stable particle; that is, it does not decay. It is not obvious that this should be the case, because the proton is a composite object and there are lighter objects into which it could possibly decay. In fact, there exist theoretical models that predict that protons have a finite lifetime, and experimental searches continue. The current tests of the proton lifetime have reached the level of 10^{29} years [13].

9.3 Models and Theories

9.3.1 Feynman's Parton Model

In 1969, Richard Feynman [6] suggested looking at a proton in a reference frame where it is moving very quickly, and to view the proton as composed of "partons," pointlike constituents that move parallel to the direction of motion of the proton, sharing the proton momentum. In this model, the partons are assumed to be massless and are also assumed to carry no significant transverse momentum. The fractional momentum carried by a parton i is denoted by x_i, and conservation of momentum implies

$$\sum_i x_i = 1$$

The values of x_i are changing constantly owing to interactions between the partons. However, in the frame chosen, the interactions between partons are time dilated, while the interaction with the electron (or other probe) is not, so that an almost instantaneous sampling of the x_i is performed. It is assumed that the probability density for a parton to carry a momentum fraction x, denoted $f(x)$, does not depend on the process or on the proton energy, but is an intrinsic property of the proton, and

$$\int x\,[u(x) + \overline{u}(x) + d(x) + \overline{d}(x) + g(x) + \cdots]\,\mathrm{d}x = 1$$

where the sum extends over all quarks, antiquarks, and gluons. This expression is known as the *momentum sum rule*. It states that the sum of the average momenta carried by the constituents of the proton should equal the proton momentum. For

the proton quantum numbers to come out correctly, we require

$$\int [u(x) - \overline{u}(x)]\mathrm{d}x = 2$$
$$\int [d(x) - \overline{d}(x)]\mathrm{d}x = 1$$

with $u(x)$ and $d(x)$ representing the densities of up- and down-type quarks, respectively. The symbols $\overline{u}$ and $\overline{d}$ represent the densities of the antiquarks. The net integral for all other quark species (there are six known types of quarks) should yield zero. The gluon density is given by $g(x)$. The only limitation on the number of gluons is via the momentum sum rule.

The parton model was historically very useful and gave insight into the structure of the proton. However, limitations were already clear at the outset, such as the inability to deal with partons that carry only a small fraction of the proton momentum (''wee'' partons, in the language of R. Feynman). Parallel to the development of the parton model, a full theory of the strong interactions was developed, QCD.

9.3.2 Quantum Chromodynamics

The development of the theory of strong interactions is the work of many physicists, and a review of the resulting theory, QCD, is beyond the scope of this chapter. However, we attempt to give the reader an intuitive picture of the proton, or hadronic matter in general (i.e., matter containing quarks and gluons), through a more heuristic discussion.

9.3.2.1 Confinement and Asymptotic Freedom

We note here one important experimental fact, which is that quarks and gluons are always bound in ''colorless'' clusters. Color is the name given in QCD to the analog of charge in electricity and magnetism. Particles with color will feel the strong force, whereas color-neutral particles will not. An electron, for example, has no color and does not feel the strong force. The quarks and gluons carry color and undergo strong interactions. A key feature of the strong interactions is that the strength of the force increases as the distance between colored particles increases, as opposed to electromagnetism or gravity. If the separation becomes too large, then new colored particles can be spontaneously produced by the vacuum, and these will act to neutralize the color charge. The quarks and gluons always end up bound in color-neutral systems (such as the proton). If a quark is produced with a large energy, then the production of more colored particles from the vacuum usually goes through several steps, such that jets consisting of many particles are produced. The corollary to this behavior is that the strong force becomes weaker and weaker at short distances, a feature known as *asymptotic freedom* [15]. This leads to the perhaps curious fact that it is easier to perform theoretical calculations in QCD for high-energy processes, as these involve short distances. The weakness of the force implies that calculating the first few terms in a perturbative expansion usually yields results of good accuracy.

9.3.2.2 Formation of Parton Densities

Imagine that we could place a bare quark in the vacuum. What would happen? Field theory has taught us that the quark would occasionally fluctuate into a quark–gluon system. The lifetime of the fluctuation would depend on the ''virtuality'' of the quark–gluon fluctuation. A large energy–momentum difference between the initial quark and the quark–gluon

fluctuation would lead to a short lifetime, whereas a small energy–momentum difference would lead to much longer-lived fluctuations. Eventually, the quark and gluon would recombine. If the quark and gluon from the first fluctuation had long-enough lifetimes, then they could act as the source for further fluctuations. The gluon could either split into two gluons (this is an important feature of QCD–that the force particles can interact with each other, as opposed to photons, which cannot) or split into a quark–antiquark pair.[1] The seed quark, after radiating the initial gluon, could radiate another gluon, which could become the source of further radiation. If the resulting gluons and quarks lived long enough, then they could also be the sources of further fluctuations. A ''radiation chain'' could be produced with a large number of virtual quarks and gluons surrounding the initial quark. These would be in a highly coherent quantum state, which would eventually collapse back to the original quark.

In addition to the quarks and gluons generated from the initial quark, the vacuum can also produce virtual particles, and, in fact, does so continuously. Quark–antiquark pairs can be created, which can also be the source of further virtual particles. These quarks and gluons could, and would, interact with the quarks and gluons from our seed quark. The initial bare quark that we placed in the vacuum would eventually reach some equilibrium state where it is surrounded by a cloud of virtual quarks and gluons (and, in much smaller numbers, other virtual particles that can couple to the quarks, such as photons or weak bosons, and their decay products).

What does this mean for our picture of the proton? We know that the quantum numbers of the proton are determined by the three valence quarks, two u- (or up-) type quarks and one d- (or down-) type quark. But the valence quarks are not single quarks: viewed at small distance and timescales, they are now understood to be composed of clouds of virtual particles.

9.3.2.3 Comparison to QED

A major difference between QCD and QED (quantum electrodynamics) is that the force particles can interact with each other. Photons cannot directly scatter from photons, whereas gluons can scatter directly from other gluons, or split into two gluons. This aspect, together with the fact that the strong coupling constant, reflecting the strength of the force, being much larger than the fine-structure constant, $\alpha \approx 1/137$, of QED makes theoretical predictions very difficult. In particular, it is not possible to calculate the parton densities starting from the basic equations of QCD. However, at large Q, the strong coupling constant[2] becomes small enough that some predictions are indeed possible, such as the evolution of the parton densities with Q. In fact, the gluon density is directly linked to the evolution of the quark density with Q.

9.3.2.4 Electron Structure

Consider a gedanken experiment with an electron similar to the one carried out above for a quark. A similar situation holds, with the difference being that an electron radiates first a photon or Z-boson, which can then split into further virtual particles

1) In the following, we will not distinguish between quarks and antiquarks–it should be understood that they will be present in equal numbers in the cloud of virtual particles.

2) Note that, despite their names, the fine-structure constant and the strong coupling constant are, in fact, not constant, but are energy dependent!

(electron–positron pairs, quark–antiquark pairs, etc.). As mentioned above, the lifetime of the fluctuations depend on the difference in energy–momentum generated by the fluctuation. Long-lived, almost real, photons will often be accompanied by further virtual particles. Highly virtual photons, with large Q, will have very short lifetimes and produce few daughter particles. We therefore expect small Q events to have properties similar to hadron–hadron scattering events, whereas at high Q, the photon or Z-boson will couple to a quark that originated from the proton.

9.4 Modern Description of the Proton

9.4.1 Standard Picture

As should be clear from the previous section, the description of the proton structure depends very much on the distance scale at which it is probed. With a resolution of about 1 fm, all that can be seen is that the proton is not a point particle, but has a size of about 0.7 fm. With a higher-energy probe, it becomes apparent that the proton is a complex collection of partons, and that the number resolved by the probe changes considerably as the probe becomes finer and finer. There is today no ultimate description of the proton starting at the smallest distance scale and building up the proton from fundamental constituents, and the description is largely just a parameterization of observations. The behavior of the structure as one changes the scale is, however, predicted in QCD. In this section, we give the state-of-the-art description of the structure of the proton.

The standard description of the proton structure rests on two concepts – the Bjorken frame and the factorization theorem. The Bjorken frame is a reference frame where the proton is moving very quickly (very similar to the frame that Feynman has discussed in his model), such that the scattering between the probe and the target is effectively instantaneous. In this frame, the incoming electron (or other probe particle) exchanges a photon (or W- or Z-particle) with one of the quarks or antiquarks in the proton. What is probed is the fractional momentum distribution carried by quarks or antiquarks in the proton. These distributions are measured at different values of the scale, Q^2. The results are useful thanks to the factorization theorem [16], which states that the structure at small distance scales is not affected by any long-distance physics, and it makes sense to define parton distributions for the proton. The parton densities extracted in this way can then be used to make predictions for other processes, such as proton–proton scattering, as long as the distance scales involved in the process are small. This requirement is met by requiring that large transverse momenta are involved, as, via the Heisenberg uncertainty relation, large momenta correspond to small distances.

9.4.1.1 Cross Section

The extraction of the structure functions is performed via scattering experiments. A schematic description of an electron–proton scattering event is shown in Figure 9.4. In this example, the electron scatters from one of the valence quarks, and is deflected into the detector. The struck quark is knocked out of the proton, and produces a jet of particles. The remnant of the proton also produces jets of particles, primarily in the direction of the incoming proton. Measuring the products of such a

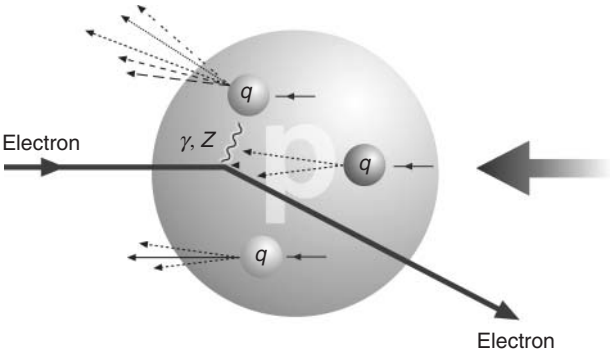


Figure 9.4 Schematic description of an electron–proton scattering event.

scattering reaction allows for the extraction of the structure of the proton.

The cross section for the reaction can be written as

$$\frac{d^2\sigma}{dx dQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[\left(1+(1-y)^2\right)F_2 - y^2 F_L + xF_3\right] \quad (9.5)$$

where F_2, F_L, and F_3 are structure functions. They depend on the relativistic invariants x and Q^2. The variable Q^2 is the same as the one introduced above, but the x here is a quantity that does not depend on the reference frame. In the Bjorken frame, it effectively has the meaning of momentum fraction of the proton as imagined by Feynman. The variable y is also a frame-independent quantity and is related to the center-of-mass scattering angle of the electron, θ, as follows:

$$1 + \cos\theta = 2(1-y)$$

The dependence of the cross section on y is a consequence of angular momentum conservation and the specific form results from the spins of the electron and quarks. The kinematic variables are bounded as follows:

$$0 \le Q^2 \le E_{cm}^2$$
$$0 \le x \le 1$$
$$0 \le y \le 1$$

9.4.1.2 Structure Functions

The structure functions are related to the density of quarks and gluons, with F_2 and F_3 given by

$$F_2(x,Q^2) = \sum_q e_q^2 x\left[q(x,Q^2) + \bar{q}(x,Q^2)\right] \quad (9.6)$$

$$F_3(x,Q^2) = \sum_q e_q^2\left[q(x,Q^2) - \bar{q}(x,Q^2)\right] \quad (9.7)$$

where e_q is the electric charge and $q(x,Q^2)$ and $\bar{q}(x,Q^2)$ are the densities for quark and antiquarks of type q. As can be seen from these formulae, the structure function F_2 gives the total momentum density of quarks and antiquarks at a given x and Q^2, weighted by the square of the quark charges. F_3 gives the net number of quarks, weighted by the square of the quark charges. The structure function F_L is more complicated, and is given by

$$F_L(x,Q^2) = \frac{\alpha_S}{4\pi}x^2\int_x^1\frac{dz}{z^3} \times\left[\frac{16}{3} + 8\sum_q^2\left(\frac{1-x}{z}\right)zg(z,Q^2)\right] \quad (9.8)$$

Here α_S is the strong coupling constant. In the parton model, $F_L = 0$ because of

angular momentum conservation. However, QCD predicts it should be nonzero, and early measurements showing $F_L > 0$ were a strong indication that QCD was, in fact, the correct theory of the strong interactions. An important point about F_L is that it depends on the gluon density, g, which is difficult to measure directly. Access to the gluon density can also be gained from the derivative of the structure function F_2. This can be understood intuitively: the number of quarks and antiquarks is determined from the number of gluons via the splitting process $g \to q + \bar{q}$, so that what appears to be a gluon at one resolution, Q_1^2, can be resolved as a quark–antiquark pair at a higher resolution $Q_2^2 > Q_1^2$.

Very extensive data exist on the structure functions. An example of the available data is shown in Figure 9.5. Here, the data are shown as a function of Q^2 for fixed values of x. F_2 (here the "reduced cross section," σ_r is shown, which is very close to F_2) is seen to increase with Q^2 at the smaller values of x, while it decreases with Q^2 at the larger values of x. We understand this as follows: as Q^2 increases, we resolve more and more structure, so that large-x objects seen at a coarse scale are actually composed of many smaller x objects when viewed at a finer scale. The strongly rising behavior of F_2 for small values of x implies that the number of quarks and antiquarks increases rapidly with Q^2 at this x.

9.4.1.3 Specific Quark Densities

It is possible to extract more information on the individual parton densities by studying specific reactions. For example, neutrino–nucleon scattering can only proceed via the exchange of a "weak boson," because the neutrino carries no electric charge. In a charged current interaction,

$$\nu_\mu + N \to \mu + X$$

a W^- boson is exchanged between the neutrino and the nucleon. The incoming neutrino is converted to a muon in the scattering process. The target nucleus, N, breaks up and a number of final state particles, described by X, are produced. Charge conservation implies that only specific quark flavors can take part in the interaction. In this case, only quarks and antiquarks with positive charge are allowed.

Another technique for accessing a particular quark flavor is to study the production of specific particles that contain the quark. For example, observing hadrons in the system X containing heavy quarks such as charm or beauty gives sensitivity to these quark densities. This technique does not work for the light quarks because they are abundantly produced in the formation of X, while the heavier quarks are most likely already present in N.

9.4.1.4 Parton Density Functions

The multitude of datasets available for scattering off of protons can be analyzed in a "global fit" by parameterizing the parton density distributions of the proton. A typical form used for the parameterization is

$$xq(x, Q_0^2) = Ax^B(1-x)^C(1 + Dx + Ex^2) \tag{9.9}$$

with A, B, C, D, and E being free parameters. The valence quark distributions are usually given individual density function, as are the gluon and sea quarks (the quarks and antiquarks generated by gluon radiation). There are also different parameterizations and approaches for handling

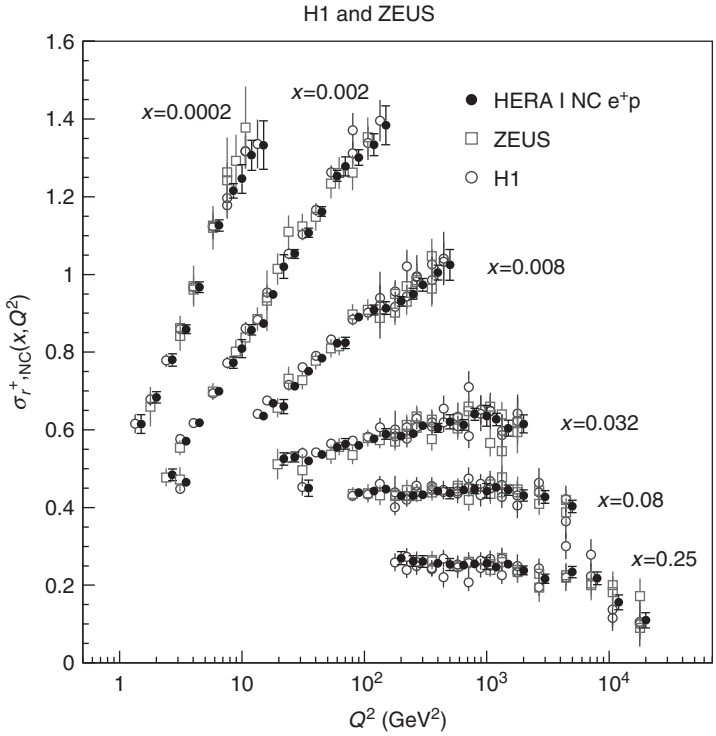


Figure 9.5 A small sample of data for $\sigma_r \approx F_2$ as a function of Q^2 for fixed values of x, (Source: From [17].)

heavier quarks, such that the fits have a large number of parameters and quite some flexibility. The small x behavior of the parton density is driven by the parameter B. For valence quarks, $B > 0$ and the parton density dies out at small x. On the other hand, $B < 0$ for the gluon and sea quarks, so that these densities increase as x decreases.

The parameterizations are defined at a scale Q_0^2, which is usually taken in the range $1 - 4\,\mathrm{GeV}^2$, and the structure functions are directly calculated from the chosen parton densities at that scale. The variation of the parton densities with scale is then given by the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi equation [18]:

$$\frac{\mathrm{d}F_2}{\mathrm{d}\ln\ Q^2} = \frac{\alpha_S(Q^2)}{2\pi}\left[\int_x^1 \left(\frac{x}{z}\right) P_{qq}\left(\frac{x}{z}\right) F_2(z,Q^2) + \int_x^1 2\sum_q e_q^2 \frac{\mathrm{d}z}{z}\left(\frac{x}{z}\right) P_{qg}\left(\frac{x}{z}\right) zg(z,Q^2)\right] \tag{9.10}$$

where P_{qq} (P_{qg}) are the "splitting functions." They represent the probabilities of finding a quark with fractional momentum x inside a quark (gluon) with momentum fraction z, and can be calculated from the basic rules of QCD. Note that the expressions written here are from a leading order analysis in perturbation theory.

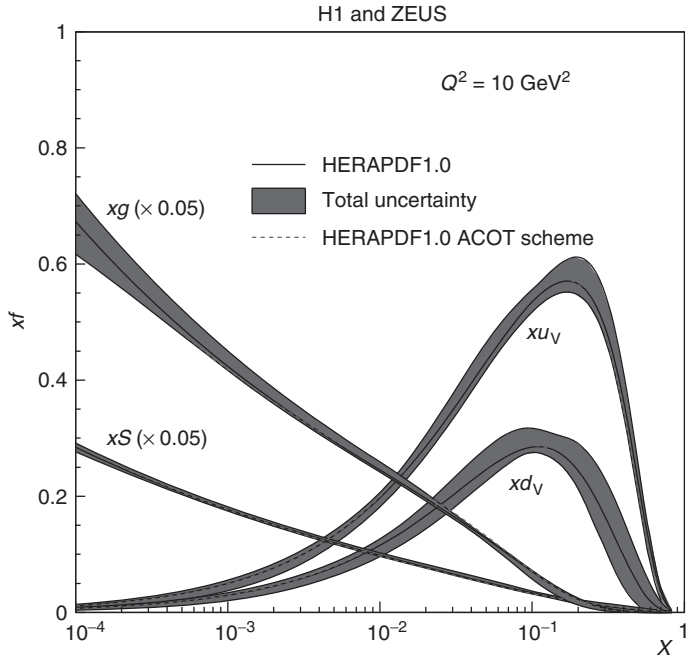


Figure 9.6 The parton densities extracted from a fit to large datasets of scattering cross sections. The valence quark momentum densities are given by xu_V, xd_V while the gluon momentum density is given by xg and the combination of quarks and antiquarks making up the "sea" are given by xS. Note that the densities for the gluons and sea are reduced by a factor 20 for visibility. (Source: From [17].)

The result of a global fit is shown in Figure 9.6. The parton momentum density functions are shown as a function of x for a fixed $Q^2 = 10\,\text{GeV}^2$. The valence quarks densities are shown as u_V and d_V, and peak near $x = 0.2$. The bands on the plots represent an estimate of the uncertainties on the density functions, and are typically around 10%. The valence quark densities die out at small values of x, while the gluon and sea (quark+antiquark) densities rise quickly. Approximately 50% of the proton momentum is carried by the small-x partons.

9.4.1.5 The "Wee" Partons

The gluon and sea quark densities are seen in Figure 9.6 to increase rapidly with decreasing x at $Q^2 = 10\,\text{GeV}^2$. The steepness of the rise changes dramatically as Q^2 is varied. This behavior is quantified by fitting the structure function data for $x < 0.01$ by a function of the form $F_2 \propto x^{-\lambda}$. The extracted values of λ are given as a function of Q^2 in Figure 9.7. The figure includes a band, which represents the effective λ seen in hadron–hadron interactions. What is observed is that the HERA data are in agreement with the band up to $Q^2 \approx 1\,\text{GeV}^2$. Above this Q^2, the value of λ increases. This is interpreted as a sign that the sea of quarks and antiquarks become visible above this Q^2, and this leads to the increasing value of λ. Using the

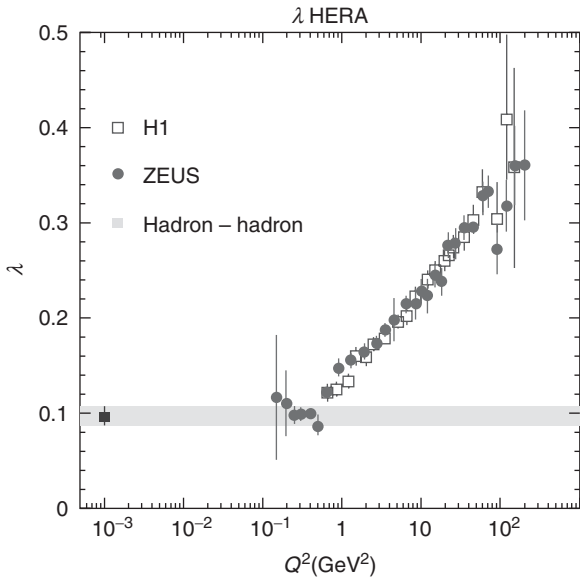


Figure 9.7 The power of $1/x$ observed in the structure function data at small x, λ, plotted as a function of Q^2. The grey band represents an effective λ observed in hadron–hadron scattering experiments. (Source: From [19].)

Heisenberg uncertainty principle, we can relate Q^2 to a distance, via $d = \frac{\hbar c}{Q} \approx 0.2$ fm. This distance is considerably smaller than the charge radius of the proton measured via elastic scattering and indicates that there is an intermediate level of structure in the proton. This is presumably the size scale of the valence quarks surrounded by the cloud of virtual particles, the ''wee-parons,'' inside the proton.

9.4.1.6 Spin Structure of the Proton

As the spin crisis has shown, the proton intrinsic angular momentum (spin) is not simply the sum of the angular momenta carried by the valence quarks. In general, we can write the spin as a sum of terms:

$$\frac{1}{2} = \frac{1}{2}\Delta\Sigma + L_Z^q + \Delta G + L_Z^g \quad (9.11)$$

where the first term is the angular momentum coming from the net spin of the quarks in the proton, the second term is the angular momentum carried by the quarks and antiquarks, the third term is from the net spin of the gluons, and the last term is from the angular momentum carried by gluons. The sum is set to 1/2, which is the net intrinsic angular momentum of protons in units of $\hbar$. Quarks and antiquarks have spin 1/2, and this factor has already been explicitly written, so that $\Delta\Sigma$ represents the fraction of the spin of one quark. Gluons have spin one.

Measuring the components of the proton spin requires using polarized beams and polarized targets (i.e., beams and targets where the spin is in a preferred direction). The scattering cross section for different polarization settings is then measured and analysis of these cross sections allows the extraction of the different components in Eq. (9.11). As an example, we consider the term due to the quark spins, $\Delta\Sigma$. This term is usually broken down into components due to the valence quarks, and components due to sea quarks:

$$\Delta\Sigma = \Delta u_V + \Delta d_V + \Delta q_S \quad (9.12)$$

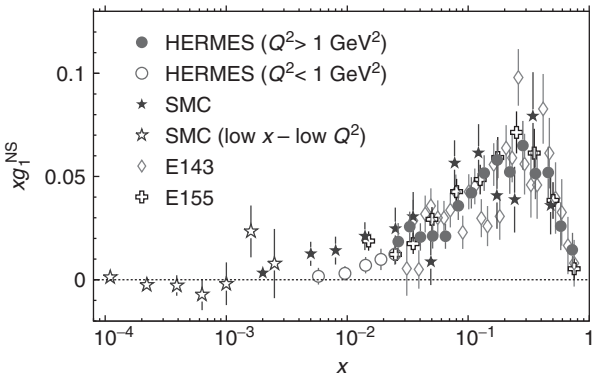


Figure 9.8 The g_1 structure function measurement for the proton as a function of x. HERMES, SMC, E143, and E155 are four different experiments that have studied polarized scattering using a variety of beam energies. (Source: Figure from [20].)

The different terms appearing in Eq. (9.12) can be accessed from an analysis of the polarized structure function g_1, defined as

$$g_1(x) = \frac{1}{2}\sum_q e_q^2(\Delta q + \Delta\overline{q}) \qquad (9.13)$$

with $\Delta q = q^+(x) - q^-(x)$. The superscript on the quark density corresponds to quark distribution functions with the helicities parallel and antiparallel to the proton helicity.

A summary of measurements of the g_1 structure function is shown in Figure 9.8. No significant Q^2 dependence has been observed, so that data with different Q^2 are plotted together as a function of x. The integral of g_1 over x can be used to extract $\Delta\Sigma$, and recent results give a value of about 1/3 with about 10% uncertainty, that is, only about 1/3 of the proton spin is accounted for by summing the projections of the spins of the individual quarks and antiquarks in the proton. The remainder has to come from the other terms. Recent measurements indicate that the sum of the gluon spin projections can only account for a small fraction of the proton spin, leaving the bulk to be explained by orbital angular momentum. A new round of experimentation with much higher precision measurements will be needed in order to verify this description.

9.5 Wee Partons and Other Mysteries

In this section, we discuss what is perhaps the most interesting aspect of the proton – the ''small-x'' physics, or, in the words of Richard Feynman , the physics of the ''wee'' partons [6]. This is the physics of very short-lived virtual quarks, antiquarks, and gluons, found at the heart of all matter, which has been made visible by the HERA data. For Feynman, a parton was a basic constituent of a hadron. We now understand Feynman's partons to be the quarks and gluons of QCD.

9.5.1 Space–Time Picture of Electron–Proton Interactions

The data described in the previous section, and used to extract the parton densities in the proton, can be interpreted differently. In the Bjorken frame, the unknown parton densities were all lumped into the proton structure functions. However, in the rest frame of the proton, which could be considered the appropriate one to discuss

proton structure, it will be seen that the probe is also evolving. To see this, consider electron–proton scattering in the rest frame of the proton. If the electron is in its bare state when traversing the proton, then nothing happens. If, however, the electron is accompanied by a virtual photon, then an interaction can occur, leading to a change in momentum of the electron and (often) the breakup of the proton. The photon energy, E_γ, is given in the proton rest frame by

$$E_\gamma = \frac{Q^2}{2Mx}$$

where we have used the standard units of particle physics, $\hbar = c = 1$. Here x is the quantity defined by Bjorken and M is the mass of the proton.

The photon can either interact directly with a quark in the proton, or it can radiate quarks, antiquarks and gluons and one of these then subsequently interacts with a quark or gluon in the proton. The distance over which a quark–antiquark pair produced by the virtual photon can survive is

$$l \approx \frac{1}{2Mx} \approx \frac{0.1\text{fm}}{x}$$

that is, for $x \leq 0.1$, the coherence lengths of the photon fluctuations will be longer than the size of the proton. We therefore see that when x is not too small, the interactions are between the photon and a quark or antiquark in the proton, whereas at smaller x, the interactions are between a quark or gluon from the photon and a quark or gluon in the proton.

9.5.2 Photon–Proton Cross Section

The electron–proton interaction cross section can be written as a product of a flux of photons from the electron and the interaction cross section for a photon with the proton. The extracted photon–proton cross section is displayed as a function of l in Figure 9.9. At each Q^2, the cross section increases monotonically with l. This can be understood as a result of giving the photon more and more time to emit virtual particles before the reaction. The slope of the increase with l is Q^2 dependent. At lower Q^2, the cross section rises more slowly than at high Q^2. This is presumably a result of the fact that the quark–antiquark dipole at high Q^2 is very small, and its effective area therefore increases by a big factor with each radiation, whereas the smaller Q^2 dipoles are already "large" and therefore grow relatively slowly. The cross sections become comparable, independent of Q^2, for $l \approx 10^9$ fm, indicating that if photons fluctuations are allowed to develop for such a distance, then the initial quark–antiquark configuration becomes irrelevant and the cross sections become universal, that is, the photon states have had enough time to evolve into what is presumably a universal size.

The detailed theoretical understanding of how this radiation cloud is produced is not yet available, although interesting models are currently under discussion.

9.5.3 Models of the "Wee" Partons

The behavior of the total cross section at small-x is strikingly simple. The cross section rises as a power of l which is Q^2 dependent, and is otherwise featureless. This is in contrast to the large-x behavior, where the intrinsic structure of the proton is seen. However, the simplicity of the data has not led to a simple theoretical modeling. Despite the efforts of a large number of physicists over a period

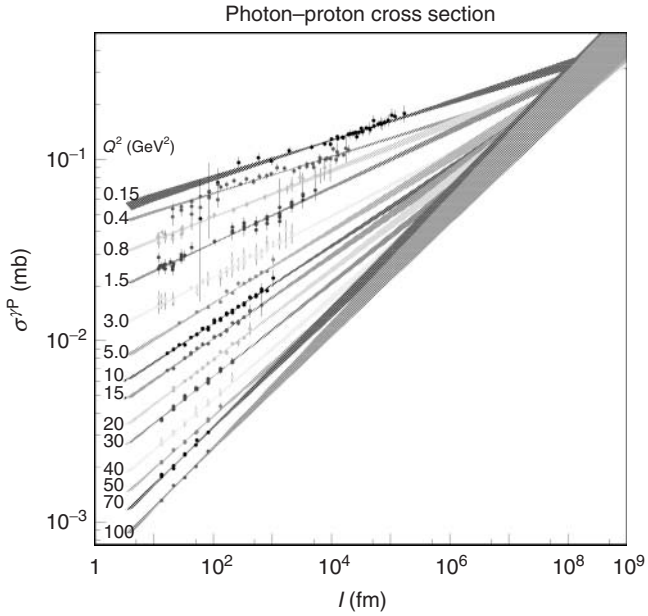


Figure 9.9 The $\gamma^* P$ cross section as a function of l (see text) for different Q^2.

covering 30 years, we still do not have a satisfactory theoretical understanding of the high-energy behavior of the scattering cross sections. Highlights during this period include the development of a new perturbative expansion of the equations of QCD, which is particularly suited for the small-x partons, named the *Balitsky–Fadin–Kuraev–Lipatov* (*BFKL*) *evolution equation* in honor of its creators [21], and the extension of the simple splitting scheme of gluons into quarks and antiquarks to allow for recombination effects [22] and subsequent developments. The Golec-Biernat, Wüsthoff model [23] has helped tremendously, in that it provided a framework for explaining simultaneously the total as well as diffractive and exclusive reactions. Further developments of this model have allowed for better fits and for the mapping of the hadronic matter profile in the proton. The discovery that this model contained a new type of scaling [24] – geometric scaling–has prompted many theorists to look for new ways to attack the small-x physics. The color-glass condensate [25] is an approach to the small-x physics that provides an appealing physical picture. It predicts that there should exist a kinematical region where perturbative calculations are possible although the Q^2 scale is small. The large gluon density at small-x introduces a new scale, the saturation scale, Q_S, which can be large enough so that the tools of classical field theory can be used. This model is now under intense development, and has been applied successfully to explain nucleus–nucleus collision data.

Despite these exciting and promising developments, we still do not have a satisfactory theoretical understanding of the small-x structure of the proton. As this small-x structure is a universal feature of all matter (viewed at high energies), it is imperative that the theoretical work

continues, and that new data is made available that could point the way to the correct interpretation.

9.5.4 Summary and Conclusion

We have reviewed the historical development in the understanding of the proton, as well as current state of knowledge. The history began 100 years ago with the demonstration by Rutherford and colleagues that atoms contain a dense core, or nucleus. The proton is one of two particles making up nuclei, and as such one of the most important particles in nature. With the advent of particle accelerators, energies were high enough to reveal that the proton has structure. Today, the basic proton features are known: it has charge $+1$, intrinsic angular momentum (or spin) $\hbar/2$, is about 1 fm across, and weighs $1.67{\cdot}10^{-27}$ kg. The proton quantum numbers are determined by the valence quarks (2 up quarks with charge $+2/3$ and one down quark with charge $-1/3$). However, when we look at the proton more closely, we find an amazingly complex structure. At a fine resolution scale, one sees that the proton is filled with a dense soup of virtual quarks, antiquarks, and gluons. The description of the proton in terms of these constituents is possible with the use of structure functions and the evolution of these structure functions with energy scale is theoretically understood. However, how the particular densities of quarks, antiquarks, gluons and even mysterious pomerons comes about, and how all these particles act together to give the proton its mass, angular momentum, size, and so on, are open research topics. The proton is far from understood at a fundamental level, and is today a laboratory for forefront research in particle and nuclear physics. Some of the open issues in the quest to understand the proton more fully were described in the last section. The answer to these open questions will have far-reaching consequences as the basic physics underlying the proton is fundamental and universal in nature.

Glossary

Cross section: Effective size of a scattering center (process dependent).

Hadrons: Particles made up of quarks, antiquarks, and gluons.

Helicity: The projection of the spin vector in the direction of motion.

Impact parameter: Distance of closest approach of a line to a point.

Parton: Generic term for the quarks, antiquarks, and gluons making up hadrons.

Resonance: Higher mass excitation of quarks and gluons.

Rutherford scattering: Coulomb-scattering process off very heavy target.

Sea: Nonvalence quarks and antiquarks in a hadron.

References

1. Rutherford, E. (1911) *Philos. Mag. Ser. 6*, **21**, 669.
2. McAllister, R.W. and Hofstadter, R. (1956) *Phys. Rev.*, **102** (3), 851.
3. Mohr, P.J. *et al.* (2008) *Rev. Mod. Phys.*, **80**, 633–730
4. Pohl, R. *et al.* (2010) *Nature*, **466**, 213–216
5. Bloom, E.D. *et al.* (1969) *Phys. Rev. Lett.*, **23** (16), 930.
6. Feynman, R. (1972) *Photon-Hadron Interactions*, Addison-Wesley Publishing Company, Inc.
7. Bjorken, J.D. (1969) *Phys. Rev.*, **179** (5), 1547.
8. Gell-Mann, M. (1964) *Phys. Lett.*, **8** (3), 214.

9. Zweig, G. (1964) CERN Report, **No.8182/TH.401**.
10. For a review of the history, see Söding, P. (2010) *Eur. Phys. Jou.*, **H 35** (1), 3.
11. Ashman, J., *et al.*, EMC Collaboration (1988) *Phys. Lett.*, **B206**, 364.
12. For a review, see Abramowicz, H. and Caldwell, A. (1999) *Rev. Mod. Phys.*, **71**, 1275.
13. Beringer, J., *et al.*, (Particle Data Group) (2012) *Phys. Rev.*, **D86**, 010001.
14. See e.g. Dürr, S. *et al.* (2008) *Science*, **322**, 1224.
15. (a) Gross, D.J. and Wilczek, F. (1973) *Phys. Rev. Lett.*, **30** (26), 1343; (b) Politzer, H.D. (1973) *Phys. Rev. Lett.*, **30** (26), 1346.
16. Collins, J.C., Soper, D.E., and Sterman, G. (1989) in *Perturbative Quantum Chromodynamics* (ed. A.H. Mueller), World Scientific, Singapore, pp. 1.
17. Aaron, F.D. *et al.* H1 and ZEUS Collab., (2010) JHEP01, p. 109.
18. (a) Gribov, V.N. and Lipatov, L.N. (1972) *Sov. J. Nucl. Phys.*, **15 438**, 675; (b) Altarelli, G. and Parisi, G. (1977) *Nucl. Phys.*, **B 126**, 298; (c) Dokshitzer, Y.L. (1977) *Sov. Phys. (JETP)*, **46**, 641.
19. Levy, A. (2009) *Acta Phys. Polon. B*, **40**, 1775.
20. Airapetian, A. *et al.* (2007) *Phys. Rev.*, **D 75**, 012007.
21. (a) Balitzki, Y. and Lipatov, L.N. (1978) *Sov. J. Nucl. Phys.*, **28**, 822; (b) Kuraev, E., Lipatov, L.N., and Fadin, V. (1977) *Sov. Phys. (JETP)*, **45**, 199.
22. Gribov, L.V., Levin, E.M., and Ryskin, M.G. (1983) *Phys. Rep.*, **100**, 1.
23. Golec-Biernat, K. and Wüsthoff, M. (1999) *Phys. Rev. D*, **59**, 014017.
24. Stasto, A.M., Golec-Biernat, K.J., and Kwiecinski, J. (2001) *Phys. Rev. Lett.*, **86**, 596.
25. (a) McLerran, L.D. and Venugopalan, R. (1994) *Phys. Rev. D*, **50**, 2225; (b) McLerran, L.D. and Venugopalan, R. (1994) *Phys. Rev. D*, **49**, 3352; (c) McLerran, L.D. and Venugopalan, R. (1994) *Phys. Rev. D*, **49**, 2233.

Further Readings

Abramowicz, H. and Caldwell, A. (1999) HERA collider physics. *Rev. Mod. Phys.*, **71**, 1275.

Cahn, R. and Goldhaber, G. (1989) *The Experimental Foundations of Particle Physics*, Cambridge University Press.

Roberts, R.G. (1990) *The Structure of the Proton*, Cambridge University Press.

10
Physics of the Neutron

Klaus Schreckenbach

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

10.1 Introduction

Neutrons are frequent particles in nature as they are part of the atomic nucleus and, for heavier nuclei, provide more than half of the nuclear mass. On the other hand, free neutrons are rather scarce because the neutrons are readily absorbed in contact with matter and, as free particles, they undergo β-decay with a lifetime of about 15 min. The neutron was discovered rather late, by Chadwick in 1932, although it was already argued earlier that a neutral particle might be a constituent of the atomic nucleus.

Since the discovery of the neutron, the field of neutron science has rapidly grown. Today, it covers a broad range of investigations in particle, nuclear, astro-, and condensed matter physics including applications in material analysis, medicine, and energy production. Several types of neutron sources and beams are available for basic science and applications.

As a neutral particle, the neutron can penetrate matter and approach the nucleus without passing through a Coulomb barrier. It is subject to the strong force, weak force, and gravitation. Although, on an average, it is electrically neutral, it has a nonuniform charge distribution and thus interacts as a spin 1/2 fermion with an electromagnetic field by its magnetic moment.

Nuclei not too far away from the range of stable isotopes gain energy by capturing a neutron, with the exception of ^{4}He. The subsequent de-excitation modes of the excited nucleus allow a number of studies in nuclear physics. Neutron capture is also important for the understanding of processes in astrophysics and cosmology.

From nuclear reactors and from spallation sources, intense neutron beams are available for experiments. The neutrons are emitted in these reactions in the megaelectronvolt range but can be moderated to thermal and subthermal energies with relatively small absorption losses, forming a neutron gas of ambient temperature. The corresponding de Broglie wavelengths of thermal neutrons are comparable to atomic distances and phonon wavelength in matter. This feature offers excellent possibilities in studying the structure and phonon excitations of matter by neutron scattering, including the study of magnetic structures. Furthermore, neutron wave properties are the subject of a variety of fundamental studies on matter waves.

The free neutron by itself is an important object of study in particle physics. The semileptonic β-decay of the free neutron is of fundamental importance in weak interaction theory and allows high precision

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

tests of the standard model (SM) of particle physics. Furthermore, the intrinsic structure and symmetry of this baryon is subject of experiments and searches beyond the SM have been carried out, such as electric dipole moment, neutron–antineutron oscillations, and neutron disappearance into a mirror world.

This chapter is focused on nuclear and particle aspects in neutron science and in particular on the mostly used neutron energies in the range from nanoelectronvolts (ultracold neutrons, UCNs) up to several megaelectronvolts. The basic properties of the neutron are outlined and neutron sources and beams are presented. For neutron–nucleus interactions, scattering and neutron capture processes are demonstrated. The field of particle physics with low-energy neutrons has made spectacular progress in the weak interaction theory and the present status is described here. This chapter concludes with a selection of applications.

10.2 Basic Properties of the Neutron

In this section, the basic properties of the free neutron are described. The actual experimentally derived values are regularly updated by the Particle Data Group (PDG, [1]). The relevance for particle physics together with the precision experiments will be given in Section 10.5.

As a fermion with spin 1/2 and as subject to the strong and weak forces, the neutron belongs to the class of baryons. The mass and the strong interaction of the neutron are close to that of the proton and thus the two particles are described as an *isospin* doublet.

10.2.1 Mass

The discovery of the free neutron by Chadwick in 1932 [2] was essentially a mass measurement of a penetrating neutral particle, which was ejected from the atomic nucleus by the reaction $\alpha + {}^{9}\mathrm{Be}$, where the α-source was polonium. In 1930, Bothe and Becker had already observed such matter-penetrating objects, which were weakly absorbed by lead but strongly so by paraffin. Chadwick analyzed this radiation by the recoil on different gas nuclei in an ionization chamber. The recoil energies were measured by signals in the ionization chamber for two different gas fillings, hydrogen and nitrogen, and the experimental values were compatible with a neutral particle of about proton mass and a kinetic energy of 5.7 MeV. Recoil by γ-rays could be excluded. Soon after Chadwick's result, it was generally accepted that the atomic nucleus is composed of positively charged protons and of neutral neutrons of similar mass.

The present precise value for the difference of the neutron mass m_{n} from the proton mass m_{p} is derived from mass spectroscopic measurements of the difference between the mass of the atomic ion of the deuterium D^{+} and the mass of a single charged hydrogen molecule ${H_2}^{+}$, and the nuclear binding energy E_{bind} of the deuteron: $m_{\mathrm{n}} - m_{\mathrm{p}} = m_{\mathrm{D+}} - m_{\mathrm{H2+}} + m_{\mathrm{e}} + E_{\mathrm{bind}}/c^2$, with corrections for the atomic shell binding energies. The nuclear binding energy of the deuteron d could be derived by the recoil-corrected γ-ray energy emitted in the thermal neutron capture reaction on the proton $\mathrm{p} + \mathrm{n_{th}} \rightarrow \mathrm{d} + \gamma$. For the absolute determination of the γ-ray energy, the high-precision double flat γ-crystal spectrometer GAMS 4 at the Institut Laue Langevin (ILL)

research reactor was used with an in-pile kapton target as hydrogen sample, yielding an absolute precision of 0.4 eV for the 2.2-MeV γ-line [3]. The present experimental value is $m_n = 939.565\,379(21)\ \mathrm{MeV}/c^2$ and $m_n - m_p = 1.293\,332\,17(42)\ \mathrm{MeV}/c^2$ [1].

10.2.2 Neutron Matter Wave

The nonrelativistic relation for the de Broglie matter wavelength of the neutron is given by

$$\lambda_n\,(\mathrm{nm}) = \frac{h}{m_n v_n} = \frac{h}{\sqrt{2 m_n E_n}} = \frac{0.0287}{\sqrt{E_n(\mathrm{eV})}} = \frac{397}{v_n(\mathrm{m\,s^{-1}})} \tag{10.1}$$

As an example, thermal neutrons of 0.025 eV energy (peak energy of a neutron gas of 293 K temperature) have a matter wavelength of 0.18 nm (=1.8 Å) and thus – as already mentioned – being close to the atomic distances and phonon wavelength in condensed matter. The very low energy regime of UCN with E_n of ~100 neV exhibits a wavelength of almost macroscopic scale of 0.1 µm.

A variety of fundamental experiments on matter wave characteristics of the neutron were carried out and are only briefly mentioned here. These investigations include experiments with neutron single-crystal interferometers, Fabry Perot interferometry with UCN, and investigations on topological phases (see, for instance, [4]). In recent years, discrete neutron levels in the Earth's gravitational field were observed with UCN, confined in one dimension above a horizontal UCN reflecting plate and the gravitational potential [5, 6]. The first discrete level has a vertical energy of 1.4 peV. For this energy, the classical turning point for neutrons when moving vertically against the gravitational field is 14 µm and the neutron wavelength without gravity is 24 µm. Experiments including dynamical excitations are going on in particular a test of Newton's law of gravitation at small distances [7].

10.2.3 Magnetic Moment

Although the neutron as a whole is a neutral particle, it exhibits a magnetic moment μ_n comparable in absolute number to that of the proton but of opposite sign. In a classical picture, this suggests that the positive charge is more concentrated in the center and the negative charge in the skin of the neutron. In Section 10.5.3.2, this property is discussed in the frame of the quark/parton model. The substantial value of μ_n opens up a wide field in neutron scattering on magnetic structures in solid-state physics.

A first precise experimental value of μ_n was measured by Alvarez and Bloch in 1940, using the Rabi nuclear resonance method on polarized neutrons. The present accurate value was obtained in measuring the Larmor frequency of a neutron in a magnetic field by the Ramsey time-separated oscillatory field magnetic resonance method [8], in comparison to the proton nuclear magnetic resonance (NMR) in the same magnetic field leading to $\mu_n = -0.684\,979\,35(17)\ \mu_p$ [9]. The Ramsey technique is also used in the search for a possible electric dipole moment of the neutron and will be covered in Section 10.5.3.3.

10.2.4 Possible Rest Charge

A nonzero charge of the neutron q_n would have far-reaching consequences for the present assumption of integer charges of

particles (in electron charge units) or, in other words, for the balance of lepton and baryon charges. A very sensitive direct search for a rest charge of the neutron was performed in 1988 [10]. A beam of slow neutrons (about $200\,\mathrm{m\,s^{-1}}$) was passed through a static electric field of $60\,\mathrm{kV\,cm^{-1}}$ perpendicular to the neutron beam and a possible beam deflection of the 9-m-long beam section was investigated by neutron multiple-slit optics. Parabolic mirrors were used for an achromatic neutron imaging of the entrance-to-exit multislits. The result of the experiment is $q_n = (-0.4 \pm 1.1) \times 10^{-21}\,\mathrm{e}$ and thus no rest charge of the neutron was observed.

10.2.5
Free Neutron β-Decay

The only baryon that is lighter than the neutron is the proton. According to the empirical law of conservation of baryon and lepton numbers in the SM together with angular momentum and energy conservation, the free neutron decays into a proton, an electron, and an antineutrino: $n \rightarrow p + e^- + \overline{\nu_e}$ with a kinetic energy release of $(m_n - m_p - m_e)/c^2 = 782.333\,\mathrm{keV}$. The rather small decay probability (lifetime 15 min) is due to this three-particle decay with rather small energy release and due to the small strength of the mediating weak force. Decay branches of the neutron into $\pi^+ e^-$ or three leptons such as e^-, e^+, and ν violating the conservation of baryon and lepton numbers are not observed although the kinetic energy release would be about 1000 times higher.

The free neutron β-decay offers a variety of sensitive tests of the SM. The free neutron lifetime and angular correlations in the emission of decay particles include a variety of parameters that can be compared with the SM prediction at a high level of sensitivity (Section 10.5.2). On the other hand, the β-decay of nuclei, where the neutrons are bound in the atomic nucleus, includes nuclear structure effects and can only be used for particle physics consideration in specific cases such as super-allowed $0^+ \rightarrow 0^+$ β-decays.

The neutron may also decay into a neutral hydrogen and a monoenergetic antineutrino. This decay branch is predicted at level of about 4×10^{-6} relative to the principal decay mode, but has not yet been observed experimentally. A further phase-space-suppressed decay is the radiative decay mode $n \rightarrow p + e^- + \overline{\nu_e} + \gamma$. It was observed in recent experiments at a level of 10^{-3} [1]. These rather small decay branches will not be discussed further in this chapter although they provide valuable information on the symmetry and radiative correction in weak interaction, respectively, once experiments achieve the necessary accuracy.

10.2.6
Summary Table

Table 10.1 summarizes the properties of the free neutron [1]. Several quantities are discussed later in this chapter.

10.3
Neutron Sources and Neutron Beams

Free neutrons can be produced by nuclear fission or fusion and by reactions disintegrating nuclei by irradiation with charged particles (p, d, α, ...) or γ-rays.

10.3.1
Laboratory Neutron Sources

In the laboratory, radioactive neutron sources produce the neutrons either by

Table 10.1 Properties of the free neutron.

Mass	$m_n =$ 939.565 379(21) MeV/c^2	Lifetime	$\tau_n = 880.1(1.1)$ s
Mass difference to proton	$m_n - m_p =$ 1.293 332 2(4) MeV/c^2	Ratio coupling constants	$G_A/G_V = -1.2701(25)$
Spin	$1/2\hbar$	Phase G_A, G_V	$\varphi_{AV} = 180.018(26)^\circ$
Isospin component	$I_z = -1$	$n - \bar{n}$ oscillation	$\tau_{n-\bar{n}} > 8.6 \times 10^7$ s
de Broglie wave length	λ_n (nm) $= 397/v_n$ (m s^{-1})	$n - n'$ oscillation	$\tau_{n-n'} > 414$ s
Magnetic moment	$\mu_n = -1.913\,042\,72(45)$ $\mu_N = -0.684\,979\,35(17)$ μ_p	Quark structure	udd
Electric dipole moment	$d_n < 2.9 \times 10^{-26}$ ecm	Baryon number, Lepton number	$B = 1$, $L = 0$
Electric charge	$q_e < 10^{-21}$ e	Intrinsic parity	$P = +1$

spontaneous fission or in nuclear reactions in a composite source, including fusion.

A frequently used spontaneous fission source is ^{252}Cf with a half-life of 2.65 years and a spontaneous fission branch of 3%. Source strengths of 10^9 n s^{-1} are commercially available. Apart from a laboratory source, it is also used as first start-up neutron source for nuclear reactors. For a safe approach to criticality, the neutron detector signal in the reactor control system should be reasonably strong. In the subcritical regime, the reactor core acts as a neutron multiplier and thus the neutron signal is proportional to the source strength of the external source.

Neutrons from a nuclear reaction can be produced in the laboratory by (α,n) reactions or as photoneutrons from (γ,n) reactions. In the (α,n) reaction, quite often, a mixture of the α-emitters americium or radium with beryllium is composed as neutron source. For the reaction $\gamma + {}^9\mathrm{Be} \rightarrow \mathrm{n} + 2\,{}^4\mathrm{He}$, the energy threshold is as low as 1.7 MeV. For instance, high-energy γ-particles up to 2.8 MeV are emitted in the radioactive β-decay of ^{124}Sb (half-life 60 days), which can be produced by thermal neutron irradiation $^{123}\mathrm{Sb(n,\gamma)}^{124}\mathrm{Sb}$ (natural abundance in antimony: 43% ^{123}Sb). A source of a mixture of antimony and beryllium is "loaded" in the neutron field of a nuclear reactor and is then available as a laboratory neutron source or with high intensity for the start-up of a nuclear reactor, replacing the ^{252}Cf source mentioned above.

A still rather small device is the so-called "neutron generator." It uses the fusion process of deuteron d and triton t: $\mathrm{d} + \mathrm{t} \rightarrow \mathrm{n} + {}^4\mathrm{He} + 17.1$ MeV, releasing monoenergetic neutrons of 14.1 MeV. Similarly, t + t or d + d yield neutrons up to 10 and 2.5 MeV respectively. These neutron sources are available as transportable commercial sources with activity up to 10^9 n s^{-1}.

10.3.2 Neutron Generation in Large-Scale Facilities

Research reactors as neutron source provide steady-state stable neutron densities and beams of high intensities. With a

compact reactor core about 3×10^{16} neutrons per megawatt of thermal power are available in the vicinity around the reactor core (about one of the 2.5 neutrons per fission may leave the core, which gives about 180 MeV thermal load per useful neutron). The neutrons are slowed down in a suitable moderator (Section 10.3.3) surrounding the reactor core and thermal neutron fluxes up to 10^{15} n $(cm^2\ s)^{-1}$ are achieved in accessible parts of the moderator (e.g., for the high-flux research reactors FRM II, HFR at ILL, and HIFIR, the thermal power is 20–85 MW). The limit for the achievable neutron flux is mainly given by the removal of the fission heat from the reactor core. A typical limit is 1 MW l^{-1} core on an average, over the reactor core. Special types of research reactors work also in pulsed mode (TRIGA reactors, IBR-2 in Russia).

The heat problem is relaxed for spallation sources. A high-energy proton beam (500–few GeV) impinges a cooled heavy-element target. The target nucleus is broken up by the proton into many parts and about 25 neutrons are evaporated. In principle, neutron fluxes about 5–10 times higher than for fission reactors are accessible when heat removal is the limitation (about 20–40 MeV thermal load per useful spallation neutron). Most of the existing spallation neutron sources work in pulsed mode with a low repetition rate of 10–60 Hz, to avoid the frame overlap of successive thermal neutron pulses in neutron beam lines that may be as long as 100 m. At present, the strongest pulsed spallation source is the SNS at Oak Ridge with 3 MW proton beam power (envisaged). The cyclotron spallation source SINQ in Switzerland is a semicontinuous source.

Neutrons can also be produced at the beam dump of an electron accelerator, using the bremsstrahlung for the (γ,n) reaction. Also energetic α-particle beams or deuteron beams could be produced by accelerators and initiate the neutron production in targets. Accelerator-based (d,t) sources yield presently 10^{13} n s^{-1} of 14 MeV neutron energy. Furthermore, fusion reactors are intense sources of neutrons.

10.3.3 Neutron Energies and Beams

Neutrons are commonly classified by their energy range as shown in Table 10.2. By elastic scattering at light nuclei with low neutron absorption cross section, neutrons are slowed down without significant losses, a process called *moderation*. Suitable moderators are D_2O, H_2O, beryllium, and graphite. By this nonconservative process, the phase-space density for neutrons is enhanced by many orders of magnitude.

Neutrons in the megaelectronvolts range are called *fast neutrons*. This energy range is

Table 10.2 Classification of neutrons by energy.

Fast neutrons	>1 MeV
Resonance neutrons	Few keV to 1 MeV
Epithermal neutrons	Electronvolt range
Thermal neutrons	Maxwell distribution for $T = 293.5$ K with maximum at 0.025 eV, $v_0 = 2200\ m\ s^{-1}$, wavelength 0.18 nm
Cold neutrons	$v_n \sim 800\ m\ s^{-1}$
Ultra cold neutrons (UCN)	$v_n = 0–7\ m\ s^{-1}$

typically for neutrons produced in neutron-induced fission (then called *fission neutrons*, typical 2 MeV, up to 10 MeV) and at somewhat higher energies (~25 MeV) at spallation sources, before the slowdown in a moderator. For a fast neutron beam, the neutron source has to be viewed directly by an evacuated or helium-filled beam tube without a moderator in between. γ-Filters and absorbers for low-energy neutrons can clean the spectrum.

When fast neutrons are slowed down in a moderator, they pass the kiloelectronvolt range (resonance neutrons) and electronvolt range (epithermal neutrons). Beam tubes for the extraction of these neutrons should be located in the moderator rather close to the neutron source and multi-scattering moderation such as at carbon is preferred. By transmission filters or by time of flight of a chopped beam, monoenergetic resonance or epithermal neutron beams can be made available. As an example, ^{56}Fe has an absorption minimum at about 24 keV due to destructive interference of neutron resonances and thus a 24-keV neutron beam with typically 1–2 keV width can be formed by transmission. In the electronvolt range and below, Bragg scattering at single crystals is used for deflection of monoenergetic neutrons out of the principal neutron beam.

Most widely used are beams of thermal and cold neutrons. When neutrons come into thermal equilibrium by scattering in the moderator, they adopt a Maxwellian energy distribution corresponding to the temperature of the moderator; for 273.2 K, the peak energy is 0.025 eV (0.18 nm = 1.8 Å). The spectrum may be distorted by the energy-dependent absorption of neutrons by the moderator nuclei. Heavy water as moderator with very low content of light water (<1%) is used to form a very clean thermal spectrum. A lower temperature of the neutrons can be achieved in a cold moderator of liquid hydrogen or deuterium. The peak energy is then shifted to about 0.003 eV (0.5 nm = 5 Å). For the extraction of a neutron beam from the moderator, an evacuated or helium-filled beam tube is inserted in the moderator and thus the neutrons can leak out. A tangential geometry of the beam tubes relative

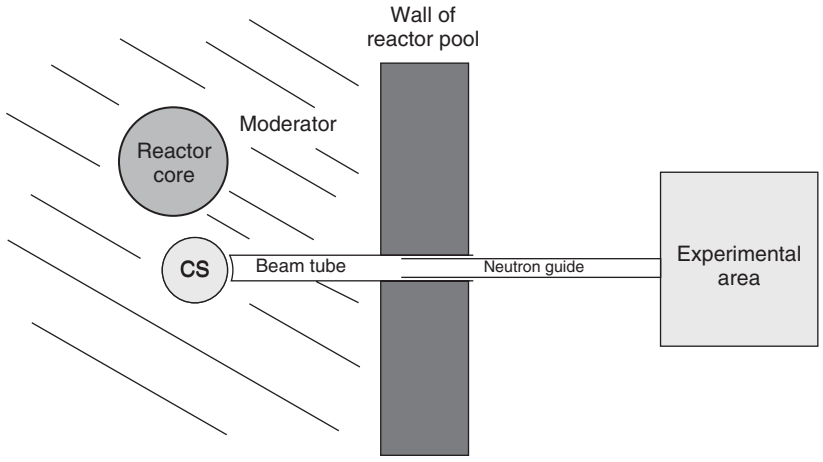


Figure 10.1 Illustration of a cold neutron beam extracted from a research reactor with cold source CS, containing liquid hydrogen or deuterium.

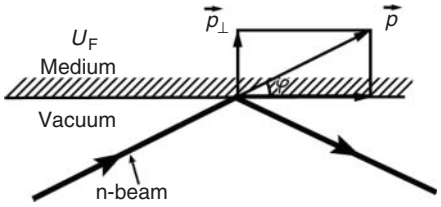


Figure 10.2 Neutron mirror reflection at a material wall with potential U_F for $p_\perp{}^2/2m_n < U_F$, where p is momentum of the neutron.

to the reactor core provides a lower background against faster neutrons and γ-rays from the reactor core (Figure 10.1).

An efficient system for the transport of cold and thermal neutrons to the experimental site are neutron guides and there has been tremendous progress in quality in the last decades. The coherent scattering of neutrons at materials of nuclear density N forms a so-called Fermi pseudopotential U_F, which is attractive or repulsive, Table 10.3, according to the sign of the coherent-scattering length b_c (Section 10.4.2). More generally, the potential has a real and imaginary part, $U_F + iW_F$, where W_F describes the absorption in the scattering process and is of minor importance for neutron guides. A refraction index n_{ref} can be defined as in optics, as the ratio of the neutron velocity in matter v over the velocity in vacuum v_0.

$$U_F = \frac{2\pi\hbar^2}{m_n} N b_c \tag{10.2}$$

$$n_{ref} = \frac{v}{v_0} = \sqrt{\frac{E_n - U_F}{E_n}} = \left(1 - \frac{\lambda_n^2}{\pi} N b_c\right)^{1/2} \tag{10.3}$$

For suitable materials, neutrons are scattered by total reflection at this potential. In Figure 10.2, the reflection at a wall is illustrated. The critical angle up to which total reflection occurs is given by

$$\varphi_c\,(\text{rad}) = \sqrt{\frac{U_F}{E_n}} = \frac{\sqrt{U_F\,(\text{neV})} \times \lambda_n\,(\text{nm})}{908} \tag{10.4}$$

Thus, neutrons are guided over long distances almost loss-free in evacuated tubes coated inside with these materials (minimal coating thickness is about 100 nm), when the scattering angle at the wall surface remains below φ_c along the beam line. As example for neutron guides with natural nickel coating with $U_F = 245$ neV, the critical angle is $\varphi_c = 0.0086\,\text{rad} = 0.5^\circ$ for a cold neutron velocity of 800 m s^{-1} ($\lambda_n = 0.5$ nm). A curved neutron guide (single curve or S-shape) is a filter against faster neutrons. With multilayer coating of continuously varying thickness (alternating Ni and Ti layers), Bragg reflection occurs in addition and the critical angle is increased by a factor m, with m up to 4.

When the neutron energy is lower than the Fermi potential of a trap wall, neutrons can be stored in "neutron bottles" for almost their β-decay lifetime. In this low-energy range, typically <250 neV,

Table 10.3 Fermi pseudopotential U_F for some materials (stainless steel, DLC diamond-like carbon).

Material	Ni	^{58}Ni	Be, BeO	Cu	Steel	Al	V	Ti	SiO_2	Fomblin	DLC
U_F (neV)	245	340	260	170	190	50	−7	−50	110	106	260

corresponding to velocities $<7\,\mathrm{m\,s^{-1}}$, they are called *UCNs*. The production of intense UCN sources is rather difficult. For many years, the UCN facility at the high flux reactor at ILL provides the highest UCN yield, with typical UCN densities of ~50 UCN $\mathrm{cm^{-3}}$ at the UCN beam ports. At ILL, cold neutrons fly in a vertical beam tube upward from the liquid deuterium source and then they are almost stopped by elastic scattering at the blades of a "neutron turbine," turning in the downstream direction [11]. Recent developments of the so-called superthermal UCN sources, where UCNs are not in thermal equilibrium with the medium, promise a yield of higher UCN densities. A converter at very low temperature is used, where cold neutrons are down-scattered by phonon excitation into the UCN range, but where the probability for up-scattering of the UCN is low because of lack of phonons in the cold converter (principle proposed in [12]). For this purpose, liquid ^{4}He at ~0.5 K or solid deuterium at ~5 K is well suited. Several UCN sources of this type are already operational with UCN densities similar to that of the UCN turbine but with the potential for significant improvements, in particular, for UCN densities in smaller UCN bottles.

10.3.4 Neutron Polarization

An important development for neutron physics is the formation of polarized neutron beams. For this purpose, scattering at the magnetized ferromagnetic surfaces with potential $V = U_F + \boldsymbol{\mu}_n\,\mathbf{B}$ or the spin-dependent cross section for neutrons at polarized nuclei is used. The optimum case for surface scattering is in analogy to optics: The Fermi and magnetic potential compensate to zero for one neutron spin direction and hence the layer is transparent for neutrons without mirror reflection, while for the opposite neutron polarization, the potentials add and the neutrons are reflected. A very high degree of polarization of up to 99% is achieved by bent magnetized multilayers ("super-mirrors"). Transmission polarizers use the spin-dependent cross section of the ^{3}He(n,p)t reaction, for thermal neutrons 10 660 and ~0 b, respectively. The ^{3}He gas is polarized by optical pumping in a magnetic field. In addition, the spin-dependent scattering cross section at polarized protons is used in transmission geometry [13]. A further polarization method is based on single-crystal reflection of neutrons up to the electronvolt range on magnetized Heusler-alloy single crystals.

For the spin flip in a polarized neutron beam, current sheets or resonant spin flippers are used, yielding a spin-flip efficiency of almost 100%.

Typical polarized neutron beam lines with spin-flip option are shown for specific experiments in Sections 10.4.4 and 10.5. The easy change of the polarization of the neutrons without changing the neutron beam profile and intensity allows precise comparative measurements. As state of the art, differences in the neutron reaction R at the target site between the two neutron polarization states as low as a few 10^{-8} can be detected in the asymmetry $(R_+ - R_-)/(R_+ + R_-)$. In addition, the absolute degree of polarization can be determined very precisely, almost at a level of 10^{-4}, using opaque ^{3}He spin filters [14].

10.3.5 Neutron Detection

Slow neutrons are detected by nuclear reactions emitting charged particles that in turn are detected as ionizing particles. The most frequently used reactions with

their thermal neutron cross sections are

$^{3}\mathrm{He} + \mathrm{n} \rightarrow \mathrm{p} + \mathrm{t} + 0.77\,\mathrm{MeV}$ 5330 b
$^{10}\mathrm{B} + \mathrm{n} \rightarrow {}^{7}\mathrm{Li} + \alpha + \gamma + 2.78\,\mathrm{MeV}$ 3840 b
$^{6}\mathrm{Li} + \mathrm{n} \rightarrow \alpha + \mathrm{t} + 4.8\,\mathrm{MeV}$ 940 b
$^{235}\mathrm{U} + \mathrm{n} \rightarrow$ fission products 586 b

Furthermore, natural gadolinium with the high (n_{th},γ) cross section of 48 800 b (^{157}Gd with 254 000 b) absorbs thermal neutrons in thin layers, while the produced internal conversion electrons and Auger electrons may leave the layer into the counter.

The reaction nuclei are added either to scintillators or to the walls or gas of gas-filled counters. Various multidetector devices are in use such as arrays of single detectors, multiwire proportional counters (MWPC), and cascade detectors. Fast neutrons are detected by their recoil on protons in an organic scintillation detector or they are moderated before the detecting reaction of the slowed-down neutron.

10.4 Neutron Interactions with the Atomic Nucleus

The basic nuclear interactions of neutrons n with the atomic nucleus comprise

- elastic scattering (n,n), including potential scattering and resonance scattering;
- inelastic scattering (n,n′), nuclear excitations up to energies $E_n - E_{n'}$ [15];
- nuclear reactions (n,x), where x stands for γ-radiation and associated effects such as internal conversion and internal pair production, for light particle emissions like protons and α-particles, and for fission.

A typical excitation mode is the compound-state formation in the resonant neutron capture process. The absorbed neutron in a nucleus of mass A produces an excited nucleus $(A+1)^*$, with typical excitation energies of 2–10 MeV according to the neutron binding energy in the nucleus and the neutron kinetic energy. The lifetime of these excited states is in the order of 10^{-14} s, where, as per definition of a compound state, the exit channels are decoupled from the entrance channel. The neutron can be emitted again (resonance scattering) or the de-excited nucleus $(A+1)$ is formed by the emission of γ-rays or associated particles.

A fundamental case is the absorption of neutrons by protons or neutrons. In analogy to the isospin formalism, the binding energy of np and nn should be similar. Only the state of the deuteron np with nuclear spin $I = 1$ (neutron and proton spin parallel) is bound, with 2.2 MeV binding energy. The corresponding dineutron $I = 1$ state is forbidden by the Pauli principle and, in analogy to the nonexistence of the $I = 0$ bound state of the deuteron, no bound dineutron $I = 0$ state is observed, but only a few kiloelectronvolts in binding energy seem to be missing.

10.4.1 Neutron Resonances

For heavier nuclei, the neutrons are absorbed into nuclear states of the compound nucleus above the neutron binding energy. A comprehensive atlas for neutron resonances of nuclei has been compiled by Mughabghab [16], including the relevant equations. At the energy of a single nuclear state of the compound nucleus, the cross section is resonance-like with a typical width of $\Gamma \sim 0.1$ eV, corresponding to the lifetime of $\tau = \hbar/\Gamma \sim 10^{-14}$ s and follows

the Breit–Wigner theorem:

$$\sigma(x,y) = \frac{\lambda_x^2}{4\pi} \frac{g\Gamma_x\Gamma_y}{(E_x - E_0)^2 + (\Gamma/2)^2} \tag{10.5}$$

$$\text{statistical factor } g = \frac{2J+1}{(2s+1)(2I+1)} \tag{10.6}$$

with s spin of x, I target spin, J spin of the resonance, x,y ingoing and outgoing reaction particles, λ_x wavelength, Γ total width, $\Gamma_{x,y}$ partial width, and E_0 resonance energy.

The exit channels have the above-mentioned possibility and, for instance, for a fissile isotope such as ^{235}U, neutron scattering, (n,γ) and (n,f), f for fission, are competing. The total decay width is then the arithmetic sum over the partial widths: $\Gamma = \Gamma_n + \Gamma_\gamma + \Gamma_f$. In the following, some aspects of the exit channels are described without going into details of nuclear structure.

10.4.2 Resonance and Potential Scattering

Firstly, the elastic scattering of the neutron at the atomic nucleus is discussed. For resonance scattering, Eq. (10.5) with $x = y = n$ leads to

$$\sigma(n,n) = \frac{\lambda_n^2}{4\pi} \frac{g\Gamma_n^2}{(E_n - E_0)^2 + (\Gamma_n/2)^2} \tag{10.7}$$

At the resonance energy $E_n = E_0$ the scattering cross section is

$$\sigma_0(n,n) = g\frac{\lambda_n^2}{\pi} \tag{10.8}$$

the so-called unitarian limit, where the amplitude of the resonance depends only on the statistical factor g, but not on the nuclear structure of the compound state.

Away from the resonance and for low neutron energies ($\Gamma_n \ll E_0$; $E_n \ll E_0$), the scattering cross section is independent of the neutron energy, because the neutron wavelength λ_n is proportional to $1/v_n$ and the phase-space term in gamma Γ_n is proportional to v_n. The absolute value of resonance cross sections varies strongly from isotope to isotope due to the nuclear structure dependence of Γ_n.

Potential scattering is defined as scattering of neutrons at the nuclear surface and thus depends on the *radius R* of the nucleus. In contrast to resonance scattering, it has a smooth dependence on the atomic number and basically results for a spherical nucleus in

$$\sigma_{pot} = 4\pi R^2 \tag{10.9}$$

In general, the total neutron scattering cross section is an interference of these two scattering modes and is described by the scattering length b as the amplitude of the outgoing radial wave of the scattered neutron. In particular, the coherent scattering length b_c is measured via the refraction index Eq. (10.3) using neutron interferometers or by the critical angle in the scattering at the Fermi potential U_F, Eq. (10.4) [17]. Experimental values are listed in Refs [17, 18]. Depending on the sign of the amplitudes, the scattering length may also be negative.

Furthermore, a correction term between the scattering length a at a free nucleus and the scattering length b at a bound nucleus has to be considered: $a = b \times A/(A+1)$. This center of mass correction is obviously most important for light nuclei.

For the scattering at an assembly of nuclei in condensed matter, the coherence and incoherence effects must be taken into account, as the assembly of scattering nuclei may be composed of different isotopes and spin directions.

For example, for the same nuclei but with different spin directions, the following scattering probabilities result. b_+ and b_- denote the scattering length for neutron parallel and antiparallel to the nuclear spin direction, respectively. They depend on nuclear structure and can be quite different. The statistical weight for the nuclear spins I is given by $w_+ = (I+1)/(2I+1)$; $w_- = I/(2I+1)$; $w_+ + w_- = 1$.

For the coherent part, the amplitudes are superposing (Eq. (10.10)). The total cross section is given by the quadratic sum and the incoherent cross section, Eq. (10.11), by the difference of these two terms:

$$\sigma_c = 4\pi(w_+ b_+ + w_- b_-)^2 \qquad (10.10)$$

$$\sigma_{inc} = 4\pi w_+ w_- (b_+ - b_-)^2 \qquad (10.11)$$

For the target spin $I = 0$, the cross sections are $\sigma_{coh} = 4\pi b^2$ and $\sigma_{inc} = 0$. As an example, for $I = 1/2$, the fundamental scattering process at an assembly of protons with statistically distributed proton spins is given here. The scattering lengths are $b_+ = 1.07 \times 10^{-12}$ and $b_- = -4.75 \times 10^{-12}$ cm and $w_+ = 3/4$, $w_- = 1/4$. Thus, the values for the scattering of low-energy neutrons (S-wave scattering) at bound protons result in $\sigma_{coh} = 1.9$ and $\sigma_{inc} = 80$ b. The complementary neutron–neutron scattering length has not yet measured and only deduced from neutron scattering at the deuteron.

A special case is that of ortho- and parahydrogen, where the protons in the H_2 molecule are orientated parallel and antiparallel, respectively. The corresponding cross sections are quite different. Thus, polarized proton targets for neutron physics can be prepared accordingly and also used as transmission polarizers.

10.4.3 Neutron Capture Reactions

The neutron may undergo a nuclear reaction with the nucleus. The following equations are shown for the (n,γ) reaction but they are equivalent for the (n,f) process. Equation (10.5) with $x = \mathrm{n}$ and $y = \gamma$ leads to

$$\sigma(\mathrm{n},\gamma) = \frac{\lambda_n^2}{4\pi} \frac{g\Gamma_n\Gamma_\gamma}{(E_n - E_0)^2 + ((\Gamma_n + \Gamma_\gamma)/2)^2} \qquad (10.12)$$

For the case of low neutron energies $E_n \ll E_0$, far off-resonance (Γ_n, $\Gamma_\gamma \ll E_0$) and as Γ_γ is independent of E_n ($E_\gamma \gg E_n$), Eq. (10.12) results in

$$\sigma(\mathrm{n},\gamma) \propto \lambda_n^2 \Gamma_n \propto \frac{1}{v_n} \qquad (10.13)$$

This is the so-called $1/v_n$ neutron absorber in the absence of close resonances and at relative low neutron energies (typically, up to epithermal energies, depending on the level density of the nucleus above neutron binding energy). For a qualitative understanding, the interaction time is longer for slow neutrons passing the nucleus.

The excitation energy of the $(A+1)^*$ nucleus may be released in the form of a cascade of γ-rays to the ground state $(A+1)$. The so-called primary γ-rays from the $(A+1)^*$ state directly feed the nuclear levels of energy $E_{level} = E_{n\_bind} + E_{n\_kin} - E_\gamma$. The γ-strength from individual neutron resonances to the excited nuclear states depends on the nuclear structure of the states involved. On an average, $E1$ transitions are preferred to $M1$ transitions and transitions get different statistical weights due to the nuclear spins involved. An elegant method for the detection and spin and parity assignments

of nuclear levels fed by primary γ-rays was developed in the 1970s, in particular at the Brookhaven National Laboratory (BNL). The BNL research reactor was a graphite-moderated neutron source offering high intensities of resonance and epithermal neutrons. Using neutron filters, a monoenergetic neutron beam of 2 keV (transmission of Sc) and 24 keV (transmission of ^{56}Fe) was formed with a width of 1–2 keV. For heavier elements, the level spacing above the neutron binding energy is in the order of a few 10 eV and thus a kiloelectronvolt neutron band can excite a statistical number of levels. The average resonance capture (ARC) averages the individual nuclear structure of the initial and final states and feeds the final nuclear levels by distinctive relative intensities, depending only on possible initial and final spins and parities. Thus, a complete set of nuclear states with their spins and parities within the accessible spin range is detected as an important base for nuclear structure work [19].

A comprehensive monograph on neutron capture work can be found in Ref. [20], which summarized the work done at and in collaboration with the ILL. It includes investigations in thermal neutron capture reactions (n,γ), (n,e), and (n,*f*) performed at the high-flux reactor of that institute and associated nuclear structure experiments elsewhere.

10.4.4 Parity Nonconservation (PNC) in Neutron–Nuclear Interactions

Weak interaction terms in neutron–nuclear reactions can be detected by parity-non-conserving symmetries in the reaction, as the weak force is the only one known to violate parity. Parity is a discrete symmetry operation that mirrors the space coordinates at its origin $\mathbf{r} \rightarrow -\mathbf{r}$. If the mirrored reaction system is different from the original one, parity is not conserved.

The well-controlled polarization of a neutron beam is an excellent tool to study even tiny parity nonconservation (PNC) effects. In a capture process, the compound state is polarized in space according to the incoming polarization of the neutron beam and the nuclear ground state spin. The emission of γ-rays, fission products, or other particles relative to the direction of the neutron spin is studied. Furthermore, the scattering of neutrons at the nuclei includes weak forces and experiments have found corresponding effects. Figure 10.3 illustrates different geometries and PNC effects in neutron–nuclear reactions.

A fundamentally interesting system is the interaction of polarized neutrons with protons. PNC in the capture process $\mathrm{p}(\vec{\mathrm{n}},\gamma)\mathrm{d}$ was searched for and PNC spin rotation search is envisaged. When polarized neutrons are captured by unpolarized protons, the 2.2-MeV γ-radiation may be emitted asymmetrically relative to the incoming neutron spin direction. This PNC is in analogy to the first observation of PNC by Wu *et al.* where β-particles are emitted asymmetrically to the ^{60}Co spin direction. Efforts to see PNC in $\mathrm{p}(\vec{\mathrm{n}},\gamma)\mathrm{d}$ achieved a sensitivity of 5×10^{-8} for the asymmetry of the detected numbers of gammas N_+ and N_- between the two spin directions of the neutron; asymmetry $= (N_+ - N_-)/(N_+ + N_-)$ [21]. Further investigations with envisaged higher sensitivity are ongoing (NPD gamma collaboration [22]). The PNC effect tests the weak interaction pion exchange force in the n–p interaction and predictions for the asymmetry are in the range of a few 10^{-8}. For heavier nuclei, PNC effects with polarized neutrons in the capture process or in transmission experiments were already

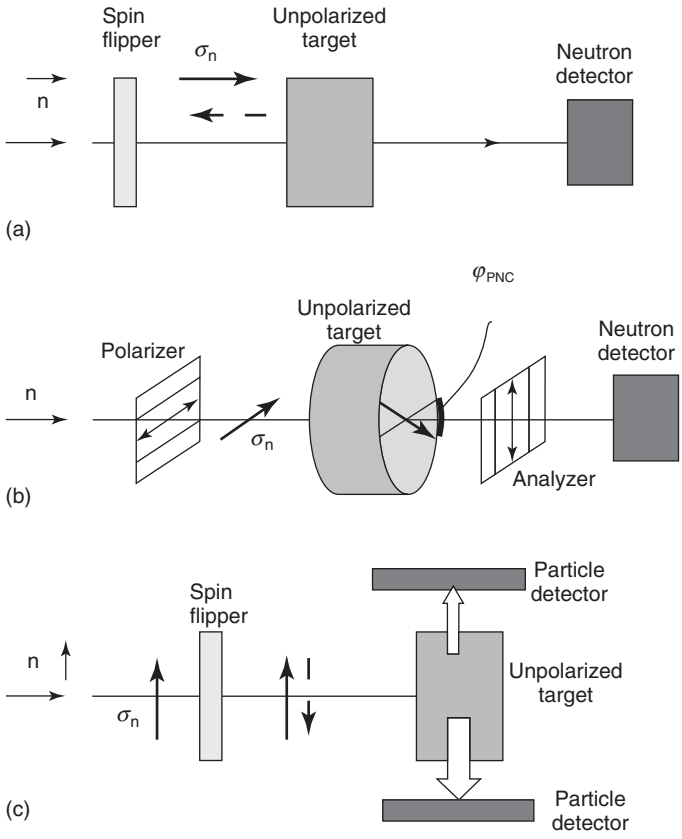


Figure 10.3 Schemes for PNC searches with polarized neutron beams. For (a) and (c), the difference by flipping the spin is measured; for (b), the spin rotation. (a) Dichroism in transmission, (b) PNC spin rotation, and (c) asymmetry in the reaction products.

observed. PNC spin rotation was first observed in Sn [23]. The PNC effect can be enhanced by the mixing of close-lying neutron resonances of opposite parity, as the mixing matrix element is proportional to $1/\Delta E$, where ΔE is the level distance, which is likely to be only in the range of a few electronvolts or may even overlap for heavier nuclei. As an example, the PNC dichroism for ^{239}La with polarized neutrons of $E_n = 0.734$ eV is in the percentage range, as the neutron energy coincides with a p-wave neutron resonance in that nucleus [24]. The PNC dichroism was also investigated for many neutron resonances by the time-of-flight method with an accelerator-based pulsed source of polarized neutrons [25].

A number of PNC effects were observed in fission with polarized neutrons. The fissile nucleus is excited and polarized after polarized neutron capture and an asymmetry of light fission products are investigated relative to the compound spin direction (see, for instance, [20]). In recent experiments, ternary fission was also studied and here the final state interaction

was investigated in a triple correlation geometry [26] similar to the *D*-coefficient in free neutron decay (Section 10.5.2.2).

The enhancement of the PNC effect in heavy nuclei allows clear detection, but fundamental weak interaction exchange terms are difficult to derive because of nuclear structure effects. The size of the effect can only be understood on the average over many nuclei and resonances.

10.5 Neutrons in Particle Physics

The field of particle physics with neutrons has rapidly developed in recent years with the availability of intense neutron beams and new experimental methods. Recent reviews illustrate the present status of the field [27, 28] and a comprehensive review [29] also includes the role of the neutron in astrophysics and cosmology. The precision achieved, particularly with experiments with cold neutrons and UCN, allows sensitive tests on the SM of particle physics and is competitive with complementary high-energy experiments. In this section, the neutron β-decay is presented in view of the SM with corresponding experimental investigations. Furthermore, experiments on the internal structure of the neutron and on possible neutron oscillation experiments are outlined.

10.5.1 Quark Structure of the Neutron

The observation of a magnetic moment of the neutral particle, the neutron, points toward an inhomogeneous distribution of the charge in the neutron and hence an internal structure. The famous experiments by Hofstadter *et al*. [30] revealed form factors for the proton and neutron in the scattering of 500 MeV electrons, which were incompatible with a single pointlike charge distribution of the nucleons. Subsequent experiments at higher energies with electrons and neutrinos showed a structure of three pointlike "partons" for the proton and the neutron. The quark model was developed and the neutron and the proton were interpreted as composite objects of three quarks of the first generation bound by gluons. The present SM describes elementary particles by three generations of quarks and leptons. Interaction by the strong, weak, and electromagnetic forces is mediated by gluons, W,Z bosons, and gammas, respectively.

In the quark picture, the neutron has the composite structure udd and the proton uud, with charges $q_u = +2/3$ e and $q_d = -1/3$ e. The quark and nucleon spins are 1/2 and identical quarks are coupled in parallel, where the Pauli principle of an antisymmetric wavefunction for identical particles is conserved by the antisymmetric wavefunction of the color quantum number of the quarks.

10.5.2 β-Decay of the Free Neutron

The β-decay of the free neutron is a fundamental process of weak interaction for semileptonic decays in the first generation of elementary particles. The dominant branch is

$$n \rightarrow p + e^- + \overline{\nu_e} \qquad (10.14)$$

The kinetic energy release in the decay is 782.333 keV, with a maximal energy of 781.582 keV for the electron and only 0.751 keV for the proton (for antineutrino mass ~0). The free neutron decay is an interference of a Fermi β-decay mode with vector-coupling constant G_V ("no

spin flip from neutron to decay proton,'' $\Delta L = 0$) and a Gamov–Teller β-decay mode with axial vector coupling constant G_A (''spin flip from neutron to decay proton,'' $\Delta L = 1$). The ratio is denoted $\lambda = G_A/G_V$ $e^{i\varphi}$ with the phase φ_{AV} between G_v and G_A observed as 180°. Thus, G_v and G_A are opposite in sign (V–A theory).

In the quark picture, the neutron decay corresponds to the transformation of a d quark into a u quark under emission of a W^- boson mediated by the weak interaction. The W^- boson decays into an electron and an antineutrino. The corresponding Feynman diagram reads as shown in Figure 10.4.

In the SM, the weak d', s', b' and mass d, s, b eigenstates of the quarks are different and connected by the unitary Cabbibo–Kobayashi–Maskawa (CKM) matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{10.15}$$

The unitarity condition for the first line requires

$$V_{ud}^2 + V_{us}^2 + V_{ub}^2 = 1 \tag{10.16}$$

Therefore, the weak interaction coupling constants have to be multiplied by V_{ud} to describe the decay strength of the free neutron decay. By comparing experimental values of the weak force deduced from purely leptonic decays such as the muon decay, the matrix element V_{ud} can be deduced. Furthermore, the individual values of the coupling constants G_V and G_A can be determined from the neutron decay alone by neutron lifetime and neutron decay angular correlation experiments. Owing to the conserved vector current (CVC) theory, G_V from the neutron should have the same value as G_V from super-allowed $0^+ \rightarrow 0^+$ β-decays. In general, the various angular correlations give insight into the symmetry of weak interaction processes. At present, these sensitive tests are in accordance with the SM and provide stringent limits for theories beyond the SM [29]. A future multipurpose device PERC [31] for free neutron decay products and a new generation of neutron lifetime experiments are supposed to improve considerably the precision of the experimental values.

10.5.2.1 Lifetime of the Free Neutron

Several specific review articles on the neutron lifetime τ_n are available [32–34] in addition to the above-mentioned more general articles on neutron particle physics.

Apart from the weak interaction theory aspects, the value of τ_n is important for a variety of applications. The inverse reaction $\overline{\nu_e} + p \rightarrow n + e^+$ is used for the

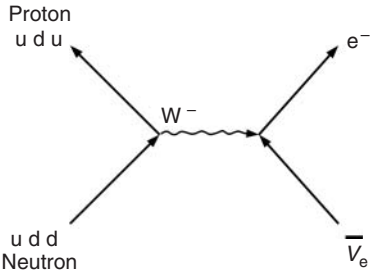


Figure 10.4 Feynman diagram for the β-decay of the neutron.

detection of antineutrinos in hydrogen-containing scintillators in neutrino oscillation experiments. The corresponding cross section of antineutrinos on the hydrogen nucleus proton is obviously proportional to $1/\tau_n$ [35]. Furthermore, processes in cosmology such as the balance between n, p, e^+, e^-, ν, $\overline{\nu}$ in the early universe include the neutron lifetime value. The formation of the first elements depends on τ_n because neutron decay competes with neutron capture on protons in the expanding and cooling down of the early universe. The $p + p \rightarrow d + e^+ + \nu_e$ fusion reaction in the sun is mediated by the axial vector-coupling constant G_A from the neutron apart from strong interaction corrections at the level of a small percentage. Only G_A comes into play here because only protons with antiparallel spins interact (the Pauli principle) to form the deuteron with parallel spins of the nucleons.

Within the V−A theory, the neutron lifetime is related to the coupling constants by

$$\tau^{-1} = \frac{c}{2\pi^3}\frac{(m_e c^2)^5}{(\hbar c)^7}(G_V^2 + 3G_A^2)f = \frac{c}{2\pi^3}\frac{(m_e c^2)^5}{(\hbar c)^7}G_F^2|V_{ud}|^2(1+3\lambda^2)f \tag{10.17}$$

The statistical weight at the coupling constants is $2L+1$ for the angular momentum transfer ΔL of 0 and 1, respectively. The coupling constant G_A is different from G_V owing to strong interaction terms in the $\Delta L = 1$ angular momentum transfer. The quantity f denotes the phase-space factor. The numerical value of Eq. (10.17) was evaluated by Ref. [36] including higher-order corrections such as radiative corrections and weak magnetism, with the Fermi constant $G_F = 1.166\,378\,7(6) \times 10^{-5}$ $(\hbar c)^3$ GeV^{-2} as determined from the purely leptonic decay of the muon:

$$\tau_n = \frac{4908.7(1.9)\,\mathrm{s}}{|V_{ud}|^2 \times (1+3\lambda^2)} \tag{10.18}$$

The free neutron decay lifetime was measured essentially by two methods. In the first method, the rate of the decay products n_d in the observed decay volume of a neutron beam and the number of neutrons in the decay volume are measured, according to $n_d = -dN/dt = N/\tau_n$, where the number of neutrons N equals to neutron density times decay volume. The neutron density can be deduced from the reaction rate n_R in a thin foil of a $1/v$ absorber, which is placed downstream behind the decay volume of length ℓ. Conceptually, the neutron lifetime is then given by

$$\tau_n = \frac{N}{n_d} = \frac{n_R}{m\sigma_0 v_0}\frac{\ell}{n_d} \tag{10.19}$$

where m denotes the mass density of the thin foil (nuclei per area unit) and σ_0 the reaction cross section at neutron velocity v_0. A group at the *national institute of standards* (NIST) achieved a precise experimental value for τ_n by the storage and measurement of the decay protons in a trap of variable length and a careful measurement of the neutron flux density of the beam. Figure 10.5 illustrates the experimental setup used by the NIST group. To avoid the border effect of the proton trap, the proton rate was measured as function of the length of the decay volume and thus n_d/ℓ in Eq. (10.19) could be deduced. The derived value is $\tau_n = 886.3(3.4)$ s [37].

In the second method, UCNs are stored in a trap and the decay of the initial neutron population N_0 is observed: $N/N_0 = \exp(-t_{st}/\tau_n)$. For this purpose, the trap is filled with UCNs and then emptied into a UCN detector after the storage time t_{st}. The filling and emptying is repeated

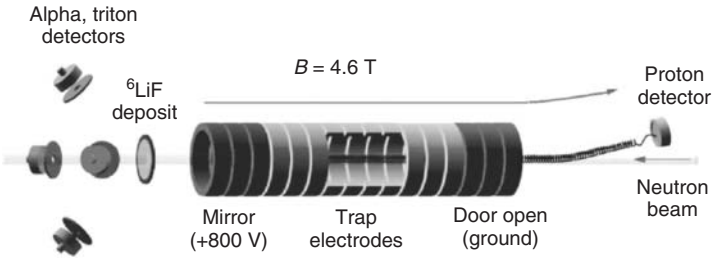


Figure 10.5 Scheme of the neutron beam τ_n experiment performed at NIST [37]. The decay protons are stored in the Penning trap of variable length and then periodically emptied to the proton detector. The absolute ^{6}Li(n,α)t reaction rate in the ^{6}LiF deposit was measured to derive the neutron density in the beam.

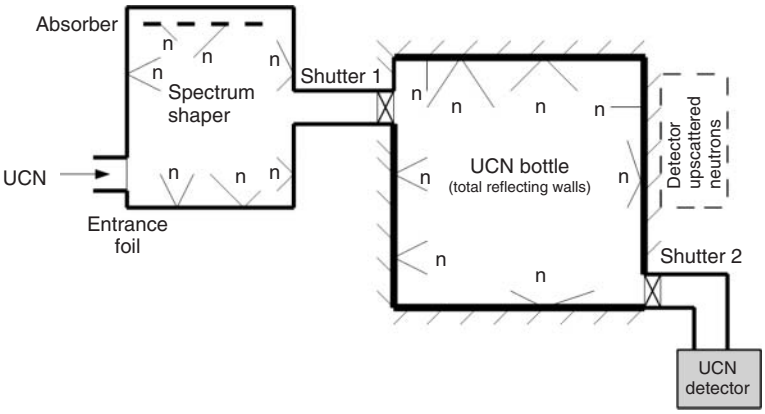


Figure 10.6 Principle of a τ_n measurement by storage of UCN. Low UCN energies are cut off by U_F of the entrance foil. The high-energy end can be limited in a shaper using the gravity potential and absorbers [38, 39]. In [40], the upscattered UCNs, which leave the bottle, were detected.

with the same initial value N_0, but different storage times. The basic setup of such an experiment is shown in Figure 10.6. In the ideal case of no losses in the trap other than by β-decay and identical initial neutron population N_0, the exponential decay law directly yields the free neutron lifetime. In this method, it is not necessary to know the absolute value of N_0. In reality, additional losses of the stored UCN must be determined and in first order, the exponential decay is faster by the loss probability $1/\tau_{loss} : \tau_{st}^{-1} = \tau_n^{-1} + \tau_{loss}^{-1}$.

The trap potential can be the Fermi pseudopotential of suitable trap walls, magnetic gradient fields, or in vertical dimension gravity. The potential height for UCN is typically ∼100 neV for the Fermi potential of walls and, for the magnetic forces, 60 neV T^{-1} change in the magnetic field. Gravity provides a potential of 103 neV m^{-1} height for the neutron mass.

The first precise UCN storage experiment for the neutron lifetime with a precision of a few seconds was carried out in 1989 (device ''MAMBO'' [41]).

An improved version (''MAMBO II'' [38]) shaped the UCN spectrum prior to filling in the main trap. To guarantee a homogeneous trap wall for UCN, the walls were covered with a hydrogen-free oil, FOMBLIN, with a UCN wall potential of 106 neV [42]. Losses in wall collisions occur due to up-scattering of the UCN above the Fermi potential and due to absorption in the trap wall. The loss probability in these experiments corresponded to about 10^{-5} per wall bounce. By varying the bottle size, the bouncing rate was varied and by an interpolation to infinite trap size and hence zero bouncing rate, the neutron lifetime could be deduced. By the UCN storage experiment ''Gravity Trap'' built at PNPI, Russia, and performed at ILL, a storage time close to τ_n was achieved with -150 °C cold PFPE oil-covered trap walls, and the extrapolation to the free neutron lifetime was substantially smaller [39]. In another UCN storage experiment, the up-scattered UCNs were counted in addition outside of the bottle [40] as illustrated in Figure 10.6. For high precision, a number of effects have to be considered. The UCN velocity-dependent loss probability in collisions with walls causes a change of the UCN spectrum during storage and hence τ_{loss} is not constant during the storage time. This effect can be handled such that for the different bottle sizes, the experimental storage times are scaled to equal number of bounces and the corresponding τ_{st} are compared [38, 41]. Further effects arise from the change of the UCN density in the trap with height due to gravity and losses by residual gases in the trap vacuum.

A magnetic storage of very slow neutrons was achieved in 1989 in the magnetic multipole torus NESTOR using superconducting coils (up to 1.2 T cm^{-1} gradient). Following an initial cleaning procedure after filling, the trap was supposed to be loss-free. The low neutron density in the trap limited this method and 877(10) s was measured [43]. Magnetic gradient storage was achieved recently by permanent magnets but final τ_n have not yet been published. Several upcoming experiments will use magnetic storage and aim at substantially higher precision for τ_n but wait for stronger UCN sources; see reviews [29, 33, 34].

The mean value for the experimental neutron lifetime is published by the PDG [1] as 880.1(1.1 s), with improved consistency of the values compared to those obtained in earlier years.

10.5.2.2 Angular Correlations in the Free Neutron β-Decay

Firstly, the angular correlation between the neutron spin and the direction of the electron is described. The angular distribution $W(\theta)$ is given by

$$W(\theta) \propto \left(1 + \frac{v_n}{c} A \cos\theta\right) \qquad (10.20)$$

with θ the angle between neutron spin and electron emission. In the V–A theory, one finds:

$$A = \frac{-2\lambda(\lambda+1)}{(1+3\lambda^2)}; \quad \lambda = \frac{G_A}{G_V} \qquad (10.21)$$

The electron asymmetry parameter A corresponds to the common manifestation of parity violation in β-decay.

Detectors of different kinds were used to detect the electrons emitted from polarized neutrons. A time projection chamber (TPC) for tracing the decay electrons or strong magnetic fields to transport the electrons to the detectors was applied. The later method with longitudinal or transversal magnetic fields produced by superconducting coils yielded the most precise experimental A

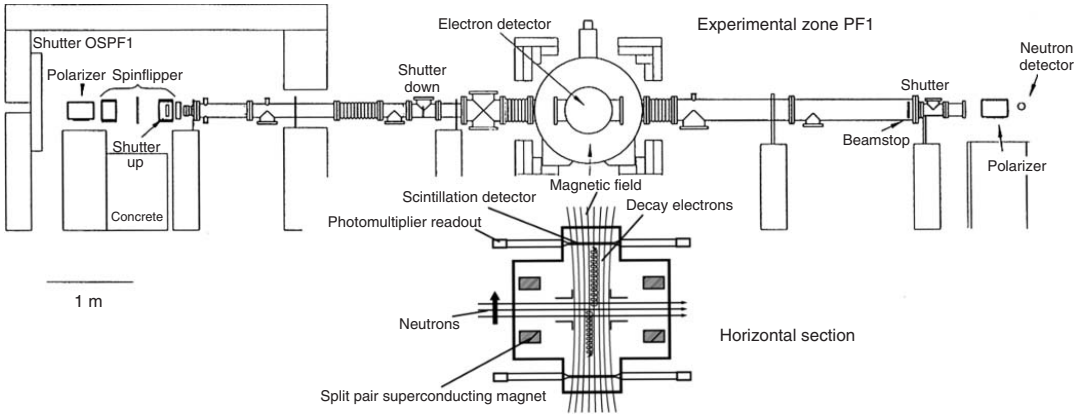


Figure 10.7 Perkeo II setup at the polarized neutron beam PF1 of the ILL. The magnetic guide field for the decay electrons (about 1 T) and the neutron spin are perpendicular to the neutron beam axis. The neutron polarization direction is frequently inverted by the spin flipper for a comparative measurement of the electron events in the two opposite scintillation electron detectors.(From [47].)

values. The backward–forward asymmetry of the decay electrons relative to the magnetic field lines is determined as function of electron energy measured in plastic scintillators. For a cold neutron beam experiment, the Perkeo II setup in a polarized neutron beam is shown in Figure 10.7. Recently, UCNs from a superthermal solid deuterium source were used for the measurement of A; see Figure 10.8 [44]. The PDG [1] value for $A = -0.1176(11)$ is governed by the results of Perkeo II [45] and UCNA [44]. A new value for A obtained with the spectrometer Perkeo II is given in Ref. [46].

Other experiments determine the angular correlation between the neutron spin and the antineutrino coefficient, B. Here, the correlation is deduced from the proton and electron emission angles. The coefficient B is sensitive on possible right-handed currents and associated masses beyond the SM. The a-coefficient describes the angular correlation between the electron and the antineutrino and is commonly deduced from a precise measurement of the spectral distribution of the decay protons. The a-coefficient yields a different way to determine λ. A comprehensive discussion of the relations between the different angular coefficients is given in [28].

The evaluation of the experiments on A led to an average value $\lambda = -1.2701(25)$ [1]. Combining this value with the experimental neutron lifetime allows an individual determination of the coupling constants from the neutron decay alone and of V_{ud} using Eqs. (10.18) and (10.21). By super-allowed β-decays of nuclei, the vector coupling constant G_V can also be derived and compared with that of the neutron. Furthermore, $V_{ud} = G_V/G_F$ from the neutron decay is a test for the unitarity of the first line of the CKM matrix, where the

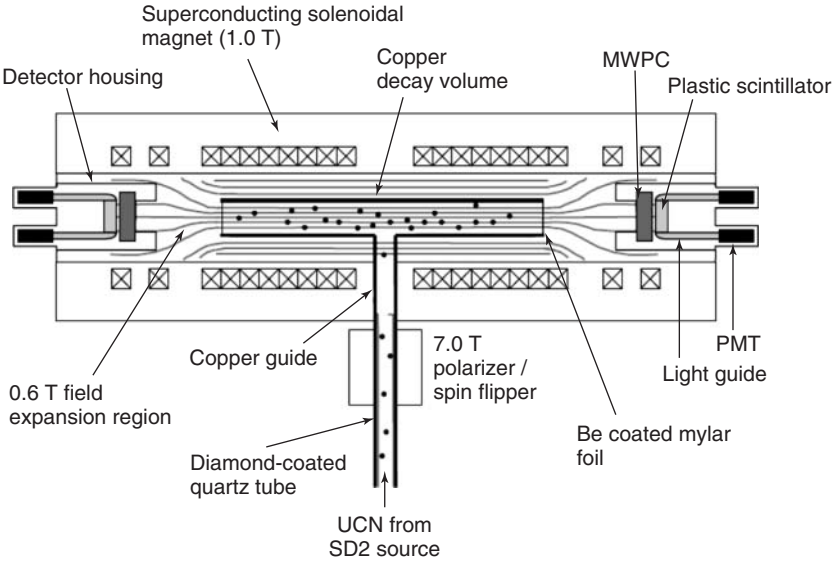


Figure 10.8 UCNA setup for a measurement of A. The magnetic field guides the electrons from the polarized UCN decay volume to the detectors composed of MWPC and scintillators [44].

terms V_{us} and V_{ub} stem from high-energy experiments. At present, unitarity Eq. (10.16) is closely fulfilled for neutron data. Using A from PDG [1] leads to a unitarity value in Eq. (10.16) of 1.0058(34).

The triple coefficient D is investigated in view of a possible violation of the time reversal symmetry T.

$$W(\theta) = 1 + D\frac{\vec{\sigma_n}(\vec{p_e} \times \vec{p_\nu})}{|\sigma_n| E_e E_\nu} \quad (10.22)$$

with D in the V–A theory:

$$D = -\frac{2\lambda}{1+3\lambda^2} \sin\varphi_{AV} \quad (10.23)$$

For the standard V–A theory, the phase between G_A and G_V is $\varphi_{AV} = 180^\circ$ and thus λ is real. A deviation from $D = 0$ would mean an complex phase φ_{AV} and thus violation of the T reversal symmetry, as long as the experimental D is higher than the so-called final-state interaction of the decay products, miming a D-value of 2×10^{-5}. The antineutrino in the angular correlation for D can be replaced for slow neutrons by the proton as $\mathbf{p_e} \times \mathbf{p_\nu} = -\mathbf{p_e} \times \mathbf{p_p}$. Thus, for a D measurement, the proton and electron are measured simultaneously. The polarization of the neutron beam should be, on an average, perpendicular to the electron–proton emission plane. A spin flip of the neutron corresponds then to the T operation, which reverses all velocities and spins. The electron–proton coincidence rate should not change with neutron spin reversal when T is conserved. Recent experiments used a multidetector geometry with 135° between proton and electron detectors (eMIT [48]) and a proton multidetector array combined with MWPC for tracing electrons (TRINE [49]), respectively. The combined result is given by PDG [1] as $\varphi_{AV} = 180.018(26)^\circ$. No T violation was observed within the experimental sensitivity.

10.5.3 Dipole Moments

10.5.3.1 Larmor Frequency of the Neutron and the Ramsey Method

The most precise experimental value for the dipole moments of the free neutron was obtained by measuring the Larmor frequency of the neutron in a magnetic field and an electric field using the Ramsey method of time-separated oscillatory field magnetic resonance [8]. The basic idea of this method is illustrated in Figure 10.9 for a neutron beam experiment. The Larmor precession frequency $\nu_L = 2\mu_n B_0/h$ in the magnetic field B_0 is compared with a radio frequency (RF) clock with $B_{RF} = B_1 \cos 2\pi \nu_{RF} t$.

Polarized neutrons are turned when passing the RF coils and thus Larmor precession is started. For a 90° turn of the neutron spin and $\nu_{RF} = \nu_L$, a time of flight in the RF coils of $\Delta t = h/(4\mu_n B_1)$ is

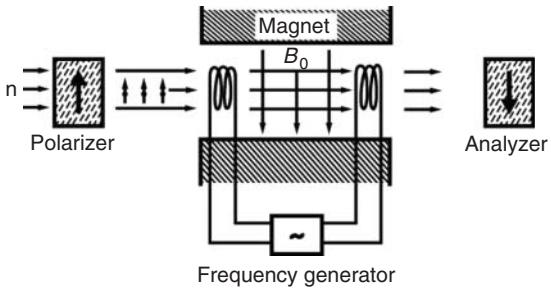


Figure 10.9 Principle of the Ramsey separated oscillatory field method.

required. After a free Larmor precession time, the neutrons pass the second RF coil, which turns the spin further in phase independently from the neutron velocity. The neutron spin direction is investigated by the polarization analyzer.

10.5.3.2 Magnetic Dipole Moment

A very precise experimental value for the ratio of the magnetic dipole moment of the neutron to that of the proton was obtained in 1979 by using a polarized neutron beam and the Ramsey method. The polarized neutron beam was passing through two coils to which the RF field was applied as shown in Figure 10.9. The Larmor precession in the magnetic field between the coils is measured by the corresponding RF in resonance. Water was running in a pipe along the same beam line and the NMR of the protons in the water molecules was measured. In the ratio of the two measurements, the static magnetic field B_0 cancels and the accuracy is given by the frequency measurement of ν_{RF}. The obtained value was very precise with $\mu_n/\mu_p = -0.684\,979\,35(17)$ [9].

This value is close to the estimate $-2/3$, which can be deduced from the quark structure of the neutron and the proton. The magnetic dipole moment is composed of the contribution of the quarks:

$$\mu_n = \text{const} \sum_i \langle \phi | q_i \sigma_z^i | \phi \rangle \qquad (10.24)$$

where

ϕ is the properly symmetrized spin wavefunction of the quark configuration,

q_i the charge of the quarks d and u, $-1/3$ e and $+2/3$ e, respectively, and

$\sigma_z{}^i$ the z-component of the quark spin.

The evaluation of Eq. (10.24) results in $\mu_n = \text{const}\ (-q_u + 4q_d)$ and $\mu_p = \text{const}\ (-q_d + 4q_u)$ and thus $\mu_n/\mu_p = -2/3$. The small difference from the experimental number is understood by a more sophisticated calculation. The absolute value of the magnetic dipole moment of the neutron is then $\mu_n = -1.913\ \mu_N$, with nuclear magneton $\mu_N = 5.051 \times 10^{-27}\ \text{J T}^{-1}$.

10.5.3.3 Possible Electric Dipole Moment

The search for a possible *electric dipole moment* (EDM) d_n of the neutron is of fundamental importance, because it would manifest time reversal T violation and, for CPT conservation, CP violation also, as illustrated in Figure 10.10. The present SM allows only a very small d_n of the order of 10^{-32} ecm stemming indirectly from the imaginary phase in the CKM matrix linked to the CP violation in higher quark generations (K_0 and B_0 mesons).

Very sensitive searches of d_n were carried out over many years essentially by groups at PNPI, Russia and at ILL using the Ramsey method with polarized neutron beams and then with polarized UCN. The change of the Larmor frequency in an overlying magnetic B_0 and electric E_0 field was investigated.

$$h\nu_L = 2(\mu_n B_0 \pm d_n E_0) \qquad (10.25)$$

The layout of the ILL experiment with UCN is shown in Figure 10.11. The free

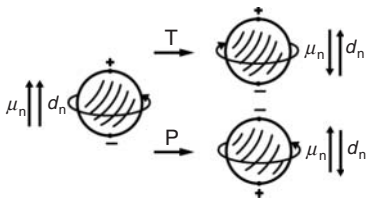


Figure 10.10 Illustration of T violation for a nonzero neutron electric dipole moment. Also P violation is shown.

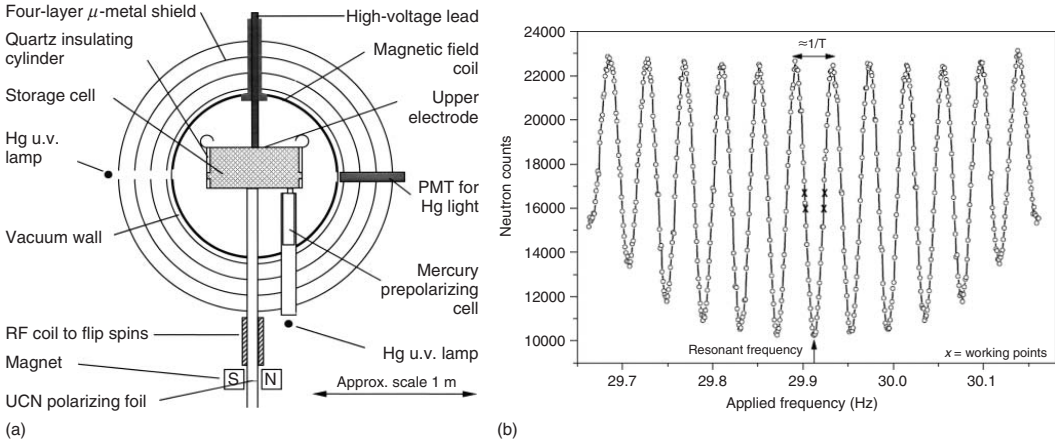


Figure 10.11 (a, b) EDM experiment at the ILL with UCN, resulting in the limit $d_n < 2.9 \times 10^{-26}$ ecm [50]. A measured resonance curve is shown. During the experiment, the working point was at the marked frequencies.(From [51].)

precession time in the UCN experiment is much longer (about 120 s) than in a neutron beam experiment (about 1 ms) and thus the resolution of the experiment is correspondingly higher (in analogy to the Heisenberg uncertainty principle). In Figure 10.11, the resonance curve as a function of the RF is shown. One of the main limitations was the low UCN density available nowadays with present UCN sources (about one polarized UCN per cubic centimeter in the storage cell).

The UCNs are filled into the storage bottle through a magnetized iron foil, which serves as a polarization filter. The spin flipper allows a measurement with opposite neutron spin direction for comparison. The storage bottle is placed in a magnetic field B_0 and a parallel or antiparallel electric field E_0. The coil for the RF field (frequency only about 30 Hz; coil not shown in Figure 10.11) is perpendicular to B_0. An RF pulse of $\Delta t = 2$ s duration is applied to the stored UCN to start Larmor precession. After about 120 s storage time, a second coherent pulse is applied and then the polarization is measured by emptying the bottle through the iron foil. The Larmor frequency with the parallel and antiparallel electric fields E_0 according to Eq. (10.25) is compared and thus d_n is investigated. As working point, the RF was set at values marked in Figure 10.11, where the change in Larmor frequency is most sensitively observed. The PNPI experiment used two flat UCN traps in the same magnetic field, but with opposite electric field directions to limit possible systematic errors [52].

At present, the measured value for d_n is compatible with 0, but a very stringent limit is obtained, which excludes already a number of possible theories beyond the SM or is close to those non-SM predictions (see for instance [53]). To get a feel of the already achieved precision: the d_n limit of $< 2.9 \times 10^{-26}$ ecm [50] in the applied electric field of $E_0 = 10$ kV cm^{-1} corresponds to an energy of 3×10^{-22} eV and a difference of four neutron precessions per year in the EDM apparatus (Eq. (10.25)).

In future, a significant progress is expected by a variety of new EDM experiments, using new and more intense UCN sources and setups with better control on magnetic field stability. An already running experiment in the test mode "cryoEDM" produces the UCN by the superthermal process [12] in the storage bottle with liquid helium, where neutron Larmor precession is also investigated in the liquid helium [54].

10.5.4 Neutron Oscillations

As a neutral particle, the neutron can oscillate into another neutral particle of similar mass without violating charge conservation. The ideas have phenomenological similarity to the recently observed neutrino oscillations.

Essentially, two classes of neutron oscillation scenarios were experimentally considered, an oscillation of a neutron into an antineutron $n - \overline{n}$ and, more exotically, an oscillation of a neutron into a hypothetical mirror world $n - n'$.

10.5.4.1 Neutron–Antineutron Oscillation

An oscillation of a neutron into an antineutron would be possible if the strong interaction eigenstate of the neutron is not identical to its mass eigenstate. The concept is similar to neutrino mixing and mixing in the kaon sector. As the neutron has a magnetic moment opposite in sign to that of the antineutron, magnetic fields hinder this oscillation. Residual

magnetic fields must be limited according to the Heisenberg uncertainty principle to (energy $\mu_n B$) × (neutron survival time) < $\hbar$.

The $n-\bar{n}$ oscillation would violate the conservation of baryon number B by two units. It requires theories that allow this $\Delta B = 2$ violation but suppresses the $\Delta B = 1$ violation as not yet observed for the proton decay.

Conceptually, the $n-\bar{n}$ oscillation experiment is straightforward as a search for $\bar{n}$ appearance. Free neutrons well shielded from magnetic fields may oscillate into $\bar{n}$ during their time of flight (survival time) to a foil in the beam, where $\bar{n}$ annihilates with the foil atoms. The experiment was running at the ILL with a cold neutron beam of total flux of 10^{11} n s^{-1} and a neutron flight path over 74 m. A sophisticated high-energy detector viewed the beam-passing foil. The $\bar{n}$ annihilation with the foil nuclei releases a total energy of 2 GeV. The isotropic momentum distribution in the laboratory system of such an event helps in the discrimination against cosmic rays. The experiment shows that the oscillation time must be longer than $\tau_{n-\bar{n}} > 8.6 \times 10^7$ s [55].

10.5.4.2 Neutron Oscillation into a Mirror World

In contrast to the $n-\bar{n}$ oscillation experiment, the neutron is supposed to disappear into a hidden mirror world [56] and the experimental method is a search for disappearance, which is conceptually less sensitive than an appearance experiment. As it is not clear whether the channel of disappearance closes in magnetic fields, the layout and interpretation are different. In a scenario where the possible oscillation is hindered by a magnetic field, a comparative UCN bottle experiment with and without a magnetic field yields the information. The result of such an experiment sets a lower limit of $\tau_{n-n'} > 414$ s [57]. If the magnetic field blockage is relaxed, the limit stems just from neutron losses along a neutron beam, with lower limit $\tau_{n-n'} > 12$ s [58].

10.6 Some Applications and Trends

The neutron will be applied with more and more sophistication in many fields of natural and life sciences. Several future developments in the fields of this chapter have already been mentioned in the corresponding sections, such as superthermal UCN sources, spallation neutron sources, and new generation of EDM and neutron decay experiments.

Beyond this nuclear physics chapter is the application in condensed matter, which is and will remain by far the broadest field of neutron research. As already mentioned, the good transmission of these neutral particles, their magnetic moment, the wavelength of thermal neutrons in the range of interatomic distances, their energies in the range of phonons in matter, and their mass similar to atomic masses make them an ideal probe for matter properties.

Neutrons are used widely for analytical purpose. Neutron imaging with thermal and fast neutrons allows nondestructive tomographic images of the inside of materials. Captured neutrons produce prompt γ-rays and often radioactive nuclei. The γ-lines yield on-line prompt gamma activation analysis (PGAA) and off-line neutron activation analysis (NAA) of isotopes. As an example, in the NAA of the scintillator for the neutrino detector BOREXINO in the Gran Sasso underground laboratory, a sensitivity of 10^{-17} g uranium and few 10^{-16} g for some other

elements per one gram scintillator was achieved [59].

Reactions with thermal neutrons are used for particle production in beam tubes close to the reactor core. A well-collimated intense beam of those particles is then available at the experimental site and can be used for high-resolution spectroscopy. Examples are neutron-induced γ-rays, electrons, and fission products (see [20] for corresponding ILL facilities) and recently at the FRM II the production of slow positrons by γ-ray absorption [60]. The reactor core itself is an intense source of electron antineutrinos steming from the β-decay of fission products. A detector for antineutrinos, placed about 180 km distant from nuclear reactors, clearly observed antineutrino oscillations [61].

Actually, modern multi-detector balls are placed also at (cold) neutron beams yielding detailed insight into nuclear structure for a wide range of nuclei using the neutron capture reactions (n_{th},f) including decay of fission products and (n_{th},γ).

In medicine and technology, radioactive tracers produced by neutron capture are frequently applied. Fast neutrons are used for cancer therapy, where the neutrons scatter at the protons in tumor tissue and thus these recoil protons damage the tumor cells.

Neutron physics is important for the understanding of cosmological and astrophysical processes. Some aspects regarding τ_n were briefly outlined in Section 10.5.2.1. A recent review highlights more generally the role of weak interaction and of neutron physics in this field of science [29].

Acknowledgment

The author is grateful to Dirk Dubbers, Frank Schreckenbach, Reinhard Stock, and Torsten Soldner for their valuable comments. The support by the DFG Cluster of Excellence ''Origin and Structure of the Universe'' is acknowledged.

Glossary

Antiparticle: Required by theory as counterpart for each particle. Particle and antiparticles are opposite in charge but otherwise identical.

Baryon: Particles consisting of three quarks, such as protons and neutrons.

Baryon Number *B* and Lepton Number *L*: Scalar quantities describing the empirical law of conservation of baryon and lepton numbers, respectively. The quarks have baryon number 1/3; all other particles have baryon number 0. Leptons have $L = 1$. Negative numbers are assigned to antiparticles. In present experimental observations, the baryon and lepton numbers are conserved in reactions.

Boson: Particle with integer spin. The fundamental forces are mediated by bosons.

Compound Nucleus: Excited nuclear state, which lives long enough such that the nuclear matter comes into equilibrium and forgets by which reaction channel it was excited.

Electric Dipole Moment EDM: Two separated charges of opposite polarity form an electric dipole with moment **d**, measured in units ecm. The corresponding energy in an electric field is $\mathbf{d} \times \mathbf{E}$.

Electromagnetic Interaction: Interaction due to electric and magnetic forces. Gamma quanta are the mediating bosons.

Fermi Pseudopotential: Potential for neutrons formed by the coherent neutron scattering at nuclei.

Fermion: Particles with half integer spin.

Fundamental Forces: The four known forces in nature: gravitation, electromagnetic force, weak force, and strong force.

ILL: Institut Laue Langevin, Grenoble, France running the high flux reactor (HFR), including neutron facilities for nuclear and particle physics research.

Isospin Component I_z: Quantum number characterizing the close equality in the strong interaction of protons and neutrons and in the quark model those of u and d quarks. The quantity is not an angular momentum, but the formalism is spinlike. Values of I_z: u quark 1/2, d quark $-1/2$, proton $+1/2$, neutron $-1/2$.

Larmor Precession: Rotation of the particle spin axis in an external field.

Lepton: Elementary fermions in the standard model, subject only to the electromagnetic and weak interactions, and to gravitation. There are three families: electron, muon, and tau with their corresponding neutrinos (plus antiparticles).

Lifetime τ: Time constant in the exponential decay of N particles: $N/N_0 = \exp(-t/\tau)$. Half of the initial population N_0 is decayed at the half-life $t_{1/2} = \tau \ln 2$.

Mirror Symmetries (Also Called Local Symmetries): The symmetries C charge conjugation (mirror particle – antiparticle), P parity (mirror in space $\mathbf{r} \rightarrow -\mathbf{r}$), and T time reversal (mirror time forward to time backward). The successive operations of the three mirror symmetries CPT lead, under very general assumptions, to the original system. CPT violation has never been observed.

Magnetic Moment of Particle: Magnetic dipole due to the ''spinning'' charge of a particle; energy in a magnetic field $= -\mu \times \mathbf{B}$.

Neutrinos: Neutral leptons. The observation of oscillations between neutrinos from different lepton families shows analogy to the mixing in the quark sector and proves their very small but different rest masses.

Nucleon: A collective term for proton and neutron as constituents of the atomic nucleus.

Phase Space: Six-dimensional quantity characterizing the occupation probability in space and momentum of a beam or in reaction channels.

Phonon: Quantized vibrational modes, in particular in solid states.

Pauli Principle: Exclusion principle, which states that identical fermions are not allowed in the same quantum state.

Quark: Constituents of hadrons. Hadrons are subdivided into baryons composed of three quarks and mesons composed of quark–antiquark pairs. Quarks do not occur as free particles and the charge of the quarks is not integer (values $\pm 1/3$ or $\pm 2/3$ e). Neutrons and protons are composed of the three quarks udd and uud, respectively.

Spontaneous Fission: Nuclear fission without external energy input, frequent decay branch for actinides.

Standard Model (SM) of Particle Physics: Widely accepted model of elementary particles: three families of quarks (together six quarks) and three families of leptons (together six leptons) together with their antiparticles. The fundamental forces, electromagnetic, weak, and strong, are mediated by the bosons gamma quanta, ($W^{\pm}$, Z), and gluons, respectively.

Strong Interaction: Nuclear force, responsible for the binding of hadrons and the binding of nucleons in a nucleus. Mediating bosons are the gluons and between nucleons the mesons.

Symmetry Violation (for P, C, and T): A mirror symmetry is violated, when the original and mirrored systems are not identical, that is, they cannot be made identical by rotation or linear shift in space. In other words, symmetry violation is observed when an experiment can definitely distinguish whether it occurs in the original or mirrored system.

UCNs: Ultracold neutrons, for practical reasons defined as neutrons with velocities of somewhat below $7\,\mathrm{m\ s^{-1}}$ as these neutrons can be confined by the Fermi pseudopotential in material wall traps.

$W^{\pm}$ Boson: Charged boson mediating the weak interaction. The third boson of weak interaction is the neutral Z boson.

Weak Interaction: Force between leptons and hadrons, mediated by the $W^{\pm}$ and Z bosons. These bosons are very massive and hence the weak force is short-ranged. In the quark sector, the weak and strong interaction eigenstates are different, which allows decays between the quark generations.

References

1. Particle Data Group (2012) *Phys. Rev. D*, **86**, 73.
2. Chadwick, J. (1932) *Nature*, **129**, 312.
3. Kessler, E.G. *et al.* (1999) *Phys. Lett. A*, **255**, 221.
4. Hasegawa, Y. and Rauch, H. (2011) *New J. Phys.*, **134**, 115010.
5. Nesvizhevsky, V. *et al.* (2002) *Nature*, **415**, 297.
6. Nesvizhevsky, V.V. (2012) *Mod. Phys. Lett. A*, **27**, 1230006.
7. Abele, H. and Leeb, H. (2012) *New J. Phys.*, **14**, 055010.
8. Ramsey, N.F. (1950) *Phys. Rev.*, **78**, 695.
9. Greene, G.L. *et al.* (1979) *Phys. Rev. D*, **20**, 2139.
10. Baumann, J., Gaehler, R., Kalus, J., and Mampe, W. (1988) *Phys. Rev. D*, **37**, 3107.
11. Steyerl, A. *et al.* (1986) *Phys. Lett. A*, **116**, 347.
12. Golub, R. and Pendlebury, J.M. (1975) *Phys. Lett. A*, **53**, 133.
13. Zimmer, O. *et al.* (1999) *Phys. Lett. B*, **455**, 62.
14. Klauser, C. *et al.* (2012) *J. Phys. Conf. Ser.*, **340**, 012011.
15. McGrath, C.A. *et al.* (1999) *Nucl. Intrum. Methods A*, **421**, 458.
16. Mughabghab, S.F. (2006) *Atlas of Neutron Resonances*, Elsevier, Amsterdam.
17. Koester, L., Rauch, H., and Seymann, E. (1991) *At. Data Nucl. Data Tables*, **49**, 65.
18. Sears, V.F. (1992) *Neutron News*, **3** (3), 26.
19. Casten, R.F. *et al.* (1980) *Phys. Rev. Lett.*, **45**, 1077.
20. Börner, H.G. and Gönnenwein, F. (2012) *The Neutron*, World Scientific, Singapore.
21. Cavaignac, J.F. *et al.* (1977) *Phys. Lett. B*, **67**, 148.
22. Gericke, M.T. *et al.*, and NPDGamma Collaboration (2011) *Phys. Rev. C*, **83**, 015505.
23. Forte, M. *et al.* (1980) *Phys. Rev. Lett.*, (45), 2088.
24. Heil, W. *et al.* (1999) *Physica B*, **267–268**, 289.
25. Crawford, B.E. *et al.* (1998) *Phys. Rev. C*, **58**, 1225.
26. Goennenwein, F. *et al.* (2007) *Phys. Lett. B*, **652**, 13.
27. Abele, H. (2008) *Prog. Part. Nucl. Phys.*, **60**, 1.
28. Severijns, N., Beck, M., and Naviliat-Cuncic, O. (2006) *Rev. Mod. Phys.*, **78**, 991.
29. Dubbers, D. and Schmidt, M.G. (2011) *Rev. Mod. Phys.*, **83**, 1111.
30. Hofstadter, R. and Herman, R. (1961) *Phys. Rev. Lett.*, **6**, 293.
31. Dubbers, D. *et al.* (2008) *Nucl. Instrum. Methods A*, **596**, 238.
32. Schreckenbach, K. and Mampe, W. (1992) *J. Phys. G*, **18**, 1.
33. Paul, S. (2009) *Nucl. Instrum. Methods A*, **611**, 157.
34. Wietfeldt, F.E. and Greene, G.L. (2011) *Rev. Mod. Phys.*, **83**, 1173.
35. Vogel, P. and Beacom, J.F. (1999) *Phys. Rev. D*, **60**, 053003.
36. Marciano, W.J. and Sirlin, A. (2006) *Phys. Rev. Lett.*, **96**, 0320002.
37. Nico, J.S. *et al.* (2005) *Phys. Rev. C*, **71**, 055502.
38. Pichlmaier, A. *et al.* (2010) *Phys. Lett. B*, **693**, 221.
39. Serebrov, A.P. *et al.* (2008) *Phys. Rev. C*, **78**, 035505.

40. Arzumanov, S. *et al.* (2000) *Phys. Lett. B*, **483**, 15; revised (2012) *JETP Lett.* **95**, 224.
41. (a) Mampe, W. *et al.* (1989) *Phys. Rev. Lett.*, **63**, 593; (b) Steyerl, A. *et al.* (2012) *Phys. Rev. C*, **85**, 065503.
42. Bates, J.C. (1982) *Phys. Lett. A*, **88**, 427.
43. Paul, W. *et al.* (1989) *Z. Phys. C*, **45**, 25.
44. Plaster, B. *et al.*, and UCNA Collaboration (2012) *Phys. Rev. C*, **86**, 055501.
45. Abele, H. *et al.* (2002) *Phys. Rev. Lett.*, **88**, 211801.
46. Mund, D. *et al.* (2013) *Phys. Rev. Lett.*, **110**, 172502
47. Reich, J. *et al.* (2000) *Nucl. Instrum. Methods A*, **440**, 535.
48. Mumm, H.P. *et al.* (2011) *Phys. Rev. Lett.*, **107**, 102301.
49. Soldner, T. *et al.* (2004) *Phys. Lett. B*, **581**, 49.
50. Baker, C.A. *et al.* (2006) *Phys. Rev. Lett.*, **97**, 131801.
51. Harris, P.G. *et al.* (1999) *Phys. Rev. Lett.*, **82**, 904.
52. Altarev, I.S. *et al.* (1996) *Phys. At. Nucl.*, **59**, 1152.
53. Pospelov, M. and Ritz, A. (2005) *Ann. Phys.*, **318**, 119.
54. Baker, C.A. *et al.*, and CryoEDM Collaboration (2010) *J. Phys. Conf. Ser.*, **251**, 012055.
55. Baldo-Ceolin, M. *et al.* (1994) *Z. Phys. C*, **63**, 409.
56. Berezhiani, Z. and Bento, L. (2006) *Phys. Rev. Lett.*, **96**, 081801.
57. Serebrov, A.P. *et al.* (2008) *Phys. Lett. B*, **663**, 181.
58. Altarev, I. *et al.* (2009) *Phys. Rev. D*, **80**, 032003.
59. Hentig, R.V. *et al.* (1999) *Nucl. Phys. B: Proc. Suppl.*, **78**, 115.
60. Hugenschmidt, C. *et al.* (2005) *Nucl. Instrum. Methods A*, **554**, 384.
61. Abe, S., *et al.*, and KamLAND Collaboration (2008) *Phys. Rev. Lett.*, **100**, 221803.

Further Readings

Schofield, P. (ed.) (1982) *The Neutron and its Applications*, Conference Series, Vol. 64, Institute of Physics, Bristol.
Conference contributions celebrating 50 years since the discovery of the neutron

Byrne, J. (1995) *Neutrons, Nuclei and Matter*, 3rd edn, Institute of Physics, Bristol.
Covers the whole field of theoretical and experimental neutron science.

Bacon, G.E. (1975) *Neutron Diffraction*, 3rd edn, Clarendon Press, Oxford.
Classical book on neutron scattering.

Arif, M., Dewey, M.S., Gentile, T., Huffman, P., and Nico, J. (eds) (2005) Precision measurements with slow neutrons, NIST. *J. Res. Inst. Stand. Technol.*, **110** (3–4), 327–311.

Börner, H.G. and Gönnenwein, F. (2012) *The Neutron*, World Scientific, Singapore,
Experiments in nuclear and particle physics carried out at the ILL, France

Soldner, T., Nesvizhevsky, V., Plonka-Spehr, C., Protasov, K., Schreckenbach, K., and Zimmer, O. (eds.) (2009) Particle Physics with slow neutrons, *Nucl. Instrum. Methods A*, **611** (2–3) 111Proceedings of an international workshop at the ILL

Rauch, H. and Werner, S. (2000) *Neutron Interferometry*, Oxford University Press, Oxford.

Holstein, B.R. (ed.) (2009) *J. Phys. G*, **36**, 100301,104001–104006.
Comprehensive review articles on neutron lifetime, EDM, PNC, chiral structure of the neutron, gravitational quantum states, neutron–antineutron oscillation.

11
Neutrino Astrophysics

Wick C. Haxton

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

11.1 Introduction

The neutrino [1] is an elementary particle that scatters only through the weak interaction, and consequently rarely interacts in matter. Neutrinos are neutral, carry spin-1/2, and are members of the family of elementary particles called *leptons*. Thus, they differ from the quarks of the standard model (spin-1/2 particles that participate in strong and electromagnetic interactions) and from the other leptons, which are charged and thus interact electromagnetically. Neutrinos and their antiparticles come in three types – or flavors – labeled according to the charged partners, the electron, muon, or tauon, that accompany neutrino production in charge-changing weak interactions. The most familiar of such reactions is β-decay

$$(N, Z) \rightarrow (N-1, Z+1) + \mathrm{e}^- + \overline{\nu}_\mathrm{e}$$

in which a nucleus containing N neutrons and Z protons decays to a lighter nucleus by converting a neutron to a proton, with the emission of an electron and an electron antineutrino. Indeed, it was the apparent absence of energy conservation in nuclear β-decay that first lead Wolfgang Pauli, more than 80 years ago, to suggest that some undetected neutral particles (the ν_e) must be escaping from nuclear β-decay experiments.

Neutrinos play a very special role in astrophysics [2]. First, they are the direct by-products of the nuclear reaction chains by which stars generate energy: each solar conversion of four protons into helium produces two neutrinos, for a total of $\sim 2 \times 10^{38}$ neutrinos each second. The resulting flux is observable on Earth. These neutrinos carry information about conditions deep in the solar core, as they typically leave the Sun without further interacting. They also provide experimentalists with opportunities for testing the properties of neutrinos over long distances. Second, they are produced in nature's most violent explosions, including the Big Bang, core-collapse supernovae, and the accretion disks encircling supermassive black holes. Recent discoveries of neutrino mass show that primordial neutrinos comprise a small portion of the "dark matter" that influences how large-scale structure – the pattern of voids and galaxies mapped by astronomers – formed over cosmological times. Third, they dominate the cooling of many astrophysical objects, including young neutron stars and the degenerate helium cores of red giants. Neutrinos can be radiated from deep within such bodies, in contrast to photons, which are trapped within stars, diffusing outward only slowly.

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

Finally, neutrinos are produced in our atmosphere and elsewhere as secondary by-products of cosmic-ray collisions. Detection of these neutrinos can help constrain properties of the primary cosmic ray spectrum. Neutrinos produced by reactions of ultrahigh-energy cosmic rays (UHECRs) can provide information on otherwise inaccessible cosmic accelerators.

Neutrinos also mediate important astrophysical processes. While the site of the r-process – the rapid-neutron-capture process thought to be responsible for the nucleosynthesis of about half of the neutron-rich nuclear species heavier than iron – remains uncertain, the needed neutrons may be generated by the extraordinary neutrino fluxes found in core-collapse supernovae. Possible r-process sites include the neutron-rich ''neutrino winds'' that blow off the proto-neutron star surface as well as the ^{4}He zones of metal-poor supernovae, where neutrons are produced by neutrino-induced spallation reactions. Supernova neutrinos can also directly synthesize certain rare nuclei, such as ^{11}B and ^{19}F.

Nuclear and particle physicists are exploiting astrophysical neutrino fluxes to do important tests of the standard model. These tests include neutrino oscillations (the process by which a massive neutrino can be produced in one flavor state but detected later as a neutrino with a different flavor), neutrino decay, the cosmological effects of neutrino mass, and searches for nonzero neutrino electromagnetic moments.

11.2 Solar Neutrinos

The first successful effort to detect neutrinos from the Sun began four decades ago. Ray Davis, Jr and his collaborators constructed a 650 ton detector in the Homestake Gold Mine, one mile beneath Lead, South Dakota [3]. This radiochemical detector, based on the chlorine-bearing cleaning fluid C_2Cl_4, was designed to capture about one of the approximately 10^{18} high-energy neutrinos that penetrated it each day – the rest passed through the detector, without interacting. The neutrino-capture reaction was inverse-electron-capture:

$$^{37}\mathrm{Cl} + \nu_{\mathrm{e}}^{\mathrm{solar}} \rightarrow {}^{37}\mathrm{Ar} + \mathrm{e}^{-}$$

The product of this reaction, ^{37}Ar, is a noble-gas isotope with a half-life of about 1 month. It can be efficiently removed from a large volume of organic fluid by a helium gas purge, then counted in miniature gas proportional counters as ^{37}Ar decays back to ^{37}Cl. Davis typically exposed his detector for about 2 months, building up to nearly the saturation level of a few dozen argon atoms, then purged the detector to determine the number of solar neutrinos captured during this period (Figure 11.1).

Within a few years, it became apparent that the number of neutrinos detected was only about one-third that predicted by the standard solar model (SSM) [2, 4], that is, the model of the Sun based on the standard theory of main-sequence stellar evolution. Some initially attributed this ''solar neutrino problem'' to uncertainties in the SSM: as the flux of neutrinos most important to the Davis detector vary as $\sim T_{\mathrm{c}}^{22}$, where T_{c} is the solar core temperature, a 5% theory uncertainty in T_{c} could explain the discrepancy. In fact, the correct explanation for the discrepancy proved much more profound. Davis was awarded the 2002 Nobel Prize in Physics for the chlorine experiment.

Figure 11.1 The Homestake Mine's chlorine detector, which Ray Davis Jr and colleagues operated for over three decades.

The solar neutrino problem stimulated a series of follow-up experiments to measure the different components of the solar neutrino flux and to determine the source of the Cl experiment discrepancy. The SAGE and GALLEX/GNO experiments, in which radiochemical detectors similar to Cl but using ^{71}Ga as a target, were designed to measure the flux of neutrinos from the dominant low-energy branch of solar neutrinos, the pp neutrinos. The first detector to measure neutrinos event by event, recording neutrino interactions in real time, was the converted proton decay detector, Kamiokande. The detector contained 3 ktons of very pure water, with solar neutrinos scattering off the electrons within the water. Phototubes surrounding the tank recorded the ring of Cerenkov radiation produced by the recoiling relativistic electrons. Kamiokande measured the high-energy neutrinos most important to the Davis detector, and thus confirmed the deficit that Davis had originally observed. New and very massive water (Super-Kamiokande) and heavy-water [Sudbury Neutrino Observatory (SNO)] detectors were constructed. Finally, Borexino, a detector using liquid scintillator, was constructed to measure low-energy solar neutrino branches in real time. These experiments – most particularly SNO, because of its multiple detection channels sensitive to different combinations of neutrino types – showed that solar neutrinos were not missing, but rather hidden by a change of flavor occurring during their transit from the Sun to the Earth, as will be described later in this chapter.

Of these experiments, SAGE, Super-Kamiokande, and Borexino remain in operation.

11.2.1 The Standard Solar Model

The Sun belongs to a class of ''main sequence'' stars that derive their energy from burning protons to He in their cores. The SSM employs the standard theory of main-sequence stellar evolution, calibrated by the many detailed measurements only possible for the Sun, to follow the Sun from the onset of thermonuclear reactions 4.6 Gyr ago to today, thereby determining the present-day temperature and composition profiles of the solar core. These profiles govern solar neutrino production and other properties of the modern Sun. The SSM is based on four basic assumptions:

- The Sun evolves in hydrostatic equilibrium, maintaining a local balance between the gravitational force and the pressure gradient. To describe this condition in detail, one must specify the electron – gas equation of state as a function of temperature, density, and composition. This requires attention to issues such as the incomplete ionization of metals, the contribution of radiation to pressure, and the influence of screening.
- Energy is transported by radiation and convection. While the solar envelope is convective, radiative transport dominates in the core region where thermonuclear reactions take place. In addition to elementary processes such as the scattering of photons off electrons and fully ionized H and He, more complex processes such as bound-free scattering off metals are important contributors to the opacity in the Sun's interior regions.
- The Sun produces its energy by fusing protons into ^{4}He,

$$2e^- + 4p \rightarrow {}^4He + 2\nu_e + 26.73MeV$$

via the pp chain (99%) and CN I cycle reactions of Figure 11.2. The Sun is a large but slow reactor: the core temperature, $T_c \sim 1.5 \times 10^7$K, results in typical center-of-mass energies for reacting particles of $\sim$ 10keV, much less than the Coulomb barriers inhibiting charged-particle nuclear reactions. Solar reaction rates are so slow that laboratory measurements of these rates are not, in most cases, feasible at solar energies, but instead must be made at higher energies and then extrapolated to the solar Gamow peak.
- The model is constrained to produce today's solar radius, mass, and luminosity. An important assumption of the SSM is that the proto-Sun passed through a highly convective phase, rendering the Sun uniform in composition until main-sequence burning began. The initial composition by mass is conventionally divided into hydrogen (X_{ini}), helium (Y_{ini}), and everything else (the metals, denoted Z_{ini}), with $X_{ini} + Y_{ini} + Z_{ini} = 1$. The relative abundances of the metals are determined from a combination of meteoritic and solar photospheric data. The absolute abundance Z_{ini} can be taken from the modern Sun's surface abundance Z_S, after corrections for the effects of diffusion over 4.6 Gyr of solar evolution. Finally, Y_{ini}/X_{ini} is adjusted along with α_{MLT}, a parameter describing solar mixing, until the model reproduces the modern Sun's luminosity and radius. The resulting ^{4}He/H mass fraction ratio is typically 0.27 ± 0.01, which can be compared to the Big Bang value of 0.23 ± 0.01, showing that the Sun was formed from previously processed material.

Three cycles with quite different temperature dependences, reflecting the relative ease or difficulty of Coulomb barrier penetration, comprise the pp chain of Figure 11.2. The competition between the cycles is very sensitive to the solar core temperature T_c. The initial interest in solar neutrinos came from the observation that each of the three cycles is associated with a characteristic neutrino. Thus, by measuring solar neutrinos, specifically the pp, ^{7}Be, and ^{8}B neutrinos, one can determine the relative importance of the ppI, ppII, and ppIII cycles, and consequently determine T_c to an accuracy of $\lesssim$1%.

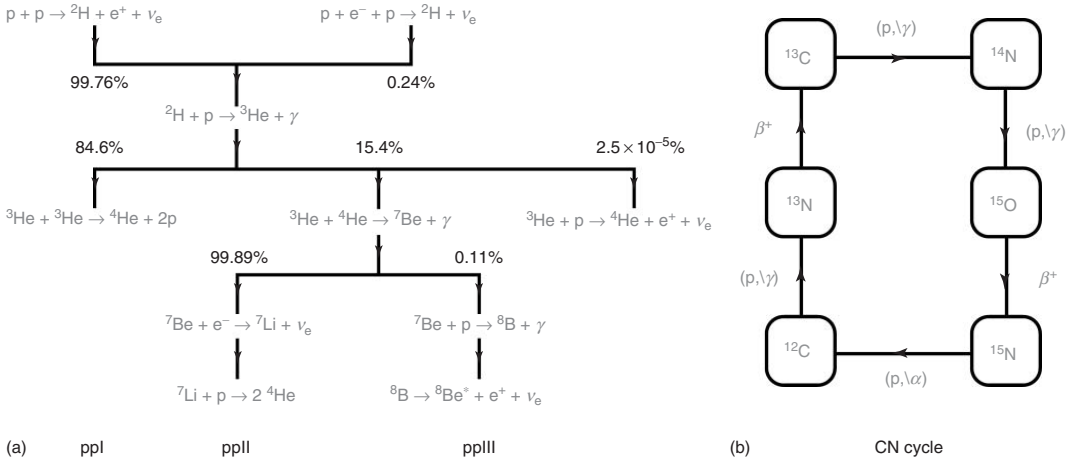


Figure 11.2 (a) The three principle cycles comprising the pp chain (denoted ppI, ppII, and ppIII). Associated neutrinos "tag" the three branches. The SSM branching ratios come from the GS98-SFII SSM [4]. Also shown is the minor branch $^3\mathrm{He} + \mathrm{p} \rightarrow {}^4\mathrm{He} + \mathrm{e}^+ + \nu_\mathrm{e}$, which generates the Sun's most energetic neutrinos. (b) The CN I cycle, which produces solar neutrinos from the β-decays of ^{13}N and ^{15}O.

The neutrino-producing reactions of the pp chain and CN I cycle are summarized in Table 11.1. The β-decay sources produce neutrinos with continuous spectra, while the electron capture reactions produce lines with widths ∼2 keV characteristic of the temperature of the solar core. The table shows two solar models, denoted GS98-SFII and AGSS09-SFII, which differ in their assumptions about solar surface metallicity due to the use of 1D or 3D models, respectively, to interpret photospheric absorption lines. The predictions of the higher metallicity (∼+30%) GS98-SFII SSM are generally in excellent agreement with solar helioseismic properties, including interior sound speeds and the location of the base of the convective zone. This is not the case for the AGSS09-SFII SSM, which nevertheless uses a more sophisticated treatment of the photosphere. The unresolved conflict between SSMs that agree with our best description of the solar interior and those that employ our best model of the solar surface is known as the *solar abundance problem*.

The last line of the table shows that the neutrino flux predictions of the two SSMs are almost identical in the quality of their agreement with fluxes derived from experiment. A best fit would be obtained for a metallicity $Z_{\rm ini}$ intermediate between GS98 and AGSS09.

11.2.2 SNO, Super-Kamiokande, and Borexino

Solar neutrino detection requires the combination of a large detector volume (to provide the necessary rate of events), very low backgrounds (so that neutrino events can be distinguished from backgrounds due to cosmic rays and natural radioactivity), and a distinctive signal. The first requirement favors detectors constructed from inexpensive materials and/or materials having large cross sections for neutrino capture. The second generally requires a deep-underground location for the detector, with sufficient rock overburden to attenuate the flux of penetrating muons produced by cosmic-ray interactions in the atmosphere. It also requires very careful attention to detector cleanliness, including tight limits on dust or other contaminants that might introduce radioactivity, use of low-background construction materials, control of radon, and often the use of fiducial volume cuts so that the outer portions of a detector become a shield against activities produced in the surrounding rock walls.

There are several possible detection modes for solar neutrinos, interesting because of their different sensitivities to flavor. The early radiochemical experiments using ^{37}Cl and ^{71}Ga targets were based on the charged-current (CC) weak reaction

$$\nu_{\rm e} + (N, Z) \rightarrow {\rm e}^- + (N-1, Z+1)$$

where the signal for neutrino absorption is the growth over time of very small concentrations of the daughter nucleus $(N-1, Z+1)$ in the detector. As the spectrum of solar neutrinos extends only to about 15 MeV, well below the threshold for producing muons, this reaction is sensitive only to electron neutrinos.

A second possible nuclear detection channel is neutral-current (NC) scattering

$$\nu_x + (N, Z) \rightarrow \nu_x' + (N, Z)^*$$

a process independent of the neutrino flavor. If this scattering leaves the nucleus in an excited state, the observable would be the de-excitation of the nucleus, such as a decay γ-ray or the breakup of the nucleus. (An example will be given below,

Table 11.1 SSM neutrino fluxes from the GS98-SFII (high-Z) and AGSS09-SFII (low-Z) SSMs, which differ in their assumptions about photospheric metallicity.

ν flux	E_ν^{max}(MeV)	GS98-SFII	AGSS09-SFII	Solar	Units
$p+p\rightarrow {}^2H+e^+ +\nu$	0.42	$5.98(1\pm0.006)$	$6.03(1\pm0.006)$	$6.05(1^{+0.003}_{-0.011})$	10^{10} cm^{-2} s^{-1}
$p+e^- +p\rightarrow {}^2H+\nu$	1.44	$1.44(1\pm0.012)$	$1.47(1\pm0.012)$	$1.46(1^{+0.010}_{-0.014})$	10^{8} cm^{-2} s^{-1}
${}^7Be+e^-\rightarrow {}^7Li+\nu$	0.86 (90%)	$5.00(1\pm0.07)$	$4.56(1\pm0.07)$	$4.82(1^{+0.05}_{-0.04})$	10^{9} cm^{-2} s^{-1}
	0.38 (10%)	–	–	–	–
${}^8B\rightarrow {}^8Be+e^+ +\nu$	∼15	$5.58(1\pm0.14)$	$4.59(1\pm0.14)$	$5.00(1\pm0.03)$	10^{6} cm^{-2} s^{-1}
${}^3He+p\rightarrow {}^4He+e^+ +\nu$	18.77	$8.04(1\pm0.30)$	$8.31(1\pm0.30)$	–	10^{3} cm^{-2} s^{-1}
${}^{13}N\rightarrow {}^{13}C+e^+ +\nu$	1.20	$2.96(1\pm0.14)$	$2.17(1\pm0.14)$	≤ 6.7	10^{8} cm^{-2} s^{-1}
${}^{15}O\rightarrow {}^{15}N+e^+ +\nu$	1.73	$2.23(1\pm0.15)$	$1.56(1\pm0.15)$	≤ 3.2	10^{8} cm^{-2} s^{-1}
χ^2/P^{agr}	–	3.5/90%	3.4/90%	–	–

In cases where associated uncertainties are asymmetric, an average has been used. The solar values come from a luminosity-constrained analysis of all available data by the Borexino Collaboration. For references see [5].

in the discussion on SNO.) Alternatively, neutrino elastic scattering (ES) (without nuclear excitation) is a coherent process at low energies, with a cross section proportional to the square of the weak charge, which is approximately the neutron number N of the nucleus. The signal then is the small recoil energy of the nucleus after scattering.

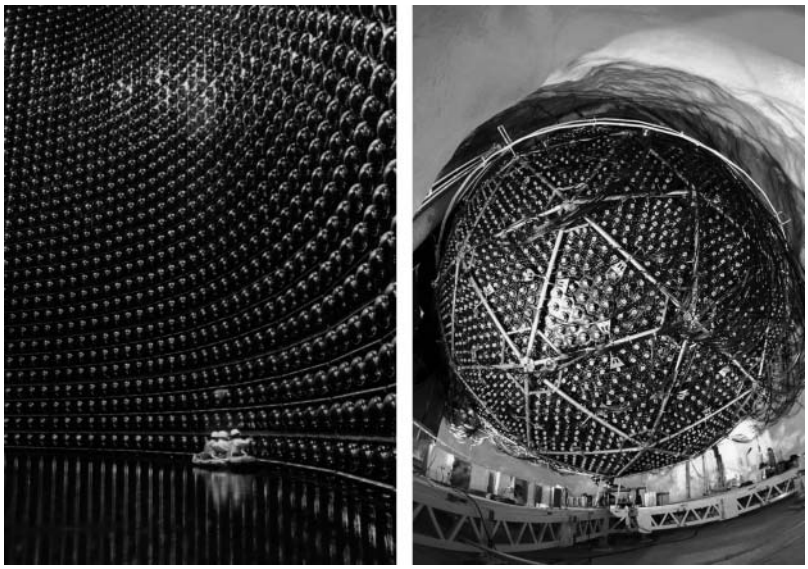
A third possibility is the scattering of neutrinos off electrons,

$$\nu_x + e^- \rightarrow \nu_x' + e^-$$

with detection of the recoiling scattered electron. Both electron- and heavy-flavor (ν_μ, ν_τ) solar neutrinos can scatter off electrons, the former by charge and NCs, and the latter by NCs only. Consequently, the cross section for scattering heavy-flavor neutrinos is only about 0.15 that for electron neutrinos. This process provides a third way of probing neutrino flavor, owing to this differential sensitivity. An important aspect of electron-neutrino scattering is its directionality: for solar neutrino energies much above the electron mass of 0.511 MeV, the electron scatters into a narrow cone along the incident neutrino's direction. This directionality provides a powerful tool for extracting solar neutrino events from the background: neutrino events correlate with the direction of the Sun, while background events should be isotropic. Thus neutrino events can be identified as the excess seen at forward electron angles.

These various detection channels were exploited in two large-volume water Cerenkov detectors that recorded events in real time and provided flavor sensitivity, as well as in the liquid scintillator experiment Borexino.

Super-Kamiokande *et al.* [6] (Figure 11.3) is a detector consisting of 50 ktons of ultrapure water within a cylindrical stainless steel tank, 39m in diameter and 42m tall. Two meters inside the walls a scaffold supports a dense array of 50 cm diameter hemispherical PMTs, which face inward and view the inner 32 ktons of water. Additional 20 cm tubes face outward, viewing the outer portion of the



(a) (b)

Figure 11.3 (a) The Super-Kamiokande detector during filling, with scientists cleaning photomultiplier tube (PMT) surfaces as the water rises. (b) Fish-eye photo of the SNO detector and cavity, showing the PMTs and support structure before cavity and detector filling.

detector that serves as a shield and as a veto. A solar neutrino can interact in the inner detector, scattering off an electron. The recoiling, relativistic electron then produces a cone of Cherenkov radiation, a pattern that can be reconstructed from the triggering of the phototubes that surround the inner detector. The detector is housed deep within Japan's Kamioka Mine, approximately 1 km underground.

The Super-Kamiokande detector began operations in 1996, progressing from phase I to the current phase IV. During Super-Kamiokande I, the detector was instrumented with an array of 11,146 50 cm PMTs, corresponding to about 40% coverage. In November 2001, during a shutdown period for repairs and upgrades, one of the 50 cm PMTs imploded, creating a powerful shock wave that propagated through the tank, destroying 60% of the phototubes. The detector was subsequently rebuilt with about half the original number of phototubes, evenly spaced over the scaffold, so that the coverage was reduced to 20%. The detector operated in this SK-II phase from late 2002 until 2005. Following a second reconstruction in which the phototube coverage was restored to 40% and other improvements made, SK-III data were obtained from October 2006 to August 2008. Preliminary results for 1069 days of running in the current SK-IV phase (in which the threshold for detecting electrons has been lowered to a total energy of 4 MeV) were reported in summer, 2012. The SK-III observed event rate of scattered electrons between 5.0 and 20.0 MeV is equivalent to an unoscillated ^{8}B flux of $[2.39 \pm 0.04\ (\text{stat}) \pm 0.05\ (\text{sys})] \times 10^6\ \text{cm}^{-2}\ \text{s}^{-1}$ [7], well below the SSM flux given in Table 11.1.

The SNO [8] (Figure 11.3) was constructed at an extraordinary depth, 2 km underground in the INCO Creighton nickel mine in Ontario, Canada. The detector took data from May 1999 to November 2006, operating in three different modes over its 7.5 year lifetime. SNO employed a 1 kton target of heavy water, contained within a spherical acrylic vessel 6 m in radius. This sphere was surrounded by an additional 5 m (7 kton) of very pure ordinary water, filling the rock cavity that housed the entire detector. An array of 9600 20 cm PMTs, mounted on a geodesic sphere surrounding the inner vessel, provided 56% coverage. As in the case of SK-III, SNO operated with a threshold of 5 MeV through much of its lifetime, detecting the portion of the ^{8}B solar neutrino spectrum from 5 to 15 MeV. (Recently, a low-energy reanalysis has been completed that employed a electron kinetic energy threshold of 3.5 MeV).

The choice of a heavy-water target allowed SNO experimentalists to exploit all three of the reaction channels described above, with their varying flavor sensitivities

$$\nu_x + e^- \rightarrow \nu'_x + e^- \quad \text{ES : elastic scattering}$$
$$\nu_e + d \rightarrow p + p + e^- \quad \text{CC : charged current}$$
$$\nu_x + d \rightarrow \nu'_x + n + p \quad \text{NC : neutral current}$$

The ES reaction is the same as that employed by Super-Kamiokande, with its differing sensitivities to electron- and heavy-flavor neutrinos. The CC reaction on deuterium is sensitive only to electron-flavor neutrinos, producing electrons that carry off most of the incident neutrino's energy (apart from the 1.44 MeV needed to break a deuterium nucleus into p + p). Thus, from the energy distribution of the electrons, one can reconstruct the incident ν_e spectrum (and possible distortions discussed below) more accurately than in the case of ES.

The NC reaction, which is observed through the produced neutron, provides no spectral information, but does measure the total solar neutrino flux, independent of flavor. The SNO experiment has used three techniques for measuring the neutrons. In the initial pure-D_2O phase, the neutrons captured on deuterium, producing 6.25 MeV γs. In the second phase, 2.7 tons of salt were added to the heavy water so that Cl would be present to enhance the capture, producing 8.6 MeV γs. In both of these approaches, the NC and CC events can be separated reasonably well because of the modest backward peaking ($\sim 1 - \cos\theta/3$) in the angular distribution of the latter. This allowed the experimenters to determine the total and electron-neutrino fraction of the solar neutrino flux. Finally, in the third phase, direct neutron detection was provided in pure D_2O by an array of ^{3}He-filled proportional counters, cleanly separating this signal from CC scattering.

SNO was constructed at very great depth and under clean-room conditions because of the need to suppress backgrounds. In particular, a minute amount of dust in the detector could have introduced environmental radioactivities that would have obscured the NC signal, a single neutron. The great advantage of the SNO detector was its three distinct detection channels, sensitive to different combinations of electron and heavy-flavor neutrinos. Furthermore, because the ES- and CC-scattered electrons are measured in the same detector, several important systematic effects cancel in the ratio of events. The ES reaction provided an important cross check on the consistency of the results from the CC and NC channels.

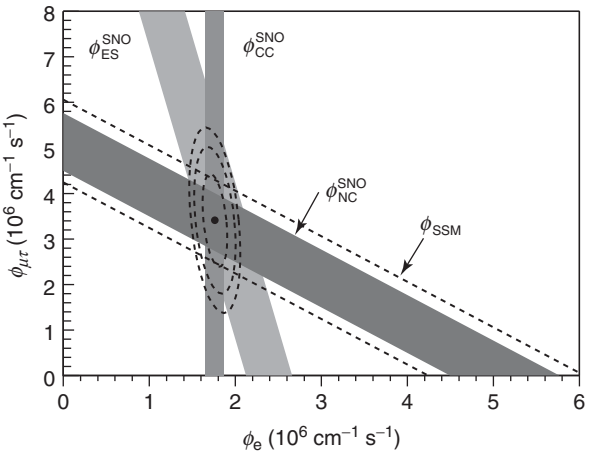


Figure 11.4 Results from the D_2O phase of the SNO experiment [8]. The allowed bands for CC, NC, and ES reactions of solar neutrinos intersect to show a flux that is one-third ν_es and two-thirds heavy-flavor neutrinos. There is agreement between the NC total-flux measurement and the predictions of the SSM (band indicated by the dashed lines). (Figure courtesy of the Sudbury Neutrino Observatory collaboration.)

The results from the three phases of SNO are in generally good agreement, separately and in combination, establishing a total flux of active neutrinos from ^{8}B decay of $\phi_{NC}(\nu_{active}) = [5.25 \pm 0.16(stat)^{+0.11}_{-0.13}(syst)] \times 10^6$ cm^{-2} s^{-1}, in good agreement with SSM predictions. SNO also established $\phi_{CC}(\nu_e) \sim 0.34\phi_{NC}(\nu_{active})$. Thus, as Figure 11.4 illustrates, about two-thirds of the electron neutrinos produced in the Sun arrive on Earth as heavy-flavor (muon or tauon) neutrinos. The Davis detector and the CC channel in SNO are blind to these heavy flavors, seeing only the portion with electron flavor. Thus the solar neutrino problem was not a matter of missing neutrinos, but rather one of neutrinos in hiding. The implications of this discovery – that neutrinos are massive and violate flavor – are profound, indicating that our standard model of particle physics is incomplete.

Other potential signals of neutrino oscillations in matter, such as an energy-dependent distortion in the ν_e survival probability or day – night differences due to neutrino passage through the Earth, were not seen in SNO, nor have they emerged from SK analyses to date at a convincing level of confidence.

The Borexino experiment [9] (Figure 11.5), located in Italy's Gran Sasso Laboratory, is the first to measure low-energy ($\lesssim$1MeV) solar neutrinos in real time. The detector is housed within

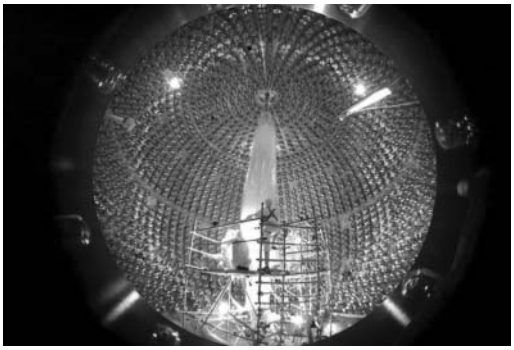


Figure 11.5 The Borexino inner vessel during installation.

a 16.9-m domed tank containing an outer layer of ultrapure water that provides shielding against external neutrons and γ-rays. At the inner edge of the water, a stainless steel sphere serves as a support structure for an array of PMTs that view both the inner detector and the outer water shield, so that the Cherenkov light emitted by muons passing through the water can be used to veto those events. Within the steel sphere, there are two regions, separated by thin nylon vessels, containing high-purity buffer liquid, within which is sequestered a central volume of 278 tons of organic scintillator. The fiducial volume consists of ∼100 tons of the liquid scintillator at the very center of the detector. Scintillation light produced by recoil electrons after ES events is the solar neutrino signal. The 862-keV ^{7}Be neutrinos produce a recoil electron spectrum with a distinctive cutoff edge at 665 keV.

Results reported by the Borexino Collaboration from 2008 to 2012 [9] constrain three low-energy solar neutrino branches. The Collaboration

1. found a ^{7}Be solar rate equivalent to an unoscillated flux of $(3.10 \pm 0.15) \times 10^{9}$ cm^{-2} s^{-1}, or about 62% of the GS98-SFII SSM central value;
2. made the first direct, exclusive determination of the pep flux, $(1.6 \pm 0.3) \times 10^{8}$ cm^{-2} s^{-1} (95% c.l.); and
3. established a limit on the carbon – nitrogen – oxygen CNO neutrino flux, $\phi_{\mathrm{CNO}} < 7.7 \times 10^{8}$ cm^{-2} s^{-1} (95% c.l.)

11.2.3 Neutrino Mass and Oscillations

The phenomenon by which a massive neutrino of one flavor changes into one of a second flavor is called *neutrino oscillations*. Neutrino oscillations have been shown to be responsible not only for the missing solar neutrinos in Davis's experiment but also for the missing atmospheric neutrinos that will be discussed in Chapter 12. Neutrino oscillations can be altered by the presence of matter or the presence of other neutrinos. For this reason, astrophysical "laboratories" for studying neutrinos – the Sun, supernovae, the early universe – are of great interest because of the unique conditions they provide, including very long "baselines" over which neutrinos propagate and enormous matter densities and neutrino fluences.

Neutrino oscillations originate from two distinct sets of labels carried by neutrinos. One is flavor, a property of the weak interaction: an electron neutrino is defined as the neutrino accompanying a positron in β-decay. The other possible label is mass. If a neutrino has a mass m, it propagates through free space with an energy and three momentum related by $\omega = \sqrt{k^2 + m^2}$. Thus neutrino states can be labeled according to flavor, and also labeled according to their masses.

However, nothing requires the neutrinos of definite flavor to be coincident with the neutrinos of definite mass. (In fact, in the analogous case of the quarks, it has long been known that the flavor (or weak interaction) eigenstates are not identical to the mass eigenstates: that is, the up quark decays not only to the down quark but also occasionally to the strange quark.) Neutrino oscillations occur when the mass eigenstates $|\nu_1\rangle$ and $|\nu_2\rangle$ (with masses m_1 and m_2) are related to the weak interaction eigenstates by

$$|\nu_e\rangle = \cos\theta_v |\nu_1\rangle + \sin\theta_v |\nu_2\rangle$$
$$|\nu_\mu\rangle = -\sin\theta_v |\nu_1\rangle + \cos\theta_v |\nu_2\rangle$$

where θ_v, the (vacuum) mixing angle, is nonzero. (Here, for simplicity, we consider just two neutrinos – the generalization to three flavors will be described later.)

In this case, a state produced as a $|\nu_e\rangle$ or a $|\nu_\mu\rangle$ at some time t – for example, a neutrino produced by β-decay in the Sun's core – does not remain a pure flavor eigenstate as it propagates away from the source. The different mass eigenstates comprising the neutrino will accumulate different phases as the neutrino propagates downstream, a phenomenon known as *vacuum oscillations* (vacuum because the experiment is done in free space.) While at time $t = 0$, the neutrino is a flavor eigenstate

$$|\nu(t=0)\rangle = |\nu_e\rangle = \cos\theta_v|\nu_1\rangle + \sin\theta_v|\nu_2\rangle$$

the accumulated phases depend on the mass

$$e^{i(\vec{k}\cdot\vec{x}-\omega t)} = e^{i[\vec{k}\cdot\vec{x}-\sqrt{m_i^2+k^2}\ t]}$$

If the neutrino mass is small compared to the neutrino momentum/energy, one finds

$$|\nu(t)\rangle = e^{i(\vec{k}\cdot\vec{x}-kt-(m_1^2+m_2^2)t/4k)} \times(\cos\theta_v|\nu_1\rangle e^{i\delta m^2 t/4k} + \sin\theta_v|\nu_2\rangle e^{-i\delta m^2 t/4k})$$

There is a common average phase (which has no physical consequence) as well as a beat phase that depends on

$$\delta m^2 = m_2^2 - m_1^2$$

From this one can find the probability that the neutrino state remains a $|\nu_e\rangle$ at time t

$$P_{\nu_e}(t) = |\langle\nu_e|\nu(t)\rangle|^2 = 1 - \sin^2 2\theta_v \sin^2\left(\frac{\delta m^2 c^4 x}{4\hbar c E}\right)$$

The probability oscillates from 1 to $1 - \sin^2 2\theta_v$ and back to 1 over an oscillation length scale

$$L_o = \frac{4\pi\hbar c E}{\delta m^2 c^4}$$

as depicted in Figure 11.6a. In the case of solar neutrinos, if L_o were comparable to or shorter than one astronomical unit, a reduction in the solar ν_e flux would be expected in terrestrial detectors.

The suggestion that neutrinos could oscillate was first made by Pontecorvo in 1958, who pointed out the analogy with $K_0 \leftrightarrow \overline{K}_0$ oscillations. If the Earth – Sun separation is much larger than L_o, one expects an average flux reduction due to oscillations of

$$1 - \frac{1}{2}\sin^2 2\theta_v$$

For a 1 MeV neutrino, this requires $\delta m^2 c^4 \gg 10^{-12}\text{eV}^2$. But such a reduction – particularly given the initial theory prejudice that neutrino mixing angles might be small – did not seem sufficient to account for the factor-of-three discrepancy that emerged from Davis's early measurements.

The view of neutrino oscillations changed when Mikheyev and Smirnov [10] showed in 1985 that neutrino oscillations occurring in matter – rather than in vacuum – could produce greatly enhanced oscillation probabilities. This enhancement comes about because neutrinos propagating through matter acquire an additional mass due to their interactions with the matter. In particular, because the Sun contains many electrons, the electron neutrino becomes heavier in proportion to the local density of electrons. An enhanced probability for oscillations can result when an electron neutrino passes from a

high-density region (such as the solar core) to a low-density one (such as the surface of the Earth.) This matter enhancement is called the MSW mechanism after Mikheyev, Smirnov, and Wolfenstein [11] (who first described the phenomenon of neutrino effective masses).

To explain this enhancement, consider the case where the vacuum mixing angle θ_v is small and $m_2 > m_1$. Then in vacuum $|\nu_e\rangle \sim |\nu_1\rangle \equiv |\nu_L(\rho = 0)\rangle$ where $\rho = 0$ is the electron density in vacuum, that is, the ν_e and the light vacuum eigenstate $|\nu_L(\rho = 0)\rangle$ are almost identical. (Correspondingly, the heavy eigenstate $|\nu_2\rangle \equiv |\nu_H(\rho = 0)\rangle \sim |\nu_\mu\rangle$ in vacuum.) Now what happens in matter? As matter makes the ν_e heavier in proportion to the electron density, if that density is sufficiently high, clearly the electron neutrino must become the (local) heavy mass eigenstate. That is, $|\nu_e\rangle \sim |\nu_H(\rho \to \infty)\rangle$ (and consequently $|\nu_\mu\rangle \sim |\nu_L(\rho \to \infty)\rangle$). That is, we conclude that there must be a local mixing angle $\theta(\rho)$ that rotates from $\theta_v \sim 0$ in vacuum to $\theta(\rho) \sim \pi/2$ as $\rho \to \infty$.

MSW enhancement occurs when the density changes between neutrino production and detection. In particular, electron neutrinos produced in the high-density solar core are created as heavy mass eigenstates. If these neutrinos now propagate to the solar surface adiabatically – this means that the solar density scale height $\rho[d\rho/dx]^{-1}$ is large compared to the local neutrino oscillation length, at all points along the neutrino trajectory – they will remain on the heavy-mass trajectory, and thus exit the Sun as $|\nu_H(\rho = 0)\rangle = |\nu_2\rangle \sim |\nu_\mu\rangle$. That is, there will be an almost complete conversion of the ν_es produced in the solar core to ν_μs. The MSW mechanism is an example of an avoided level crossing, a familiar phenomenon in quantum mechanics.

A schematic comparison of vacuum (a) and matter-enhanced (b) oscillations is shown in Figure 11.6. The matter transition between electron and muon flavors is centered around a density where the vacuum mass difference is just compensated by the matter contributions.

The discussion above was presented for two neutrino flavors, and thus a single vacuum mixing angle θ_v and mass difference δm^2. But three neutrino flavors appear in the standard model of particle physics. In this case, the relationship between flavor ν_e, ν_μ, ν_τ and mass ν_1, ν_2, ν_3 eigenstates is described by the Pontecorvo – Maki – Nakagawa – Sakata (PMNS) matrix [12, 13]

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \\ |\nu_\tau\rangle \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \begin{pmatrix} e^{i\alpha_1/2}|\nu_1\rangle \\ e^{i\alpha_2/2}|\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix} \tag{11.1}$$

where $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. This matrix depends on three mixing angles θ_{12}, θ_{13}, and θ_{23}, of which the first and last are the dominant angles for solar and atmospheric oscillations, respectively; a Dirac phase δ that can induce CP-violating differences in the oscillation probabilities for conjugate channels such as $\nu_\mu \to \nu_e$ versus $\overline{\nu}_\mu \to \overline{\nu}_e$; and two Majorana phases α_1 and α_2 that will affect the interference among mass eigenstates in the effective neutrino mass probed in the lepton-number-violating process of neutrinoless double β-decay. There are

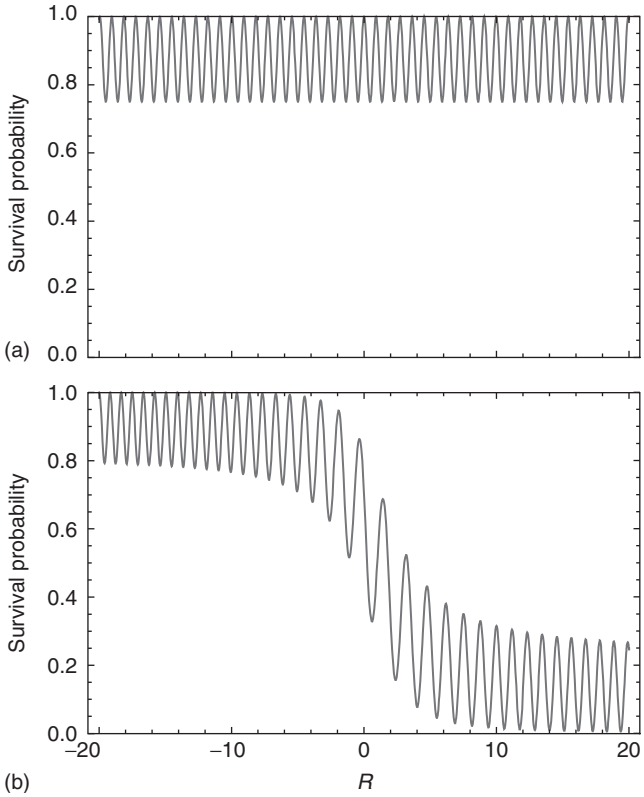


Figure 11.6 A simple example illustrating the MSW mechanism. (a) Vacuum oscillations for a ν_e created at $R = -20$ and propagating to the right, for $\theta_v = 15^\circ$. The average ν_e survival probability is large, 87.5%. [Here the distance R is given in units related to the oscillation length, $4E\cos 2\theta/(\delta m^2 \sin^2 2\theta)$.] In (b) an electron density $\rho(R)$ has been added proportional to $1 - (2/\pi)\tan^{-1} aR$, with a chosen to guarantee adiabaticity, and normalized so that (i) $\rho(r) \to 0$ as $R \to \infty$; (ii) the matter effects cancel the vacuum mass difference for $R \sim 0$ (the MSW crossing point); and (iii) the matter effects reverse the sign of δm^2 as $R \to -\infty$. Thus these are the MSW conditions described in the text. A ν_e created at high density ($R = -20$), where it approximately coincides with the local heavy-mass eigenstate, adiabatically propagates to low density ($R = +20$), where it approximately coincides with the ν_μ. Thus the ν_e survival probability at $R = 20$ is much reduced, compared to the vacuum case. Note that the local oscillation length is maximal near the crossing point.

also two independent mass^2 differences, $\delta m^2_{21} \equiv m^2_2 - m^2_1$ and $\delta m^2_{32} \equiv m^2_3 - m^2_2$.

In this framework, the dominant oscillation affecting solar neutrinos is that described by δm^2_{21} and θ_{12}. The results from SNO, Super-Kamiokande, Borexino, and earlier solar neutrino experiments, and from the reactor experiment KamLAND [14], have determined these parameters quite precisely, as discussed below. Unlike the MSW example given above, the dominant mixing is characterized by a large mixing angle, $\theta_{12} \sim 34^\circ$. Thus the vacuum oscillation probability is significant. The mass^2 difference, $\delta m^2_{21} \sim 7.6 \times 10^{-5}\text{eV}^2$, leads to important matter effects in the

higher-energy portion of the solar neutrino spectrum, thus influencing the rates found in the SNO, Super-Kamiokande, and chlorine experiments. These effects produce a characteristic, energy-dependent distortion of the solar ν_e spectrum.

11.2.4 Solar Neutrinos: Oscillation Parameters and Outlook

Neutrino oscillations proved to be responsible for both the solar neutrino problem and the atmospheric neutrino problem, the subject of Section 11.3. This phenomenon requires both flavor mixing and neutrino mass, phenomena that can be accommodated in various extensions of the standard model of particle physics. The great interest in neutrino astrophysics stems in part from the expectation that newly discovered neutrino properties may help us formulate the correct extension. Indeed, the tiny neutrino mass differences deduced from the solar and atmospheric problems are compatible with theories where the neutrino mass is inversely proportional to a scale for new physics of about 10^{15}GeV. This is close to the grand unified scale where supersymmetric extensions of the standard model predict that the strengths of the fundamental forces unify.

The precision with which fundamental parameters of the standard model of particle physics were determined by solar neutrino experiments, supplemented by KamLAND's measurements to better constrain δm^2_{21}, is really quite remarkable. The global analysis performed by the SNO group of this set of experiments yielded

$$\begin{aligned}\sin^2\theta_{12} &= 0.308 \pm 0.014\\ \delta m^2_{21} &= (7.41^{+0.21}_{-0.19})10^{-5}\text{eV}^2\\ \sin^2\theta_{13} &= 0.025^{+0.018}_{-0.015}\end{aligned}$$

where θ_{13} is the mixing angle for the subdominant 1–3 oscillation. The first two results are in excellent agreement, in value and uncertainty, with the corresponding values from the Bari [15] and Valencia [16] global analyses that include all accelerator and reactor data. The main impact of terrestrial neutrino experiments on solar neutrino analyses has come from the Daya Bay, RENO, and Double Chooz reactor experiments, which have significantly tightened the constraints on θ_{13},

$$\sin^2\theta_{13} = \begin{cases} 0.0243^{+0.0027}_{-0.0026} & \text{Bari} \\ 0.0248^{+0.0031}_{-0.0029} & \text{Valencia} \end{cases}$$

where the error bars include the uncertainty in the mass hierarchy (i.e., the sign of δm^2_{32}). Yet even in this case, the solar neutrino analysis gave the correct central value for θ_{13}.

The effects of matter on neutrino oscillations have also been convincingly demonstrated from the comparison of the Borexino result for ^{7}Be neutrinos and those of SNO and Super-Kamiokande for ^{8}B neutrinos. The ^{7}Be neutrinos lie in the vacuum oscillation region where the survival probability is larger; the SNO and Super-Kamiokande detectors are sensitive to more energetic ^{8}B neutrinos, where matter effects enhance the oscillation into heavy-flavor states.

There are also interesting developments involving the SSM, as we have noted in our discussion of two competing SSMs, GS98-SFII and AGSS09-SFII. One of the important validations of the SSM has come through helioseismology, the measurement of solar surface fluctuations, as deduced from the Doppler shifts of spectral lines. The observed patterns can be inverted to determine properties of solar interior, including the speed of sound as a function of the solar radius. For

about a decade the agreement between SSM predictions and observation had been excellent.

But recently improved three-dimensional models of the photosphere, when applied to the analysis of absorption lines, led to a ∼30% downward revision in convective-zone metal abundances (and thus $Z_{\rm ini}$). If the new abundances are employed in the SSM, the resulting changes in the opacity alter both neutrino flux predictions and the sound speed profile. This solar abundance problem could indicate that the SSM assumption of a homogeneous zero-age Sun – which is not obviously correct due to the efficiency with which planet formation swept metals from the protosolar disk – may have to be reexamined. One of the goals of the next solar neutrino experiment – SNO+, a larger, deeper version of Borexino under construction in the cavity that formerly held SNO – is to use CN neutrinos to measure the core abundance of carbon and nitrogen directly, thereby allowing a direct comparison between core and surface metal abundances [5].

11.3 Atmospheric Neutrinos

The atmospheric neutrino problem developed very much in parallel with the solar neutrino problem and also involved missing neutrinos. The first definitive claim that neutrinos are massive came from the atmospheric neutrino group associated with Super-Kamiokande, in 1998. The oscillations seen in atmospheric neutrinos differ from those seen in solar neutrinos, resulting from the coupling of a different pair of neutrinos.

11.3.1 The Neutrino Source

When primary cosmic-ray protons and nuclei hit the upper atmosphere, the ensuing nuclear reactions with atmospheric oxygen and nitrogen nuclei produce secondaries such as pions, kaons, and muons. Atmospheric neutrinos arise from the decay of these secondaries. For energies $\lesssim 1$GeV, the secondaries decay before reaching the Earth's surface

$$\pi^{\pm}(K^{\pm}) \rightarrow \mu^{\pm} + \nu_{\mu}(\overline{\nu}_{\mu}),$$
$$\mu^{\pm} \rightarrow e^{\pm} + \nu_{e}(\overline{\nu}_{e}) + \overline{\nu}_{\mu}(\nu_{\mu}) \qquad (11.2)$$

Consequently, one expects the ratio

$$r = (\nu_{e} + \overline{\nu}_{e})/(\nu_{\mu} + \overline{\nu}_{\mu}) \qquad (11.3)$$

to be approximately 0.5 in this energy range. Detailed Monte Carlo calculations, including the effects of muon polarization, give $r \sim 0.45$. This ratio should be rather insensitive to theoretical uncertainties. It does not depend on absolute fluxes, and as a ratio of related processes, one expects many sources of systematic error to cancel. Indeed, while various groups have estimated this ratio, sometimes starting with neutrino fluxes that vary in magnitude by up to 25%, agreement in the ratio has been found at the level of a few percent. This agreement persists at higher energies, where r decreases because higher-energy muons survive passage through the atmosphere, owing to the effects of time dilation.

Atmospheric neutrinos are a very attractive astrophysical source for experimenters. Apart from relatively minor geomagnetic effects, atmospheric neutrino production is uniform over the Earth. Thus an experimenter, operating an underground detector at some location, can make use of a

set of nearly equivalent neutrino sources at distances ranging from tens of kilometers (directly overhead) to 13 000 km (directly below, produced on the opposite side of the Earth). Effects such as neutrino oscillations, which depend on the distance from the source to the target, might show up as a characteristic dependence of neutrino flux on the zenith angle, provided the relevant oscillation length is comparable to or less than the Earth's diameter. (Note that the solar neutrino δm^2_{21}, for atmospheric neutrinos of energy $\sim$ 1GeV, would not satisfy this condition, as the oscillation length is several times the Earth's diameter.)

11.3.2 Atmospheric Neutrinos and Proton Decay Detectors

The atmospheric neutrino anomaly grew out of efforts to build large underground detectors for proton decay, one of the phenomena expected in the grand unified theories that were formulated in the late 1970s and early 1980s. As atmospheric neutrinos and proton decay would deposit very similar energies in such detectors, studies of atmospheric neutrinos were a natural second use of such detectors. Significant indications of an anomaly came from the Irvine – Michigan – Brookhaven (IMB) [17] and Kamiokande [18] proton decay detectors. IMB first noticed a possible deficit of neutrino-induced muon events in 1986, while Kamiokande established a deficit in excess of 4σ by 1988. By 1998, this anomaly was also apparent in data from the Soudan detector and from Super-Kamiokande.

The quantity determined in such experiments is a ratio (observed to predicted) of ratios

$$R = \frac{(\nu_\mu/\nu_e)_{\text{data}}}{(\nu_\mu/\nu_e)_{\text{Monte Carlo}}}$$

where the numerator is determined experimentally, and the denominator calculated. Agreement between data and theory thus requires $R \sim 1$. Early experimenters faced a difficulty in evaluating this ratio owing to limited statistics: the counting rates were too low to allow a detailed analysis based on the zenith angle, that is, based on the neutrino path length. This changed with the construction of Super-Kamiokande, which provided a fiducial volume of about 20 ktons. An early analysis from Super-Kamiokande found

$$R = 0.61 \pm 0.03(\text{stat}) \pm 0.05(\text{syst})$$

for sub-GeV events that were fully contained in the detector and

$$R = 0.66 \pm 0.05(\text{stat}) \pm 0.08(\text{syst})$$

for fully and partially contained multi-GeV volts events. In addition, in 1998, the collaboration presented an analysis based on 33 kton years of data, showing a zenith angle dependence inconsistent with theoretical calculations of the atmospheric flux, in the absence of oscillations [19]. This indicated a distance dependence in the muon deficit, a signature of oscillations. Furthermore, the parameters of the oscillation, particularly $5 \times 10^{-4}\text{eV}^2 < |\delta m^2_{32}| < 6 \times 10^{-3}\text{eV}^2$, differed from those that would later be determined from solar neutrinos. The collaboration concluded that the data were consistent with the two-flavor oscillation $\nu_\mu \rightarrow \nu_\tau$. This was the first definitive claim for massive neutrinos.

SK-I collected approximately 15,000 atmospheric neutrino events in nearly 5 years of running. The collaboration's zenith angle analysis of the data found evidence of a first oscillation minimum at $L/E \sim 500$ km GeV^{-1}, so that $L_o \sim 1000$ km for a 1 GeV muon neutrino. Current

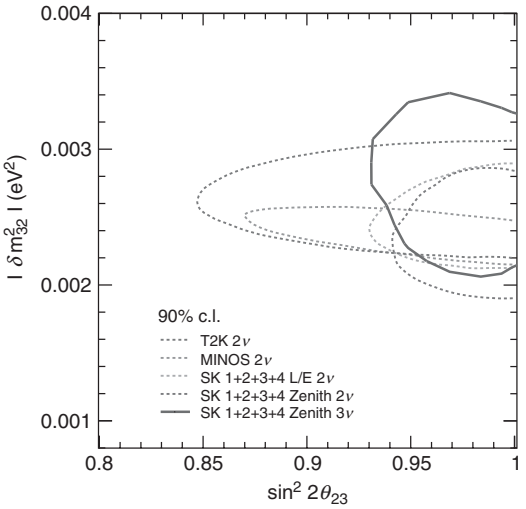


Figure 11.7 The Super-Kamiokande I–IV three-neutrino analysis of atmospheric neutrino oscillations (solid curve), including comparisons to the long-baseline accelerator results of MINOS and T2K (the two ellipses that extend to the left). (Source: From [20], with permission of the Super-Kamiokande Collaboration.)

results are based on 3903 days of data from SKI–IV, including 1097 days from the current phase IV. The entire dataset has been reanalyzed with a common set of improved tools, in a three-flavor analysis in which $\sin^2\theta_{13}$ was fixed to 0.025, based on the recent reactor neutrino results from the Daya Bay, RENO, and Double Chooz experiments. The analysis yielded [20]

Normal hierarchy:

$$\delta m^2_{32} = (2.66^{+0.15}_{-0.40}) \times 10^{-3}\ \text{eV}^2 (1\sigma)$$
$$\sin^2\theta_{23} = 0.425^{+0.194}_{-0.034}\ (90\%\text{c.l.})$$

Inverted hierarchy:

$$\delta m^2_{32} = -(2.66^{+0.17}_{-0.23}) \times 10^{-3}\ \text{eV}^2 (1\sigma)$$
$$\sin^2\theta_{23} = 0.575^{+0.055}_{-0.182}\ (90\%\text{c.l.})$$

The results continue to allow maximal mixing, $\sin^2\theta_{23} \sim 0.5$. The significant difference between the normal and inverted hierarchy best values for $\sin^2\theta_{23}$ reflects the flatness of the χ^2 fit around $\sin^2\theta_{23} \sim 0.5$, with slight local minima appearing above and below this value. The atmospheric value for δm^2_{32} is consistent with the somewhat more precise values from the MINOS (and T2K) long-baseline experiments. The MINOS beam-neutrino analysis, which includes atmospheric neutrino data obtained with that detector, determined $|\delta m^2_{32}| = 2.39^{+0.09}_{-0.10} \times 10^{-3}\ \text{eV}^2$ [21], a number consistent with that from the SK atmospheric analysis (Figure 11.7).

11.3.3 Outlook

While a great deal of new physics has been learned from experiments on atmospheric and solar neutrinos, several important questions remain [22]:

- Oscillation experiments are sensitive to differences in the squared masses. They are not sensitive to absolute neutrino masses. We do know, from the atmospheric $|\delta m^2_{32}|$, that at least one neutrino must have a mass $\gtrsim$0.05 eV. But the best laboratory bound, from tritium β-decay experiments, would allow neutrino masses 50 times greater. That is, the three light neutrinos might be nearly equal in mass, split by the tiny

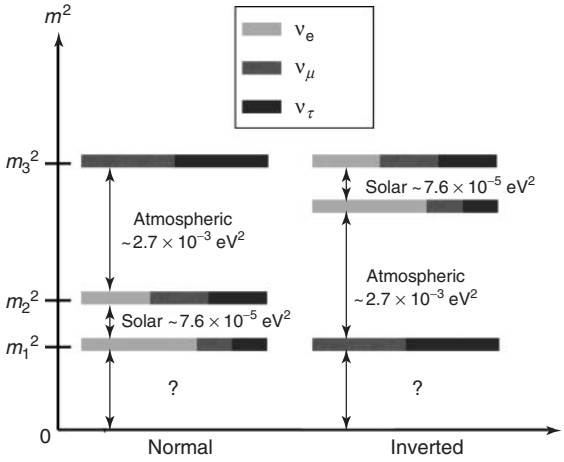


Figure 11.8 Illustration of the two level schemes that are possible for the light neutrinos, given that no matter effects have yet been seen in atmospheric neutrinos.

mass differences indicated by solar and atmospheric neutrino oscillations.

- Matter effects (from passage through the Earth) have not been seen in atmospheric neutrino experiments. This leaves open two possible orderings (normal, inverted) of the mass eigenstates, as indicated above and as illustrated in Figure 11.8. Atmospheric neutrino detectors that can determine the sign of the produced lepton are of potential interest in resolving the hierarchy question.
- Very little is known about the three CP-violating phases that appear in the PMNS matrix. The Dirac phase could be determined by looking for differences between certain conjugate oscillation channels, such as $P(\nu_e \to \nu_\mu)$ and $P(\overline{\nu}_\mu \to \overline{\nu}_e)$, one of the goals of future long-baseline neutrino experiments. Finding such a difference could be important to theories that attribute the excess of matter over antimatter in our universe to leptonic CP violation (although the precise connection between low-energy CP violation and the high-energy mechanism responsible for baryogenesis may be difficult to define).
- The neutrino, lacking an electric or any other charge that must flip sign under particle – antiparticle conjugation, is unique among standard model particles in that it may be its own antiparticle. So far no measurement has been made that can distinguish this possibility (a Majorana neutrino) from the case where the ν and $\overline{\nu}$ are distinct (a Dirac neutrino). Next-generation neutrinoless double β-decay experiments

$$(N, Z) \to (N - 2, Z + 2) + 2e^-$$

could settle this issue, however. This process requires lepton number violation and Majorana masses. The two remaining CP-violating phases are Majorana phases that can affect neutrinoless double β-decay rates.
- No compelling argument has been given to account for the large mixing angles deduced from atmospheric and solar neutrino oscillations. These angles differ markedly from their measured counterparts among the quarks.

While some of these questions may be answered in terrestrial experiments,

neutrino astrophysics will continue to offer unique environments for probing fundamental neutrino properties. Several examples are given in the chapter on neutrino cooling.

11.4 Supernova Neutrinos and Nucleosynthesis

The bursts associated with a core-collapse supernova are among the most interesting sources of neutrinos in astrophysics [23]. A massive star, in excess of 10 solar masses, begins its lifetime burning the hydrogen in its core under the conditions of hydrostatic equilibrium. When the hydrogen is exhausted, the core contracts until the density and temperature are reached where $3\alpha \rightarrow {}^{12}C$ can take place. The helium is then burned to exhaustion. This pattern (fuel exhaustion, contraction, heating, and ignition of the ashes of the previous burning cycle) repeats several times, leading finally to the explosive burning of Si to Fe. For a heavy star, the evolution is rapid because of the amount of energy the star must produce to support itself against its own gravity. A 25 solar mass star would go through the set of burning cycles in about 7 My, with the final explosive Si burning stage taking a few days. The result is an ''onion skin'' structure of the precollapse star in which the star's history can be read by looking at the surface inward: there are concentric shells dominated by H, ^{4}He, ^{12}C, ^{16}O and ^{20}Ne, ^{28}Si, and ^{56}Fe at the center.

11.4.1 The Explosion Mechanism and Neutrino Burst

The source of energy for this evolution is nuclear binding energy. A plot of the nuclear binding energy δ as a function of nuclear mass shows that the minimum is achieved at Fe. In a scale where the ^{12}C mass is picked as zero:

^{12}C $\quad \delta/\text{nucleon} = 0.000$ MeV
^{16}O $\quad \delta/\text{nucleon} = -0.296$ MeV
^{28}Si $\quad \delta/\text{nucleon} = -0.768$ MeV
^{40}Ca $\quad \delta/\text{nucleon} = -0.871$ MeV
^{56}Fe $\quad \delta/\text{nucleon} = -1.082$ MeV
^{72}Ge $\quad \delta/\text{nucleon} = -1.008$ MeV
^{98}Mo $\quad \delta/\text{nucleon} = -0.899$ MeV

Once the Si burns to produce Fe, there is no further source of nuclear energy adequate to support the star. So as the last remnants of nuclear burning take place, the core is largely supported by degeneracy pressure, with the energy generation rate in the core being less than the stellar luminosity. The core density is about $2 \times 10^9 \text{g cm}^{-3}$ and the temperature is $kT \sim 0.5$ MeV.

Thus, the collapse that begins with the end of Si burning is not halted by a new burning stage, but continues. As gravity does work on the matter, the collapse leads to a rapid heating and compression of the matter. Sufficient heating of the Fe can release αs and a few nucleons, which are bound by ∼ 8MeV. At the same time, the electron chemical potential is increasing. This makes electron capture on nuclei and any free protons favorable,

$$e^- + p \rightarrow \nu_e + n$$

As the chemical equilibrium condition is

$$\mu_e + \mu_p = \mu_n + \langle E_\nu \rangle$$

the increase in the electron Fermi surface with density will lead to increased neutronization of the matter, as long as neutrinos freely escape the star. These escaping neutrinos carry off energy and lepton number. Both the electron capture and

the nuclear excitation and disassociation take energy out of the electron gas, which is the star's only source of support. Consequently, the collapse is very rapid, with numerical simulations finding that the star's iron core (~1.2–1.5 solar masses) collapses at about 0.6 of the free-fall velocity.

While the ν_es readily escape in the early stages of infall, conditions change once the density reaches $\sim 10^{12}\,\mathrm{g\,cm^{-3}}$. At this point, neutrino scattering off the matter through both CC and coherent NC processes begins to alter the transport. The NC neutrino scattering off nuclei is particularly important, as the scattering amplitude is proportional to the total nuclear weak charge, which is approximately the neutron number. Coherent NC scattering transfers very little energy because the mass of the nucleus is so much greater than the typical energy of the neutrinos. But momentum is exchanged. Because of repeated scattering, the neutrino ''random walks'' out of the star. When the neutrino mean free path becomes sufficiently short, the time required for the neutrino to diffuse out of the high-density core begins to exceed the time required to complete the collapse. Above densities of about $10^{12}\,\mathrm{g\,cm^{-3}}$, or ~1% of nuclear density, such neutrino trapping occurs. Consequently, once this critical density is exceeded, the energy released by further gravitational collapse is trapped within the star until after core bounce. Similarly, the star no longer can lose lepton number due to neutrino emission.

For a neutron star of 1.4 solar masses and a radius of 10 km, an estimate of its binding energy is

$$\frac{GM^2}{2R} \sim 2.5 \times 10^{53}\mathrm{ergs} \qquad (11.4)$$

Thus this is roughly the trapped energy that will later be radiated in neutrinos, after core bounce, as the protoneutron star formed in the collapse cools.

The collapse produces a shock wave that is critical to subsequent ejection of the star's mantle. The velocity of sound in matter rises with increasing density. Late in the collapse the sound velocity in the inner portion of the iron core, with M_{HC} ~0.6–0.9 solar masses, exceeds the infall velocity. Any pressure variations that may develop during infall can even out before the collapse is completed. Consequently, this portion of the iron core collapses as a unit, retaining its density profile.

The collapse of the core continues until nuclear densities are reached. As nuclear matter is rather incompressible ($\sim 200\,\mathrm{MeV\,f^{-3}}$), the nuclear equation of state is effective in halting the collapse: maximum densities of 3–4 times nuclear density are reached, for example, perhaps $6 \times 10^{14}\,\mathrm{g\,cm^{-3}}$. The innermost shell of matter reaches this supernuclear density first, rebounds, sending a pressure wave out through the inner core. This wave travels faster than the infalling matter, as the inner iron core is characterized by a sound speed in excess of the infall speed. Subsequent shells follow. The resulting pressure waves collect at the edge of the inner iron core – the radius at which the infall velocity and sound speed are equal. As this point reaches nuclear density and comes to rest, a shock wave breaks out and begins its traversal of the outer core.

Initially, the shock wave may carry an order of magnitude more energy than is needed to eject the mantle of the star ($< 10^{51}$ ergs). But as the shock wave travels through the outer iron core, it heats and melts the iron that crosses the shock front, at a loss of ~ 8 MeV nucleon. Additional energy is lost by neutrino emission, which increases after the melting. These losses are comparable to the initial energy

carried by the shock wave. Most simplified (e.g., one-dimensional) numerical models fail to produce a successful "prompt" hydrodynamic explosion, for this reason. The shock stalls near the edge of the iron core, instead of propagating into the mantle.

Most of the theoretical attention in the past decade has focused on the role of neutrinos in reviving this shock wave, a process that becomes more effective in multidimensional models that account for convection. In this delayed mechanism, the shock wave stalls at a radius of 200–300 km, some tens of milliseconds after core bounce. But neutrinos diffusing out of the protoneutron star react frequently in the nucleon gas left in the wake of the shock wave, depositing significant energy. Over ∼0.5s the increasing pressure due to neutrino heating of this nucleon gas helps push the shock outward. This description is oversimplified – a variety of contributing effects are emerging from numerical simulations – but there is wide agreement that energy deposition by neutrinos is an essential ingredient for successful explosions.

Regardless of explosion details, neutrinos dominate supernova energetics. The kinetic energy of the explosion and the supernova's optical display account for $> 1\%$ of the available energy. The remaining 99% of the 3×10^{53} ergs released in the collapse is radiated in neutrinos of all flavors. The timescale governing the leakage of trapped neutrinos out of the protoneutron star is about 3s. The energy is roughly equipartitioned among the flavors (a consequence of reactions among trapped neutrinos that equilibrate flavor). The detailed decoupling of the emitted neutrinos from the matter – which occurs at a density of about $10^{11} - 10^{12}$ g cm^{-3} –does depend on flavor. This leads to differences in neutrino temperatures, with electron neutrinos being somewhat cooler ($T \sim 3.5$ MeV) than the heavy-flavor neutrinos ($T \sim 6$ MeV). The radius for neutrino–matter decoupling defines a "neutrinosphere" deep within the star, analogous to the familiar photosphere for optical emissions.

The burst of neutrinos produced in a galactic core-collapse supernova is detectable with instruments such as Super-Kamiokande and SNO. On February 23, 1987, a neutrino burst from a supernova in the Large Magellanic Cloud was observed in the proton decay detectors Kamiokande and IMB [24]. The optical counterpart reached an apparent magnitude of about 3 and could be observed easily in the night sky with the naked eye. This supernova originated 160,000 light years from Earth. Approximately 20 events were seen in the Kamiokande and IBM detectors, spread over approximately 10 s. Within the limited statistics possible with these first-generation detectors, the number of events and the burst duration were consistent with standard estimates of the energy release and cooling time of a supernova. The neutrino data from the two detectors are shown in Figure 11.9.

Temperature differences between neutrino flavors are interesting because of oscillations and nucleosynthesis. The discussion of matter effects in the solar neutrino problem was limited to two flavors. But the higher densities found in core-collapse supernovae make all three flavors relevant. The three-flavor MSW level-crossing diagram is shown in Figure 11.10. One sees, in addition to the neutrino "level crossing" $\nu_\mu \leftrightarrow \nu_e$ important for solar neutrinos, a second crossing of the ν_e with the ν_τ. The higher density characterizing this second crossing, $\sim 10^4$ g cm^{-3}, is determined by atmospheric mass difference $|\delta m^2_{32}|$ and by the typical energy of

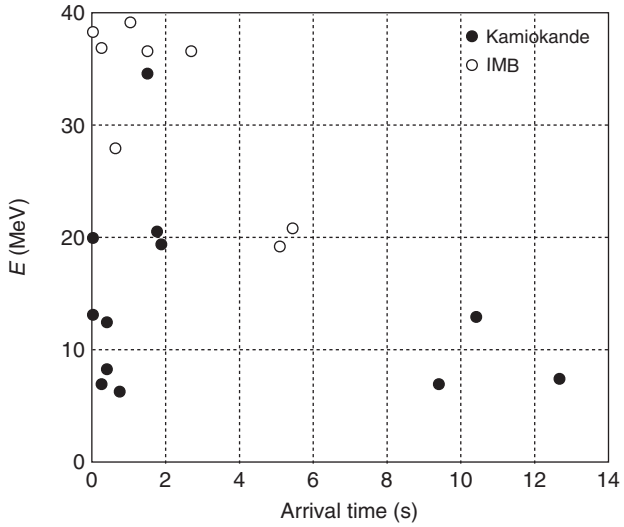


Figure 11.9 The timing and energy of neutrino events from SN1987A, Source: as observed in the Kamiokande and IMB detectors [24].

supernova neutrinos, ∼10–20 MeV. This density is beyond that available in the Sun ($\lesssim 10^2$ g cm^{-3}), but far less than that of a supernova's neutrinosphere. Consequently, this second level crossing alters neutrino flavor only after the supernova neutrinos are free-streaming out of the star, with well-defined spectra that are approximately thermal. This level crossing can exchange the flavor labels on the ν_e and ν_τ spectra, so that the ν_es become hotter than the heavy-flavor neutrinos. One would expect such an inversion to be apparent in terrestrial supernova neutrino detectors.

In fact, this description oversimplifies the neutrino physics of supernovae. The enormous neutrino densities encountered in a supernova lead to a new aspect of the MSW effect – oscillations altered not by neutrino–electron scattering, but by neutrino–neutrino scattering [25]. While the precise consequences of

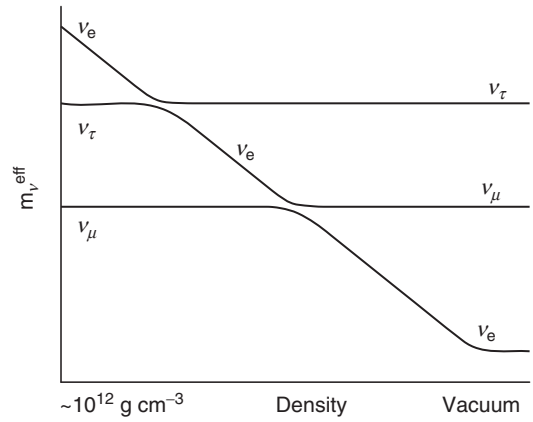


Figure 11.10 A schematic illustration of the two level crossings that the ν_e would experience, given the higher densities available in the mantle of a massive star undergoing core collapse. A normal hierarchy has been assumed.

this neutrino–neutrino MSW potential are still being explored – the problem is both nonlinear and dependent on the angles and flavors of interacting neutrinos – the effects reach much deeper into the star and alter the flavor physics in distinctive ways. Consequently, this novel flavor physics could play a role in the dynamics of the explosion. Supernovae likely provide the only environment in nature where neutrino–neutrino interactions dominate the MSW potential.

11.4.2 Supernova Neutrino Physics

This novel neutrino–neutrino MSW potential is one of many reasons core-collapse supernovae play an important role in neutrino astrophysics. Others include the following:

- As ∼99% of the collapse energy is radiated in neutrinos, one can in principle deduce the binding energy of the neutron star from neutrino flux measurements, provided other parameters (such as the distance to the supernova) are sufficiently well known. The recent observation of a 2 solar mass neutron star suggests that the nuclear equation of state is stiff at high density, disfavoring very compact neutron stars.
- Neutrinos from galactic supernovae will not be obscured by intervening matter or dust, unlike optical signals. Thus supernova neutrino bursts should, over the next few hundred years, provide our most reliable measure of the contemporary rate of galactic core collapse.
- There exists a so-far undetected diffuse background of supernova neutrinos, produced by all past supernovae occurring in the universe. Future detectors that may approach the megaton scale should be able to see a few events from this source. Detection of these neutrinos would place an important constraint on the inventory of massive stars undergoing core collapse, from the first epoch of star formation until now.
- Supernovae are one of the most important engines for nucleosynthesis, controlling much of the chemical enrichment of the galaxy. As described in the next section, neutrinos are directly and indirectly involved in this synthesis.
- While most of the energy released in core collapse is radiated as neutrinos over the first several seconds, neutrino emission at a lower level continues as the proto-neutron star cools and radiates away its lepton number. It is quite possible that phase changes in the dense nuclear matter could occur several tens of seconds after core bounce, altering the late-time neutrino ''light curve'' in a characteristic way. While there are considerable uncertainties in estimates, neutrino processes may continue to dominate neutron star cooling for $\sim 10^5$ years.
- The neutrino burst could include other sharp features in time, marking interesting astrophysics. The melting of iron to nucleons with the passage of the shock wave through the outer iron core is predicted to produce a spike in the neutrino luminosity, lasting for a few milliseconds. Continued accretion onto the neutron star surface could produce a collapse to a black hole, and consequently a sudden termination in neutrino emission.
- Supernova cooling times place constraints on new physics associated with particles that also couple weakly to matter. For example, a light scalar called the *axion* could, in principle, compete

with neutrinos in cooling a supernovae. The requirement that axion emission should not shorten the cooling time too much, which would be in conflict with SN1987A data, constrains the mass and coupling of this hypothesized particle.

11.5 Neutrinos and Nucleosynthesis

Neutrinos and nucleosynthesis are both associated with explosive environments found in astrophysics. This section discusses three examples, the Big Bang, the neutrino process, and the r-process.

11.5.1 Cosmological Neutrinos, BBN, and Large-Scale Structure

One of the classic problems in nucleosynthesis and cosmology is accounting for the mass fraction of primordial helium of 25%, as well as the abundances of D, ^{3}He, and ^{7}Li. The ^{4}He mass fraction shows that nuclei were synthesized from the early universe nucleon soup at a time when the n/p ratio was $\sim$1/7. This requires the proper coordination of two ''clocks,'' one being the weak interaction rate for decay of neutrons to protons, and the other the Hubble rate governing the expansion of the universe,

$$H(t) \equiv \frac{1}{R(t)}\frac{\mathrm{d}R(t)}{\mathrm{d}t} = \sqrt{\frac{8\pi G\rho(t)}{3}}$$

where G is the gravitational constant. This clock depends on the energy density $\rho(t)$ which, at this epoch, is dominated by relativistic particles, including the neutrinos. Big Bang nucleosynthesis (BBN) depends on one adjustable parameter, the baryon density, which is usually given in terms of the ratio of baryons to photons η. Perhaps the first quantitative result in cosmology was the BBN determination of $\eta \sim (5.9 \pm 0.8) \times 10^{-10}$ [26], from which one can deduce the baryon density $\rho_\mathrm{b} \sim 4.2 \times 10^{-31}\,\mathrm{g\,cm^{-3}} \sim 0.044\rho_\mathrm{crit}$, where ρ_crit is the closure density. The comparison of ρ_b with various astrophysical measurements of the amount of gravitating matter in the universe shows that approximately 85% of that matter is dark, nonbaryonic, and consequently outside the standard model.

An independent and more precise determination of η has been made from the pattern of acoustic peaks seen in mappings of the cosmic microwave background. The 7-year Wilkinson microwave anisotropy probe (WMAP) result is $(6.19 \pm 0.15) \times 10^{-10}$ [27]. The two determinations, one based on the nucleosynthesis that took place 3 min after the Big Bang and the second connected with the pattern of large-scale structure at the time the first atoms formed 380,000 years later, are in remarkable agreement.

The primordial abundance of ^{4}He is relatively insensitive to η, in contrast to other BBN species like D. Consequently, this abundance is a rather good test of the number of neutrino flavors contributing to early universe expansion. BBN analyzes over a number of years have generally favored $N_\nu \lesssim 3$, though calculations using $N_\nu = 3$ yield the observed abundances of both D and ^{4}He at 68%,c.l. [26]. However, recent reanalyses of the ^{4}He abundance [28] have led to an upward revision, so that $N_\nu \sim 3.74^{+0.86}_{-0.76}$. The 7-year WMAP analysis of the cosmic microwave background (CMB) constrains a similar quantity, the effective number of relativistic species, yielding $N_\mathrm{eff} \sim 4.34 \pm 0.87$, where ~ 3 is expected [27]. This modest discrepancy as well as the disagreement between the BBN primordial ^{7}Li prediction and Li

abundances deduced from old, metal-poor stars have stimulated recent interest in nonstandard BBN scenarios [29, 30].

Because neutrinos are massive and relativistic in the early universe, they are an interesting component of dark matter. The lighter the scale of neutrino mass, the longer they remain relativistic, the further they stream, and the more effective they are in suppressing the growth of large-scale structure. The evolution of neutrino mass effects on structure growth is distinctive in both redshift Z and spatial scale, altering the distribution of baryons and cold dark matter at the ∼1% level when their contribution to the critical density is just $\Omega_\nu \sim 0.1\%$. The seven-year WMAP analysis yielded [27]

$$\sum_i m_\nu(i) \lesssim 0.58 \quad \text{eV}$$

though tighter bounds have been claimed in other, combined analyses. A variety of cosmological surveys planned for the next decade have the anticipated statistical power to test neutrino mass effects with an order of magnitude or more increase in sensitivity [31]. As the atmospheric neutrino mass difference implies a lower bound on neutrino mass of ~ 50 meV, the absolute scale of neutrino mass could be determined cosmologically. A cosmological bound significantly below 100 meV would also imply that the hierarchy is normal.

These arguments assume that cosmological measurements of the energy density in relativistic species constrain standard model neutrinos. There are a number of possible exceptions to this assumption that are of significant current interest, such as the consequences of a net lepton number asymmetry in the universe, or the presence of a sterile neutrino (a neutrino lacking standard model couplings) that might mix with active species.

11.5.2 The Neutrino Process

One of the more amusing roles for neutrinos in nucleosynthesis is found in the neutrino process, the direct synthesis of new elements through neutrino reactions. Core-collapse supernovae provide the enormous neutrino fluences necessary for such synthesis to be significant. They also eject newly synthesized material into the interstellar medium, where it can be incorporated into a new generation of stars.

Among the elements that might be made primarily or partially in the ν-process, the synthesis of ^{19}F is one of the more interesting examples [32]. The only stable isotope of fluorine, ^{19}F has an abundance

$$\frac{^{19}\text{F}}{^{20}\text{Ne}} \sim \frac{1}{3100}$$

Ne is one of the hydrostatic burning products in massive stars, produced in great abundance and ejected in core-collapse supernovae. Thus a mechanism that converts ∼0.03% of the ^{20}Ne in the star's mantle into ^{19}F could account for the entire observed abundance of the latter.

The Ne zone in a supernova progenitor star is typically located at a radius of ∼20,000 km. A simple calculation that combines the neutrino fluence through the Ne zone with the cross section for inelastic neutrino scattering off ^{20}Ne shows that approximately 0.3% of the ^{20}Ne nuclei would interact with the neutrinos produced in the core collapse. Almost all of these reactions result in the

production of ^{19}F, for example,

$$\begin{aligned} {}^{20}\mathrm{Ne}(\nu,\nu'){}^{20}\mathrm{Ne}^* &\rightarrow {}^{19}\mathrm{Ne} + \mathrm{n} \\ &\rightarrow {}^{19}\mathrm{F} + \mathrm{e}^+ + \nu_\mathrm{e} + \mathrm{n} \\ {}^{20}\mathrm{Ne}(\nu,\nu'){}^{20}\mathrm{Ne}^* &\rightarrow {}^{19}\mathrm{F} + \mathrm{p} \end{aligned} \qquad (11.5)$$

with the first reaction occurring half as frequently as the second. Thus one would expect the abundance ratio to be $^{19}\mathrm{F}/^{20}\mathrm{Ne} \sim 1/300$, corresponding to an order of magnitude more ^{19}F than found in nature.

This example shows that stars are rather complicated factories for nucleosynthesis. Implicit in the reactions above are mechanisms that also destroy ^{19}F. For example, about 70% of the neutrons coproduced with ^{19}F in the first reaction immediately recapture on ^{19}F, destroying the product of interest. Similarly, many of the coproduced protons destroy ^{19}F via $^{19}\mathrm{F}(\mathrm{p},\alpha)^{16}\mathrm{O}$ – unless the star is rich in ^{23}Na, which readily consumes protons via $^{23}\mathrm{Na}(\mathrm{p},\alpha)^{20}\mathrm{Ne}$. Finally, some of the ^{19}F produced in the neon shell is destroyed when the shock wave passes through that zone: the shock wave can heat the inner portion of the Ne zone above 1.7×10^9K, the temperature at which ^{19}F can be destroyed by $^{19}\mathrm{F}(\gamma,\alpha)^{15}\mathrm{N}$.

If all of this physics is treated carefully in a nuclear network code, one finds that the desired $^{19}\mathrm{F}/^{20}\mathrm{Ne} \sim 1/3100$ is achieved for a heavy-flavor neutrino temperature of about 6 MeV. This is quite consistent with the temperatures that come from supernova models.

The neutrino process produces interesting abundances of several relatively rare, odd-A nuclei including ^{7}Li, ^{11}B, ^{138}La, ^{180}Ta, and ^{15}N. CC neutrino reactions on free protons can produce neutrons that, through (n,p) and (n,γ) reactions, lead to the nucleosynthesis of the so-called p-process nuclei from $A = 92$ to 126. The production of such nuclei has been a long-standing puzzle in nuclear astrophysics.

11.5.3 The r-Process

Beyond the iron peak, nuclear Coulomb barriers become so high that charged-particle reactions become ineffective, leaving neutron capture as the mechanism responsible for producing the heaviest nuclei. If the neutron abundance is modest, this capture occurs in such a way that each newly synthesized nucleus has the opportunity to β-decay, if it is energetically favorable to do so. Thus weak equilibrium is maintained among the nuclei, so that synthesis is along the nuclear valley of stability. This is called the *s-process* or *slow-process*. However, a plot of the s-process in the (N, Z) plane reveals that this path misses many stable, neutron-rich nuclei. This suggests that another mechanism is also at work. Furthermore, the abundance peaks found in nature near masses $A \sim 130$ and $A \sim 190$, which mark the closed neutron shells where neutron capture rates and β-decay rates are slower, each split into two subpeaks. One set of subpeaks corresponds to the closed-neutron-shell numbers $N \sim 82$ and $N \sim 126$, and is clearly associated with the s-process. The other set is shifted to smaller $N \sim 76$ and $N \sim 116$, respectively, suggestive of a much more explosive neutron capture environment.

This second process is the r- or rapid-process [33], which requires a neutron fluence so large that neutron capture is fast compared to β-decay. In this case, nuclei rapidly absorb neutrons until they approach the neutron drip line, that is, equilibrium is maintained by $(\mathrm{n},\gamma) \leftrightarrow (\gamma,\mathrm{n})$, not by weak interactions. Consequently, the nuclei participating in

the r-process are very different from ordinary nuclei – very neutron-rich nuclei that would decay immediately in the low-temperature environment of Earth. The rate of nucleosynthesis is controlled by the rate of β-decay: a new neutron can be captured only after β-decay, $n \to p + e^- + \bar{\nu}_e$, opens up a hole in the neutron Fermi sea. Consequently, one expects abundance peaks near the closed neutron shells at $N \sim 82$ and $N \sim 126$, as β-decay is slow and mass will pile up at these "waiting points." By a series of rapid neutron captures and slower β-decays, synthesis can proceed all the way to the transuranics. Typical r-process conditions include neutron densities $\rho(n) \sim 10^{18} - 10^{22}$ cm^{-3}, temperatures $\sim 10^9$K, times ~ 1 s, and ratios of free neutrons to heavy seed nuclei of $\gtrsim 100$ (so that the transuranics can be synthesized from Fe).

The path of the r-process is typically displaced by just $\sim(2-3)$ MeV from the neutron drip line (where no more bound neutron levels exist). After the r-process neutron exposure ends, the nuclei decay back to the valley of stability by β-decay. This involves conversion of neutrons into protons, shifting the r-process peaks from the parent-nucleus values of $N \sim 82$ and $N \sim 126$ to lower values and explaining the double-peak structure of the r-process/s-process closed-shell abundance peaks.

One possible neutrino role in the r-process is in producing the required explosive, neutron-rich environments. One such site is the supernova neutrino-driven wind – the last ejecta blown off the proto-neutron star. This material is hot, dominated by radiation, and contains neutrons and protons, often with an excess of neutrons. As the nucleon gas expands off the star and cools, it goes through a freezeout to α-particles, a step that essentially locks up all the protons. Then the α-particles interact through reactions like

$$\alpha + \alpha + \alpha \to {}^{12}\mathrm{C}$$

to start forming heavier nuclei. The α-capture continues, eventually synthesizing intermediate-mass "seed" nuclei. Once these seed nuclei are produced, if the requisite number of neutrons is available ($\sim$100 per seed nucleus), very heavy nuclei can be made. The scenario, as depicted in Figure 11.11, is quite similar to the Big Bang: a hot nucleon gas (with an entropy $S \sim 100$ in Boltzmann units, compared to the BBN $S \sim 10^{10}$) expand and cools, condensing into nuclei. But a detail – the neutrino wind has an excess of neutrons, while the Big Bang is proton-rich – leads to uranium in one case, and to termination of nucleosynthesis at ^{4}He (plus a few light elements) in the other. The neutrinos are crucial: they help keep the entropy of the nucleon gas high, control the n/p ratio of the gas through competing charge-current reactions, and generate the wind that ejects the r-process products.

There are some very nice aspects of this site: the amount of matter ejected is $\sim 10^{-6}$ solar masses, a production per event that if integrated over the lifetime of the galaxy gives the required total abundance of r-process metals, assuming typical supernova rates. There are also some significant problems – including great difficulties in maintaining the necessary neutron/seed ratio of $\gtrsim 100$ for realistic wind entropies [34] – that have encouraged investigations of other sites. Possibilities include a so-called cold neutrino-driven r-process operating in the ^{4}He mantles of early, metal-poor supernova progenitors [35] and neutron star mergers, which could be the dominant r-process site once our galaxy has evolved to the point that these events become common [36].

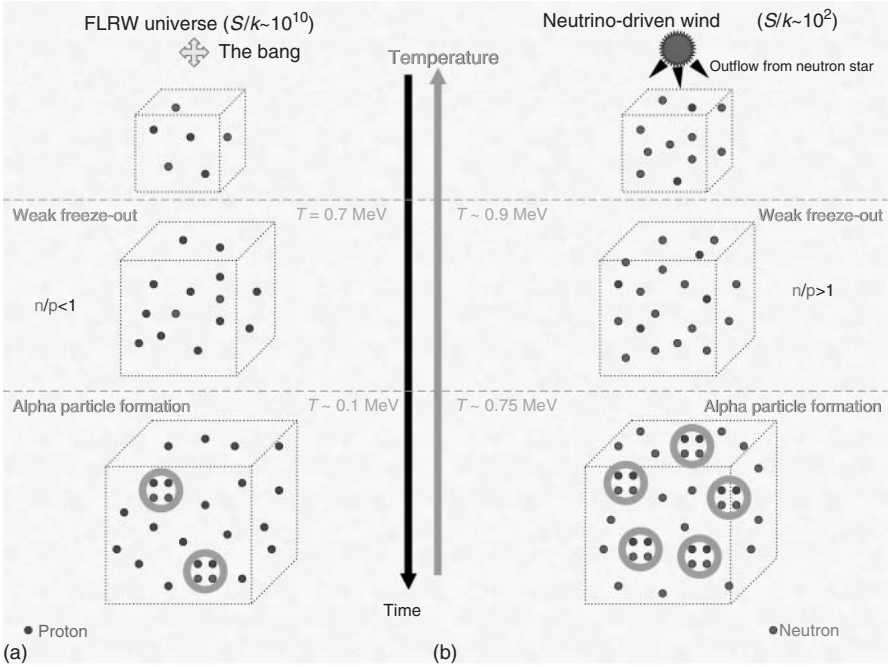


Figure 11.11 A comparison of two remarkably similar mechanisms for nucleosynthesis, (a) an expanding, cooling, high-entropy, proton-rich nucleon gas (the Big Bang) and (b) an expanding, cooling, high entropy, neutron-rich nucleon gas (the supernova wind r-process). In the former, the synthesis determinates at ^{4}He and a few other light elements. In the latter, the synthesis continues to the heavy transuranic elements. (Source: Figure from G. Fuller.)

11.6 Neutrino Cooling and Red Giants

Several neutrino cooling scenarios have already been discussed, including cooling of the proto-neutron star produce in core collapse and cooling connected with the expansion of the early universe. Red giant cooling provides an additional example of the use of astrophysical arguments to constrain fundamental properties of neutrinos.

11.6.1 Red Giants and Helium Ignition

In a solar-like star, when the hydrogen in the central core has been exhausted, an interesting evolution ensues:

- With no further means of producing energy, the core slowly contracts, thereby increasing in temperature as gravity does work on the core.
- Matter outside the core is still hydrogen rich and can generate energy through hydrogen burning. Thus a hydrogen-burning shell forms, generating the gas pressure supporting the outside layers of the star. As the ^{4}He-rich core contracts, the matter outside the core is also pulled deeper into the gravitational potential. Furthermore, the H-burning shell continually adds more mass to the core. This means the burning in the shell must intensify to generate the additional gas pressure to fight gravity. The shell

also thickens, as more hydrogen is above the burning temperature.

- The resulting increasing gas pressure causes the outer envelope of the star to expand by a large factor of up to 50. The increase in radius more than compensates for the increased internal energy generation, so that a cooler surface results. The star reddens. Stars of this type are called *red supergiants*.
- This evolution is relatively rapid, perhaps a few hundred million years: the dense core requires large energy production. The helium core is supported by its degeneracy pressure and is characterized by densities $\sim 10^6\ \mathrm{g\,cm^{-3}}$. This stage ends when the core reaches densities and temperatures that allow helium burning through the reaction

$$\alpha + \alpha + \alpha \rightarrow {}^{12}\mathrm{C} + \gamma$$

  As this reaction is quite temperature dependent, the conditions for ignition are very sharply defined. This has the consequence that the core mass at the helium flash point is well determined.
- The onset of helium burning produces a new source of support for the core. The energy released elevates the temperature and the core expands: He burning, not electron degeneracy, now supports the core. The burning shell and envelope move outward, higher in the gravitational potential. Thus shell hydrogen burning slows (the shell cools) because less gas pressure is needed to satisfy hydrostatic equilibrium. All of this means the evolution of the star has now slowed: the red giant moves along the ''horizontal branch,'' as interior temperatures increase slowly, much as in the main sequence.

The 3α process involves some fascinating nuclear physics that will not be recounted here: the existence of certain nuclear resonances was predicted based on the astrophysical requirements for this process. The resulting He-burning rate exhibits a sharp temperature dependence $\sim T^{40}$ in the range relevant to red giant cores. This dependence is the reason the He flash is delicately dependent on conditions in the core.

11.6.2 Neutrino Magnetic Moments and Helium Ignition

Before the helium flash, the degenerate He core radiates energy largely by neutrino pair emission. The process is the decay of a plasmon – which one can think of as a photon ''dressed'' by electron – hole excitations, thereby given the photon an effective mass of $\sim$ 10keV. The plasmon couples to an electron particle – hole pair that then decays via $Z_o \rightarrow \nu\overline{\nu}$.

If this cooling is somehow enhanced, the degenerate helium core would not ignite at the normal time, but instead continue to grow. When the core does finally ignite, the larger core will alter the star's subsequent evolution.

One possible mechanism for enhanced cooling is a neutrino magnetic moment. Then the plasmon could directly couple to a neutrino pair. The strength of this coupling would depend on the size of the magnetic moment.

A delay in the time of He ignition has several observable consequences, including changing the ratio of red giant to horizontal branch stars. Thus, using the standard theory of red giant evolution, investigators have attempted to determine what size of magnetic moment would produce unacceptable changes in the astronomy. The

resulting limit [37] on diagonal or transition neutrino magnetic moments,

$$\mu_{ij} \lesssim 3 \times 10^{-12} \text{ electron Bohr magnetons}$$

is about an order of magnitude more stringent than the best limits so far obtained from reactor neutrino experiments [38].

This example is just one of a number of such constraints that can be extracted from similar stellar cooling arguments. The arguments above, for example, can be repeated for neutrino electric dipole moments, or for axion emission from red giants. As noted previously, the arguments can be extended to supernovae: anomalous cooling processes that shorten the cooling time in a way that is inconsistent with SN1987A observations are ruled out. For example, large Dirac neutrino masses are in conflict with SN1987A observations: the mass term would allow neutrinos to scatter into sterile right-handed states, which would then immediately escape, carrying off energy.

11.7 High-Energy Astrophysical Neutrinos

Previous discussions focused on astrophysical neutrino sources with energies ranging from the cosmic microwave temperature to $\lesssim 10$ GeV, a range including the bulk of atmospheric neutrinos. These sources are displayed in Figure 11.12 according to their contributions to the terrestrial flux density.

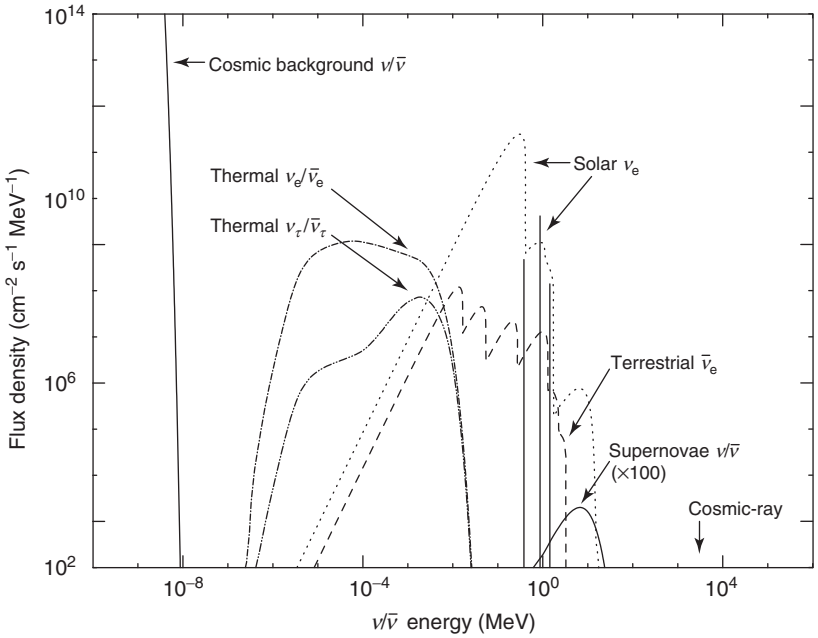


Figure 11.12 The low-energy neutrino "sky." In addition to the principal sources discussed in the text (Big Bang, solar, supernova, and atmospheric neutrinos), the natural neutrino background includes ~ 1keV thermal neutrinos of all flavors emitted by the Sun as well as $\bar{\nu}_e$s generated by the Earth's natural radioactivity. (Source: From [39].)

The figure includes low-energy sources, such as the thermal solar neutrinos of all flavors, not explicitly discussed here because of space limitations. Beyond the figure's high-energy limits, there exist neutrino sources associated with some of nature's most spectacular natural accelerators. The program of experiments to map out the high-energy neutrino spectrum is just beginning. Some guidance is provided by existing data on cosmic-ray protons, nuclei, and γ-rays, which constrain possible neutrino fluxes (Figure 11.13). This concluding section discusses some of the suggested sources and current efforts to develop high-energy neutrino telescopes appropriate to these sources. The high-energy spectrum is one of the frontiers of neutrino astronomy.

11.7.1 Neutrino Production by Cosmic-Rays

The UHECR spectrum – presumably protons and nuclei – is known to vary smoothly up to an energy $E \sim 4 \times 10^{19}$eV. The spectrum just below this point is characterized by a spectral index $\alpha \sim -2.7$: $\phi_\nu(E) \sim E^\alpha$. Higher-energy events are seen, but the flux drops off steeply beyond this point. This is consistent with the prediction of Greisen, Zatsepin, and Kuz'min [41]: Above this cutoff, UHECRs can lose energy by scattering off microwave photons, producing pions. This sharply reduces the mean free path of such UHECRs. The declining flux reflects the reduced number of sources within a mean free path of the Earth. This behavior has been mapped by

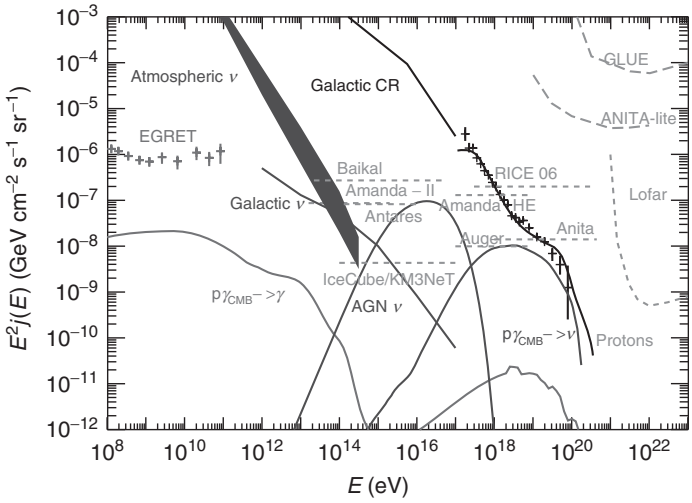


Figure 11.13 A theoretical model of high-energy neutrino "sky." Shown are the expected neutrino fluxes (solid black area and labeled gray lines), the primary galactic cosmic-ray fluxes (determined from data and a model, black), and the secondary γ-ray fluxes expected from protons interacting with the microwave background (labeled gray line). The neutrino flux is per flavor and includes only relatively certain sources: atmospheric neutrinos, neutrinos resulting from cosmic-ray interactions in our galaxy, and Greisen – Zatsepin – Kuzmin (GZK) neutrinos resulting from cosmic-ray interactions with the microwave and infrared backgrounds. The figure includes experimental data, limits, and projected sensitivities to existing and planned telescopes. (Source: Figure from G. Sigl [40].)

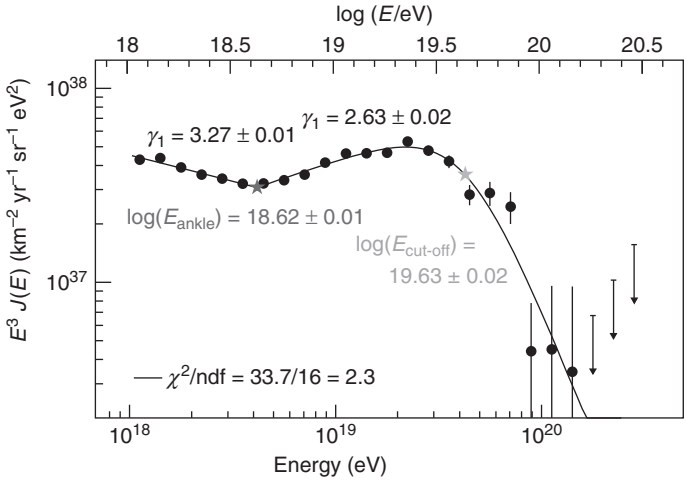


Figure 11.14 Pierre Auger Observatory cosmic-ray energy spectrum, fitted with two power laws in the lower-energy "ankle" region and with a smooth curve in the higher-energy GZK region, where a cutoff energy is identified. In addition to the statistical uncertainties shown, there is a systematic uncertainty in the energy scale of 22%. (Source: From [42], with permission of the Pierre Auger Observatory.)

the Pierre Auger Cosmic-Ray Observatory. The data shown in Figure 11.14 can be reproduced with a smooth fitting function having a cutoff in the spectrum above $\sim (4.3 \pm 0.2) \times 10^{19}$eV [42]. Above the cutoff, the primary signature of UHECR protons and nuclei would be the neutrinos they produce when interacting with the CMB.

From models of the cosmic-ray spectrum of protons and nuclei, an estimate can be made of the flux of ultrahigh-energy (UHE) neutrinos associated with the decays of pions and other secondaries produced in the GZK scattering. Uncertainties in this estimate include the flux, spectrum, and composition of the UHECRs, the behavior of the spectrum beyond the GZK cutoff (as we are blind to these UHECRs), and the spectrum's cosmological evolution. One bound was obtained by Bahcall and Waxman [43]; another is shown in Figure 11.13.

These uncertainties are connected to some very interesting astrophysical issues: the maximum energies that can be reached in astrophysical accelerators; the UHECR uniformity over time (or equivalently redshift); and the role of other background photon sources, such as the infrared and optical backgrounds, in producing UHE neutrinos from UHECRs. A more extended discussion can be found in [22].

As in the case of low-energy sources such as solar neutrinos, the detection of very high energy astrophysical neutrinos would open up new opportunities in both astrophysics and particle physics. Because the GZK cutoff limits the range of the UHECRs, neutrinos provide the only direct probe of nature's most energetic astrophysical accelerators. The cutoff mechanism itself – very high energy cosmic-ray interactions with CMB photons that can photoproduce pions and other secondaries–is a source of very energetic

neutrinos. Because neutrinos travel in straight lines through magnetic fields, they point back to their sources, allowing astronomers to correlate those sources with their optical counterparts – the accretion disks surrounding supermassive black holes, quasars, γ-ray bursts, and so on. The interactions of such energetic neutrinos with matter are untested experimentally, as terrestrial accelerators have reach only the TeV scale.

11.7.2 Cosmic-Ray Studies, Point Sources, and Neutrino Telescopes

The possibility of point sources is generally considered the astrophysical "driver" for developing instruments to measure the highest energy neutrinos. There are intensely energetic sources in the sky, including active galactic nuclei (AGNs), supernovae and associated phenomena such as γ-ray bursts, and compact objects such as black holes and neutron stars. The magnetic fields, shock waves, gravitational fields, and energy densities associated with such objects are beyond those that can be produced in the laboratory.

The measurements that have been made by the Pierre Auger Observatory, which began operations in 2004 and reached its full scope in 2008, provide an important baseline for high-energy neutrino studies. Its observation of a high-energy cutoff consistent with GZK predictions is one important step. The high-energy behavior of the flux is consistent with either a primary cosmic-ray spectrum dominated by protons or one dominated by heavier nuclei, so that additional information on the character of air-shower observables is needed to constrain composition (which influences GZK neutrino production). Changes in air-shower characteristics above $E \gtrsim 5 \times 10^{18}$eV indicate an increasingly heavier composition. The UHECR production of neutrinos can be significantly lower if the UHECR primary spectrum is dominated by CNO or Fe nuclei [44].

A second issue is the existence of point sources that might be probed in neutrinos. At very high energies, the trajectories of protons and nuclei are not strongly perturbed by magnetic fields, so that Pierre Auger results can inform high-energy neutrino studies about potential points sources. The observations and their interpretation in terms of the GZK effect suggest that the closest sources of UHECRs are within the GZK volume of $\lesssim 100$ Mpc, where the distribution of matter is inhomogeneous, and thus where there should be anisotropies in the UHECR flux [42]. Early results from the Pierre Auger Collaboration [46] showed directional correlations between UHECRs and the positions of cataloged AGNs on an angular scale of 3.1^{o} at 99%, c.l. The strength of this correlation has diminished somewhat as the Pierre Auger dataset has been enlarged, but the probability that the current data are uncorrelated with cataloged AGN positions remains below 1% [42]

The challenge in the field of UHE neutrinos is to build telescopes with the necessary sensitivity to see events, given current estimates of the fluxes (Figure 11.13). This requires instrumenting very large volumes. There have been ongoing efforts to use large quantities of water and ice as detectors, with experiments completed, operating, or in development using Antarctic ice (AMANDA, IceCube, ANITA, RICE, ARA, ARIANNA), the oceans (ANTARES, NEMO, NESTOR, KM3Net), and lakes (Baikal), and with detection methods

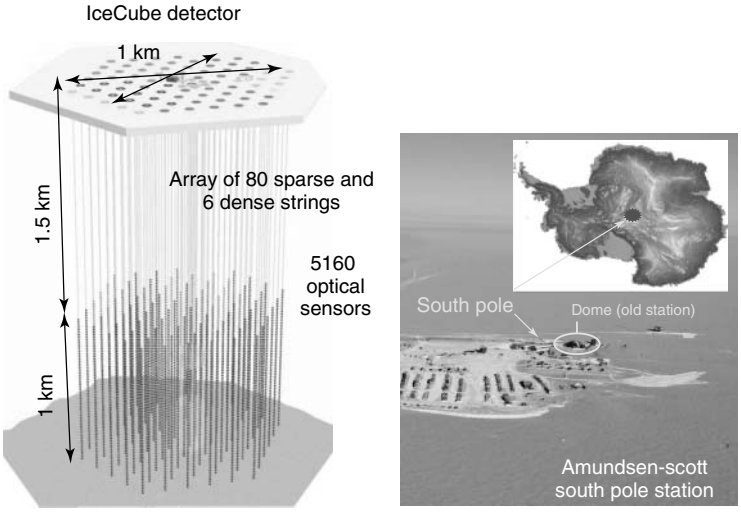


Figure 11.15 Current configuration of the IceCube high-energy neutrino detector. (Figure courtesy of the IceCube Project [45].)

including optical and coherent radio detection as well as particle production.

IceCube [45, 47], a project that extended the dimensions of high-energy neutrino detectors to a cubic kilometer, has been in operation at the South Pole for the past 5 years (Figure 11.15). This telescope views the ice through approximately 5160 optical sensors, deployed on an array of 80 sparse and 6 dense vertical strings, at a depth of 1450–2450 m. The detector is also coupled to a surface air-shower array. As the Earth is opaque to UHE neutrinos, detection must come from neutrinos incident at or above the horizon. IceCube deployment began in 2006, and the full set of 86 strings was completed in 2011: results from 354 days of 86-string operations were reported at Neutrino 2012 [45]. The number of high-energy neutrino events detected in IceCube is currently 28, with the most energetic of these exceeding 1 PeV (10^{15} eV). They appear to be of cosmogenic origin (e.g., perhaps resulting from the GZK mechanism or from on-site production from a cosmic accelerator. The experimentalists have concluded that these events are highly unlikely to be high-energy atmospheric neutrinos. IceCube sensitivity has reached the upper range of the high-energy neutrino fluxes predicted in cosmogenic neutrino models.

Acknowledgments

This work was supported in part by the US Department of Energy under contracts DE-SC00046548 (UC Berkeley) and DE-AC02-98CH10886 (LBNL) and by the Alexander von Humboldt Foundation.

Glossary

Atmospheric neutrinos: Neutrinos produced when cosmic rays strike the Earth's atmosphere.

Big Bang: The cosmological model of the origin of the universe in which the current universe resulted from the expansion of a primordial state characterized by very high temperatures and particle densities.

Borexino: A currently operating solar neutrino detector that employs a scintillator to detector low-energy solar neutrinos, such as the ^{7}Be and pep neutrinos. Borexino is located in Italy's Gran Sasso Laboratory.

CNO cycle: A series of thermonuclear reactions in which pre-existing C, N, and O act as ''catalysts'' for the conversion of four protons to helium, with release of energy. The CNO cycle is the dominant mechanism for ''proton burning'' in rapidly evolving, high-mass stars.

Chlorine detector: A solar neutrino detector that operated in the South Dakota Homestake Mine for three decades. This radiochemical detector contained about 615 tons of chlorine-bearing cleaning fluid. The discrepancy between the rate of neutrino reactions in this detector and the predictions of the SSM came to be known as the *solar neutrino problem*.

Core-collapse or Type II supernova: A type of supernova in which a massive star collapses under its own gravity to form a neutron star or black hole, with ejection of the star's mantle, producing a spectacular visual display. Core-collapse supernovae produce enormous bursts of neutrinos of all flavors.

Cosmic microwave background: The sea of low-energy photons that fills space, left over from the Big Bang. The spectrum is that of a blackbody characterized by a temperature of 2.725 K.

Cosmic rays: Energetic particles that are produced in various astrophysical sources, propagate through the interstellar medium, and can interact in the Earth's atmosphere. Sources include the winds that blow off stars and the ejecta from supernovae. About 90% of cosmic rays are protons, 9% are helium or other nuclei, and 1% are electrons.

Dark matter: The term in cosmology given to the largely unidentified matter that influences the large-scale structure of the universe through its gravitational interactions but so far has not been detected by more direct means. The majority of matter in our universe is dark. While neutrinos produced in the Big Bang account for a portion of the dark matter, the bulk of this matter appears to be something new, not currently described in the standard model of particle physics.

Dirac neutrino: A neutrino that has a distinct antiparticle, and thus has a total of four components. Lepton number is the ''charge'' that is used to distinguish the neutrino from the antineutrino.

IceCube Neutrino Observatory: A neutrino telescope under construction in Antarctica in which approximately a cubic kilometer of deep ice will be instrumented as an observatory for high-energy astrophysical neutrinos.

Majorana neutrino: A neutrino that serves as its own antiparticle, and thus has only two components.

MSW mechanism: The mechanism by which surrounding matter alters neutrino oscillation probabilities. In particular, the MSW mechanism can lead to large matter oscillation probabilities even when the oscillation probabilities in vacuum are small. The mechanism is named after the investigators who first discovered the phenomenon (S. Mikheyev and A. Smirnov) and first described matter's influence on neutrino mass (L. Wolfenstein).

Neutrino: A chargeless elementary particle that interacts only through the weak interaction, carries a spin of one-half, and has a very small mass. Neutrinos are produced in a variety of weak interactions, including nuclear β-decay. Neutrinos come in three types or flavors: electron, muon, and tauon.

Neutrino (flavor) oscillations: The quantum mechanical phenomenon in which a massive neutrino produced in one flavor state is later found, when detected some distance from the neutrino source, to be in a different flavor state.

Neutrino process: Nucleosynthesis by direct interactions of neutrinos in the mantle or wind of a core-collapse supernova.

Neutron star: The very dense compact object produced from the gravitational collapse of a massive star, made up primarily of neutrons and supported by the nuclear equation of state (in contrast to electron – gas interactions that support ordinary stars of ordinary density).

Nucleosynthesis: The production of new nuclei in cosmology and astrophysics. Nucleosynthesis sites include the Big Bang, the cores of ordinary stars, the interstellar medium through which high-energy cosmic rays pass, and explosive astrophysical environments such as supernovae.

Pierre Auger Cosmic Ray Observatory: A new international observatory for the study of UHECR interactions in the Earth's atmosphere. The observatory is located in Argentina's Mendoza Province.

pp chain: A set of thermonuclear reactions that, in the cores of stars like our sun, is responsible for the conversion of four protons to helium, with release of energy. The pp chain is the dominant mechanism for such ''proton burning'' in slowly evolving, low-mass stars.

Red giant: A later stage of evolution for low- or intermediate-mass stars characterized by large radii, low surface temperatures, and high luminosities. One class of red giants has a helium core supported by electron degeneracy pressure, surrounded by a shell where protons are burned to helium, adding to the core mass.

r-process: The process by which approximately half of the nuclei heavier than iron are synthesized, including all of the transuranics. The r- or rapid-neutron-capture process requires enormous neutron fluences and high temperatures, such as those produced in supernovae, neutron star mergers, or other explosive astrophysical environments.

Solar neutrinos: Neutrinos produced in the core of the Sun as a by-product of the thermonuclear reactions of the pp chain and CNO cycle.

Sudbury Neutrino Observatory (SNO): A detector that operated deep underground in a nickel mine located in Sudbury, Canada. The inner portion of this detector contained 1 kton of heavy water, a target that generates signals sensitive to neutrino flavor.

Super-Kamiokande: The massive (50,000 ton) water detector that operates underground in Japan's Kamioka Mine, recording solar and atmospheric neutrino interactions.

Standard solar model (SSM): The theoretical model of the evolution of the Sun, based on the standard theory of the structure of hydrogen-burning stars. Our detailed knowledge of the Sun – its age, mass, composition, vibrational modes, and so on – make it an interesting testing ground for stellar evolution theory.

Ultrahigh-energy cosmic rays (UHECRs): Cosmic rays with energies of, typically,

10^{15}–10^{20}eV, a range associated with very large cosmic accelerators and bounded, on the upper end, by an energy cutoff associated with cosmic-ray interactions with the cosmic microwave background.

References

1. Haxton, W.C. and Holstein, B.R. (2000) *Am. J. Phys.*, **68**, 15–; (2004) **72**, 18.
2. Bahcall, J.N. (1989) *Neutrino Astrophysics*, Cambridge University, Cambridge.
3. Davis, R. Jr., Harmer, D.S., and Hoffman, K.C. (1966) *Phys. Rev. Lett.*, **20**, 1205.
4. Serenelli, A.M., Haxton, W.C., and Peña-Garay, C. (2011) *Ap. J.*, **743**, 743.
5. Haxton, W.C., Robertson, R.G.H., and Serenelli, A.M. (2013) *Ann. Rev. Astron. Astrophys.*, to be published in.
6. (a) Super-Kamiokande Collaboration, Hosaka, J. *et al.* (2006) *Phys. Rev.*, **D73**, 112001; (b) Super-Kamiokande Collaboration, Cravens, J.P. *et al.* (2008) *Phys. Rev.*, **D78**, 032002; (c) Super-Kamiokande Collaboration, Abe, K. *et al.* (2011) *Phys. Rev.*, **D83**, 052010.
7. Smy, M. (2012) to appear in the Proceedings Neutrino 2012.
8. (a) SNO Collaboration, Ahmad, Q.R. *et al.* (2002) *Phys. Rev. Lett.*, **89**, 011301; (b) SNO Collaboration, Aharmim, B. *et al.* (2005) *Phys. Rev.*, **D72**, 052010–and (2010) **C81**, 055504.
9. Borexino Collaboration, Bellini, G. *et al.* (2011) *Phys. Rev. Lett.*, **107**, 141302–and (2012) **108**, 051302; (2010) *Phys. Rev.*, **D82**, 033006; (2012) *Phys. Lett.*, **B707**, 22.
10. Mikheyev, S.P. and Smirnov, A. (1985) *Sov. J. Nucl. Phys.*, **42**, 913.
11. Wolfenstein, L. (1979) *Phys. Rev.*, **D17**, 2369.
12. Maki, Z., Nakagawa, M., and Sakata, S. (1962) *Prog. Theor. Phys.*, **28**, 870.
13. Pontecorvo, B. (1967) *Zh. Eksp. Teor. Fiz.*, **53**, 1717.
14. KamLAND Collaboration, Gando, A. *et al.* (2011) *Phys. Rev.*, **D83**, 052002; KamLAND Collaboration, Abe, S. *et al.* (2011) *Phys. Rev.*, **C84**, 035804.
15. Fogli, G.L. *et al.* (2012) *Phys. Rev.*, **D86**, 013012.
16. Forero, D.V., Tórtola, M., and Valle, J.W.F. (2012) *Phys. Rev.*, **D86**, 073012.
17. Haines, T.J. *et al.* (1986) *Phys. Rev. Lett.*, **20**, 1986; Casper, D. *et al.* (1991) *Phys. Rev. Lett.*, **66**, 2561.
18. Hirata, K.S. *et al.* (1988) *Phys. Lett.*, **B205**, 416.
19. Fukuda, Y. *et al.* (1998) *Phys. Rev. Lett.*, **81**, 1562.
20. Itow, Y. (2012) talk presented at Neutrino 2012.
21. MINOS Collaboration (2012) talked presented by R. Nichol, Neutrino 2012.
22. APS Multi-Divisional Study of the Physics of Neutrinos, *http://www.aps.org/policy/reports/multidivisional/neutrino/*.
23. (a) Janka, H.-T., Langanke, K., Marek, A., Martinez-Pinedo, G., and Muller, B. (2007) *Phys. Rep.*, **442**, 38; (b) Mezzacappa, A. (2005) *Ann. Rev. Nucl. Part. Sci.*, **55**, 467; (c) Kotake, K., Sato, K., and Takahashi, K. (2006) *Rep. Prog. Phys.*, **69**, 971; (d) Woosley, S. and Bloom, J.S. (2006) *Annu. Rev. Astron. Astrophys.*, **44**, 507.
24. (a) Hirata, K. *et al.* (1987) *Phys. Rev. Lett.*, **58**, 1490; (b) Bionta, R.M. *et al.* (1987) *Phys. Rev. Lett.*, **58**, 1494.
25. Fuller, G.M., Mayle, R.W., Wilson, J.R., and Schramm, D.N. (1987) *Ap. J.*, **322**, 795; (b) Notzold, D. and Raffelt, G. (1988) *Nucl. Phys.*, **B307**, 924.
26. Steigman, G. (2007) *Ann. Rev. Nucl. Part. Sci.*, **57**, 463.
27. WMAP Collaboration, Komatsu, E. *et al.* (2011) *Astrophys. J. Suppl.*, **192**, 18.
28. (a) Izotov, Y.I. and Thuan, T.X. (2010) *Astrophys. J.*, **710**, L67; (b) Aver, E., Olive, K.A., and Skillman, E.D. (2011) *J. Cosmol. Astropart. Phys.*, **03**, 043.
29. Fields, B.D. (2011) *Annu. Rev. Nucl. Part. Sci.*, **61**, 47.
30. Pospelov, M. and Pradler, J. (2010) *Annu. Rev. Nucl. Part. Sci.*, **60**, 539.
31. Abazajian, K.N. *et al.* (2011) *Astropart. Phys.*, **35**, 177.
32. Woosley, S.E., Hartmann, D.H., Hoffman, R.D., and Haxton, W.C. (1990) *Ap. J.*, **356**, 272.
33. Qian, Y.-Z. (2003) *Prog. Part. Nucl. Phys.*, **50**, 153.
34. Roberts, L.F., Woosley, S.E., and Hoffman, R.D. (2010) *Ap. J.*, **722**, 1.
35. Banerjee, P., Haxton, W.C., and Qian, Y.-Z. (2011) *Phys. Rev. Lett.*, **106**, 201104.

36. Argast, D., Samland, M., Thielemann, F.-K., and Qian, Y.-Z. (2004) *Astron. Astrophys.*, **416**, 997.

37. Raffelt, G.G. (1990) *Phys. Rev. Lett.*, **64**, 2856.

38. Beda, A.G. *et al.* (2012) *Ad. High Energy Phys.*, **2012**, 350150.

39. Haxton, W.C. and Lin, W. (2000) *Phys. Lett.*, **B486**, 263.

40. Sigl, G. (2007) *Nucl. Phys. Proc. Suppl.*, **168**, 219.

41. (a) Greisen, K. (1966) *Phys. Rev. Lett.*, **16**, 748; (b) Zatsepin, G.T. and Kuz'min, V.A. (1966) *JETP Lett.*, **4**, 78.

42. Pierre Auger Collaboration, Kampert, K.-H. (2011) to appear in the Proceedings 3rd International Cosmic Ray Conference (arXiv:1207.4823).

43. Waxman, E. and Bahcall, J.N. (1999) *Phys. Rev.*, **D59**, 023002.

44. Ave, M., Busca, N., Olinto, A.V., Watson, A.A., and Yamamoto, T. (2005) *Astropart. Phys.*, **23**, 19.

45. Ishihara, A. (2012) talk presented at Neutrino.

46. Pierre Auger Collaboration, Abraham, J. *et al.* (2007) *Science*, **318**, 938.

47. (a) Ahrens, J. *et al.* (2004) *Astropart. Phys.*, **20**, 507; (b) Halzen, F. (2006) *Eur. Phys. J.*, **C46**, 669; (c) Resconi, E. *et al.* arXiv:0807:3891.

12
Nuclear Astrophysics

Hendrik Schatz

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

12.1 Introduction

Nuclear astrophysics aims at understanding the nuclear reaction processes that occur naturally in the cosmos and shape much of the visible universe. These nuclear processes make the stars shine, they power stellar explosions, and they are responsible for the origin of the elements defining the relative proportions of elements available for life.

At present, nuclear astrophysics is an active field at the intersection of astrophysics and nuclear physics. In addition to the long-standing open questions, such as the origin of the elements from iron to uranium or the nature of neutron stars, astronomical observations and advances in astrophysical theories continue to raise important new questions and frequently change paradigms. Examples include the recent observational discovery that the heavy elements formerly attributed to the rapid neutron capture process (r-process) are actually produced in two distinct processes, a weak and a main r-process. While one clearly observes the signatures of these processes, the actual site is known for neither. One new hypothesis based on more sophisticated astrophysical models is that the weak r-process occurs in supernovae, surprisingly not as a neutron capture process, but as a new kind of rapid proton capture process (rp-process), accelerated by neutrino interactions. Similarly, the discovery of super X-ray bursts has revealed glaring gaps in our understanding of accreting neutron stars. New theoretical work has predicted the occurrence of stellar explosions that are weaker than type Ia supernovae, but stronger than novae, so-called .Ia supernovae and observations have indeed revealed candidates for such events. Furthermore, the discovery of a neutron star with a surprisingly high mass, about twice the mass of the Sun, has challenged our understanding of nuclear matter.

Answering the open questions in nuclear astrophysics requires information on a wide, with every new discovery, expanding range of stable and unstable nuclei. However, the nuclear data needs of astrophysical models are pushing the limits of nuclear physics. In most cases, nuclear physics has fallen short of providing the nuclear data needed. There are two notable exceptions that involve light nuclei, which can be more easily studied experimentally and calculated theoretically:

The nuclear reactions forming deuterium, helium, and traces of lithium in the Big Bang are now well understood, and the remaining nuclear uncertainties are well characterized. Early on, the understanding of the nuclear physics of the Big

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

Bang has pointed to the need for non-baryonic dark matter, as the synthesis of the correct amounts or primordial isotopic abundances requires a baryon density that does not close the universe. After cosmic microwave background observations have pinned down the baryon density precisely, the focus is now on the resolution of the factor of 3–5 discrepancy between predicted Big Bang production of ^{7}Li and the level of the plateau of ^{7}Li abundances observed in very metal poor, old stars, thought to reflect the Big Bang production level. The good understanding of the Big Bang nuclear reaction rates forces us now to question our understanding of mixing processes in stars that might deplete ^{7}Li on their surface, or, alternatively, consider exotic nonstandard Big Bang scenarios.

The other astrophysical scenario where nuclear reactions are now relatively well known, thanks to decades of experimental and theoretical efforts in nuclear physics, is the Sun. The good understanding of the nuclear reactions in the hydrogen-burning pp-chains enabled the discovery of the solar neutrino problem – a difference between the predicted and observed fluxes of neutrinos emitted by these reactions. This in turn led to the profound discovery of neutrino flavor oscillations and the fact that neutrinos differ in mass.

Clearly, the understanding of the nuclear physics in these scenarios has led to major progress in science, often in unexpected ways. The goal of the field is to achieve a similar level of nuclear physics understanding for the astrophysical scenarios that involve heavier nuclei beyond boron. There are three major challenges for nuclear physics:

1. Most nuclei participating in the cosmic nuclear reaction sequences are unstable and therefore radioactive. It has turned out to be a major technical challenge to produce sufficiently intense beams of radioactive nuclei to measure the nuclear properties needed for astrophysics. In fact, the majority of nuclei participating in stellar nuclear processes currently cannot be produced in the laboratory. The construction of a new generation of powerful radioactive beam facilities such as the radioisotope beam factory (RIBF) in Japan, facility for rare isotopes beams (FRIB) in the United States, facility for antiproton and ion research (FAIR) in Germany, and Spiral II in France will address this challenge.
2. The lack of a predictive theory for all nuclei severely hampers progress in nuclear astrophysics. Even with greatly expanded experimental capabilities, nuclear theory will be needed to predict properties of out-of-reach nuclei and the modifications of nuclear properties in extreme stellar environments. The most important aspect is the prediction of nuclear structure trends such as shell closures, changes in shape with increasing proton and neutron numbers, and the limits of stability across the chart of nuclides, from the proton to the neutron drip lines, and from hydrogen to superheavy elements. Such general trends can have profound consequences on astrophysical model calculations and uncertainties due to the lack of knowledge about the basic structure of unstable nuclei are difficult to quantify.

Indeed, recent discoveries in nuclear physics clearly show that nuclear structure can change in unexpected ways when adding protons and neutrons to stable nuclei. One example is that of the changes in shell structure that alter

the traditional magic numbers for light, unstable nuclei. Another example is the limit of stability when adding neutrons to a nucleus, the so-called neutron drip line. While ^{24}O is known to be the last bound oxygen isotope, adding just a single proton allows to bind a surprisingly large number of additional neutrons – at least six. Predicting such trends for all nuclei, with a well-characterized uncertainty, is of critical importance for nuclear astrophysics.

3. Reactions among charged nuclei in stars are extremely slow, a fact that manifests itself in the long lifetimes of stars. Measuring these rates in the laboratory has therefore been a major challenge, and, with a few exceptions, measurements had to be performed at higher energies where rates are faster. However, extrapolating measurements to low stellar energies has turned out to be unreliable owing to limitations in nuclear reaction theory. An important direction in nuclear astrophysics is therefore the development of experimental approaches that increase the signal-to-background ratio in stellar reaction rate measurements and allow one to perform the measurements at lower energies. An example is the establishment of advanced accelerator facilities deep underground.

12.2 Connecting Nuclear Physics and Astrophysics

Element synthesis in the cosmos is a complex cycle that continues to operate today. Stars form out of interstellar gas and dust, create new elements through nuclear reactions, and eject them into space through winds or supernova explosions. The ejecta are mixed into the interstellar medium where the process then starts anew. The increase in the amount of heavy nuclei in the interstellar gas, and therefore in stars, over time opens up new reaction sequences that generate more new elements. The presolar nebula out of which our solar system formed 4.5 billion years ago is thought to be enriched by hundreds of such cycles. Understanding the chemical evolution of our Galaxy is therefore intimately tied to the understanding of the origin of the solar system, the Earth, and life. Stars and stellar explosions are natural nuclear reactors. Nuclear reactions therefore determine their evolution and the elements that they form in the process. The following sections give an overview of how nuclear physics influences various astrophysical environments and what kind of data are needed to address open astrophysical questions.

12.2.1 Stars

Stars mark the beginning of the cycle of nucleosynthesis. They form out of interstellar gas that has been enriched with chemical elements by previous generations of stars. The initial composition depends on the location the star forms in, but consists of mostly hydrogen (∼70%) and helium (∼23%) with a small percentage of heavier elements, the so-called metallicity.

After a star forms, it eventually settles into a configuration where nuclear reactions burn hydrogen into helium in its core, replenishing the energy the star is losing through radiation and neutrino emission. As a consequence, the star maintains a temperature and pressure profile that ensures hydrostatic equilibrium. Once hydrogen is exhausted, the star burns helium into carbon and oxygen. Because

of the higher central temperature required for this process, the star settles into a different configuration and becomes a red giant. The further evolution depends strongly on the initial mass. Low- and intermediate-mass stars with initial masses below 8–10 solar masses shed their envelopes as a planetary nebula and leave an inert, hot carbon–oxygen core behind. These compact carbon–oxygen cores are stabilized against gravitational collapse by the degeneracy pressure of the highly compressed electrons in their interior, and are observed as white dwarf stars. More massive stars become hot and dense enough in their centers to initiate carbon and oxygen fusion reactions, followed by fusion reactions of their reaction products prolonging the life of the red giant to some extent. The last possible burning stage is silicon burning, in which, within a day, silicon is burned into iron and nickel isotopes. This last burning stage consists not of a simple fusion reaction sequence anymore. Instead, temperatures are so high that nuclei get partially dissociated into protons, neutrons, and helium particles, which get rapidly recaptured. This opens up many fast reaction channels leading to partial nuclear statistical equilibrium that favors iron and nickel nuclei because of their high binding energy per nucleon. Further fusion reaction stages are not possible and would not release any energy. Silicon burning therefore marks the end of the red giant phase of a massive star and a core-collapse supernova is initiated.

Knowledge of the rates of the nuclear reactions in stars is important. Stars are difficult to model and their structure is not well understood. A particular challenge is convective mixing processes requiring sophisticated 3D models that need to include magnetic fields and rotation. Such models do not exist yet, but promising developments are underway and specific aspects of stellar models have been investigated using multidimensional approaches [1]. However, the main challenge is model validation – how does one know that the complex processes deep in the stellar interior are modeled correctly? As convective processes directly affect the elements synthesized, comparison of the predicted nucleosynthesis with observations of the composition of stars opens a pathway to validation. The composition of stars can, for example, be determined through analysis of presolar grains, pieces of stardust that have been incorporated in primitive meteorites found on the Earth. However, this validation approach is only as good as the knowledge of the nuclear reaction rates, which therefore need to be known with high precision. Understanding of the nuclear reactions in stars is also important to enable computing of realistic initial models for supernova simulations.

The hydrogen-burning reaction rates in the pp-chains and carbon, nitrogen, and oxygen (CNO) cycles are comparably well studied, but more precise rates will help extract more physics out of observations of solar neutrinos. In particular, with ongoing and planned efforts to detect the lower-energy neutrinos from the CNO cycle, for example, with BOREXINO or SNO+, there is a need to improve the precision of the CNO cycle reaction rates. The goal is to use the CNO neutrino flux together with precise cross sections to determine the absolute amount of carbon, nitrogen, and oxygen in the Sun. This will address the problem of the composition of the Sun: the latest analysis of the composition of the solar surface using spectral lines disagrees with helioseismology data [2].

The helium-burning reactions $3\alpha \rightarrow {}^{12}C$ and ${}^{12}C + \alpha \rightarrow {}^{16}O$ (in nuclear reaction notation ${}^{12}C(\alpha,\gamma){}^{16}O$) have been identified

as one of the major uncertainties in stellar evolution. They determine the ^{12}C-to-^{16}O ratio at the end of helium burning, which sets the stage for the subsequent evolution. Variations of these rates within their current uncertainties (10–12% and 25–35%, respectively, [3]) have been shown to have a large impact on the resulting nucleosynthesis and the size of the iron core, triggering the core-collapse supernova in massive stars. Because of the strong impact of the $^{12}C(\alpha,\gamma)^{16}O$ reaction on a wide range of produced elements in massive stars, astrophysicists have adjusted the rate within its uncertainties to obtain agreement with observed relative abundances of elements produced in massive stars, effectively determining the rate through astrophysical arguments [3]. However, while intriguing, there are significant systematic uncertainties in current 1D stellar models. It is therefore essential that the rate is determined experimentally with sufficient accuracy so that current stellar models can be tested.

Other fusion reactions also suffer from large uncertainties because measurements at low energies are not possible owing to the exponential drop of the rate with energy. The problem is further exacerbated by the unknown general behavior of fusion rates at low energies where measurements are not possible. Experimental data obtained for fusion reactions of heavier nuclei with negative Q-values point to the possibility of an unexpected suppression of fusion at low energies compared to expectations. Whether the same effect occurs for the astrophysically relevant fusion reactions such as $^{12}C+^{12}C$ remains to be seen. The case of the $^{12}C+^{12}C$ fusion rate is further complicated by the fact that at higher energies, this rate shows large fluctuations with energy that resemble resonances. The stellar fusion rate therefore depends strongly on whether there happens to be a resonance in the critical stellar energy range or not.

12.2.2 Core-Collapse Supernovae

Core-collapse supernovae mark the end of the life of a massive star. The core collapse is initiated once the star runs out of fuel and has accumulated an iron core of roughly 1.4 solar masses. The collapse begins in an inner section of the iron core, the so-called inner core, within which for a finite sound speed, pressure changes can travel within the collapse timescale. The inner core collapses homologously, while the remaining outer core lags behind. Once the inner core reaches nuclear density, a jump in pressure occurs because of the repulsive nature of the nuclear force at small distances, resulting in a bounce and an outward directed shock that passes through the in-falling outer core. How this outward-moving shock gains enough energy to move beyond the iron core and revert the direction of the mass flow into an explosion is one of the major open questions in nuclear astrophysics. This problem is discussed in more detail in Chapter 15.

Nuclear physics plays a central role in core-collapse supernovae and is an important piece of the physics that needs to be understood to address the problem of the supernova explosion mechanism. The important nuclear physics is the nuclear equation of state at and beyond nuclear density – this determines time and strength of the bounce that launches the shock – and the rate of electron capture on nuclei. Electron capture on nuclei occurs in the early stages of the collapse. As electrons provide the pressure, a

reduction of electron abundance affects the dynamics of the collapse. In addition, neutrino interactions with matter are important. Neutrinos are the most likely way gravitational energy is transferred from the hot, collapsed inner core to the outward-moving shock.

12.2.3 The Origin of the Heavy Elements

Elements beyond the iron–nickel region cannot be produced by stellar fusion reactions. Overcoming the considerable Coulomb repulsion of two heavy nuclei would require temperatures that would lead to the immediate dissociation of the reaction product. Fusion reactions beyond iron also become endothermic as the binding energy per nucleon starts to decrease with increasing element number. Most of the heavy elements are therefore thought to be produced by neutron capture processes requiring free neutrons and seed nuclei. Captures of neutrons by the seed nucleus will eventually produce an unstable neutron-rich isotope, which will then β-decay into a heavier element. The process repeats and successively heavier and heavier elements can be produced. Even though the seed nuclei are converted into less-bound, heavier nuclei, the consumption of free neutrons more than compensates for this, making the process exothermic. The origin of about 60% of these heavier elements can be explained with the slow neutron capture process (s-process) that has been found to operate in red giant stars. But where do the other ∼40% come from? They are usually attributed to an r-process. We still do not know where this r-process takes place and what exactly the reaction sequence is that builds up the heavy elements.

12.2.3.1 s-Process

The s-process is characterized by neutron capture timescales of 10–100 years. Its existence can be inferred from the distinct abundance peaks in the Solar System abundance distribution of isotopes (Figure 12.1) at mass numbers $A = 138$ and $A = 208$. These peaks can be easily explained by a neutron capture process that

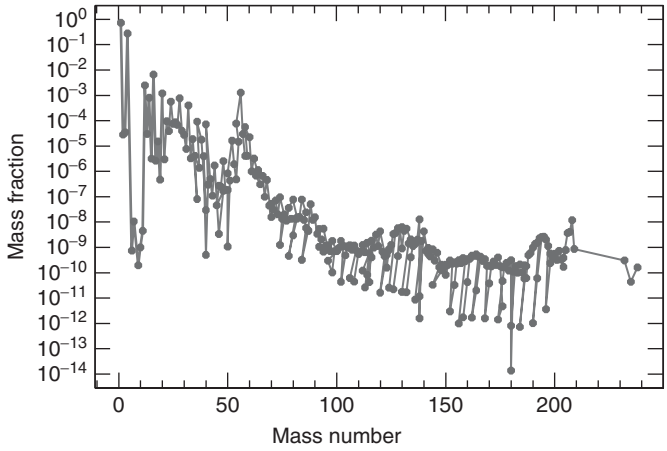


Figure 12.1 The composition of the Solar System in mass fraction as a function of the atomic number of the nuclei. (Source: Adapted from [4].)

proceeds along the stable isotopes (i.e., as soon as neutron captures reaches a typical unstable isotope, the isotope decays before it captures another neutron – hence s-process) and crosses the closed neutron shells at the magic neutron numbers $N = 82$ and $N = 126$. For these neutron numbers, the nucleus is particularly stable and neutron capture is suppressed. The decrease in destruction rate (but not in production rate) leads to a higher overall production of these isotopes. Stable isotopes at $N = 82$ and $N = 126$ are ^{138}Ba and ^{208}Pb, respectively, explaining nicely the occurrence of abundance peaks at these mass numbers. The site of the s-process has been unambiguously identified by the detection of technetium on the surface of red giant stars. The element technetium has no stable isotope and its presence therefore indicates the ongoing production in the interior of the star as well as the presence of a dredge-up process that transports the isotopes to the surface.

The s-process is not unique. s-process nucleosynthesis depends strongly on the initial mass and, as the s-process is a secondary process that requires the presence of seed nuclei, on the initial metallicity of the star. Most of the s-process elements above mass number 90 are thought to be produced in the main s-process in low-mass asymptotic giant branch (AGB) stars. These are red giant stars where unstable helium shell burning leads to periodic helium-burning flashes. The main neutron source is the ^{13}C(α,n) reaction in the so-called ^{13}C pocket – a thin layer of ^{13}C that is thought to be produced by mixing material from the star's hydrogen envelope into the helium-burning zone following a flash. Proton captures on the products of helium burning, ^{12}C, and the subsequent decay of ^{13}N produce the ^{13}C. As the hydrogen retracts, the ^{13}C is embedded in the helium layer and the ^{13}C(α,n) reaction produces neutrons for an s-process for about 50 000 years. Then the next helium flash occurs triggering again the mixing process. During the helium flashes, the ^{22}Ne(α,n) reaction produces neutrons for a short additional s-process episode that lasts a few hundred years. Stars with lower metallicities are thought to produce mostly the heaviest s-process elements, lead and bismuth, as they achieve a larger neutron-to-seed ratio and can therefore sustain more neutron captures. The light s-process elements between iron and $A = 90$ are thought to be produced in the weak s-process in the cores of massive red giant stars. In this case, the main neutron source is the ^{22}Ne(α,n) reaction rate, which is activated at the end of helium burning and during the beginning of carbon burning.

The s-process is a good example for what can be achieved once the nuclear physics of a nucleosynthesis process is well understood. Neutron capture rates on stable nuclei are the slow reactions in the s-process reaction sequence and therefore determine the produced abundance pattern for a given set of astrophysical conditions. Decades of careful precision measurements of astrophysical neutron capture rates enable detailed tests of s-process models by comparing them to the observed s-process abundance pattern in the Solar System. Of particular importance are the so-called s-process branch points. These are unstable isotopes along the s-process path that happen to have lifetimes comparable to the timescale of neutron capture. When the s-process reaches these isotopes, sometimes neutron capture occurs, and sometimes β-decay. The reaction path therefore splits. The relative frequency of each branch depends

on the stellar environment. For example, at higher neutron density, the neutron capture branch will be enhanced. In some cases, β-decay rates can be temperature sensitive when photons excite the nucleus, so it does not always decay from its ground state. In this case, the branchings will also depend on temperature. Other branch points occur when both, β^- and β^+ or electron capture decay modes are possible. Branchings can typically be determined from the observed relative abundances of the produced nuclei along each branch. Once the nuclear rates governing each branch are precisely known, one can then infer neutron density and temperature at the s-process site. A major breakthrough in the understanding of the s-process occurred when after many laboratory measurements of neutron capture rates, it became clear that a simple set of exponentially declining neutron flux exposures cannot explain all branch points simultaneously. The solution was the thermal-pulse AGB star model described above, with two neutron sources operating at very different conditions [5].

The open question concerning the main s-process today is how mixing processes in the stellar interior create the ^{13}C pocket, and how one can predict its size and dependence on metallicity. In this context, the discovery of presolar grains – pieces of SiC stardust that can be found in meteorites – is of importance. Some of these presolar grains can be shown to originate from AGB stars, and they carry abundance signatures of the s-process in one particular star. This information is invaluable, as the s-process yields vary from star to star and the Solar System abundance distribution only provides an integrated view of the output of many different s-process sources. Models of convective processes in AGB stars as a function of stellar mass, metallicity, and other parameters can be directly tested by comparing the resulting s-process yields to the range spanned by AGB star grain data. For this test to be meaningful, the neutron capture rates in the s-process need to be known with a precision of better than 5%. While this goal has been achieved for most nuclei in the main s-process, there are still many measurements needed for the weak s-process. An additional challenge is that of neutron capture rates on branch point isotopes that are too short-lived to be easily measured.

The main open question concerning the weak s-process is its end point and the actual pattern of abundances produced as a function of stellar mass and initial metallicity. The main uncertainty is the nuclear reaction rates of neutron-producing and neutron-destroying reactions. One of the most critical uncertainties is the rate of the ^{22}Ne(α,n) neutron-producing reaction.

12.2.3.2 r-Process

The r-process produces roughly 40% of the heavy elements above iron. The fact that we still do not know where and how this process operates is one of the biggest open questions in nuclear astrophysics. The existence of the r-process follows directly from the peaks at $A = 130$ and $A = 195$ in the solar abundance distribution (Figure 12.1) and the shell structure of unstable nuclei. Similarly to the s-process (Section 12.2.3.1), the peaks can easily be explained with a neutron capture process that crosses the neutron shells with 82 and 126 neutrons at mass numbers around 130 and 195. Just as in the s-process, the additional binding of nuclei with a closed neutron shell leads to a slowdown of the process and increased element production. Therefore, the r-process must pass through nuclei near ^{130}Cd (mass number 130 and

82 neutrons), a rare isotope with a half-life of just 160 ms, and through nuclei near ^{195}Tm, an extremely neutron-rich thulium isotope that has never been observed in a laboratory. Once these nuclei are produced in the r-process, they decay via chains of β^- decays. Some of these decays can be accompany by neutron emission, which can change the final mass number of the resulting stable nucleus at the end of the decay chain by at most a few mass units. With some reasonable assumptions about the physics of neutron-rich nuclei, such as the persistence of the neutron shells at neutron numbers 82 and 126 (which remain to be confirmed), one can therefore conclude that the r-process proceeds far from stability, requiring extreme neutron densities of the order of 10 kg cm^{-3} in order for neutron captures to be sufficiently fast to reach ^{130}Cd and ^{195}Tm. Identifying an astrophysical site with such conditions has been a major challenge for the field.

Advances in astronomy have led to breakthroughs in the understanding of the chemical evolution of the Galaxy and the r-process. The important observations concern the spectroscopy of very metal poor stars mostly found in the halo of the Galaxy. These stars have formed during the early stages of chemical evolution and have preserved the composition of the Galaxy at the time and location of their formation. This composition can now be determined from the spectral analysis of the starlight. The chemical age of the stellar material can be inferred from the iron contents, which tends to increase with ongoing chemical evolution. Therefore, these stars provide snapshots of the process of chemical evolution during an epoch where individual nucleosynthesis events step by step built up the chemical composition of the Galaxy – in essence providing a fossil record of chemical evolution. About a dozen stars have now been found that apparently have formed near r-process events revealing in detail the composition produced by a single r-process. In the most extensively studied star among these, CS22892-052, the abundances of 57 elements have been determined [6].

The major conclusions from these observations are that the r-process is a so-called primary process (a nucleosynthesis process that does not rely on previous nucleosynthesis events for seed production, but produces its own seed) that operates in the early galaxy. The very stable abundance pattern among the dozen or so ''r-process stars'' discovered so far, and the good agreement of the pattern with the r-process composition inferred from the Solar System by subtracting the s-process contribution at least for elements from Ba to Pb indicate that all r-process events produce exactly the same abundance pattern – a remarkable aspect that needs to be explained by a successful r-process model. Maybe most importantly, the observations also reveal the surprising fact that a second r-process contributes to the lighter r-process elements, especially Y, Sr, and Zr. Indeed stars that show the composition ejected in this new so-called weak r-process or light element primary process (LEPP) [7] have also been found [8].

Thanks to this major effort in astronomy, that required extensive use of the largest telescopes available, including the Hubble Space Telescope, Keck, SUBARU, and the VLT, as well as low- and medium-resolution surveys of millions of stars in campaigns such as SEGUE or the ongoing APOGEE, we now know well the characteristic abundance patterns produced by the two r-processes.

Major progress has also been achieved in the theoretical modeling of the r-process. Up to very recently, r-process models did not naturally produce a full r-process, but had to be adjusted to do so in order to be able to investigate the r-process theoretically. For example, in the neutrino-driven wind expelled off a forming neutron star in a supernova explosion, entropies had to be artificially increased far beyond reasonable limits to obtain an r-process. Recently, however, a number of models have been published that self-consistently lead to a robust r-process. Examples include supernovae with rapidly rotating stars that have high magnetic fields and expel matter in so-called jets [9]. Another example is that of neutron star mergers, where two neutron stars orbiting each other, move closer over time due to gravitational wave emission, and eventually merge into a black hole. During the merger, some neutron star matter is ejected, heated, and decompressed resulting in an r-process. A recent study demonstrated that the same r-process pattern is produced regardless of the initial mass distribution of the two neutron stars, even for neutron star and black hole mergers [10].

With reliable nuclear data, the new robust r-process models could now be tested by comparing their resulting abundance patterns to the new observations. The goal is to reveal deficiencies in astrophysical models that can then be addressed by improving models through additional physics or, if that is not possible, by discarding certain types of models. Indeed different r-process models are characterized by significantly different conditions, resulting in differences in the abundance patterns that could be used to identify the best model once the nuclear physics is known. Unfortunately, only a small fraction of the nuclear physics needed in r-process models is known, and most r-process nuclei are out of reach at current radioactive beam facilities. Together with the lack of a reliable predictive theory of the properties of heavy nuclei, this prevents testing of the various proposed r-process models. Here, next-generation radioactive beam facilities such as FAIR or FRIB will have a major impact.

Figure 12.2 shows the nuclear reaction sequence during a possible r-process. In a typical r-process model, temperatures are sufficiently large to make (γ,n) photodisintegration reactions (the emission of a neutron when the nucleus is excited by a γ-ray) as fast as (n,γ) neutron capture reactions. The result is an $(n,\gamma)-(\gamma,n)$ equilibrium that is established within an isotopic chain. The isotope favored in the equilibrium, the so-called waiting point, is, for a given neutron density and temperature, largely determined by nuclear masses (Section 12.3.4). At the waiting point, the r-process has to wait for the β-decay to occur. Once the β-decay has happened, the process repeats in the next isotopic chain. While nuclear masses determine the waiting point nuclei and therefore the r-process path, β-decay rates determine how quickly the process can proceed toward heavy elements. As the slowest reactions in the chain, β-decay rates also determine the abundance pattern produced for a given r-process path (Section 12.3.2). Neutron capture rates are unimportant as long as $(n,\gamma)-(\gamma,n)$ is maintained. They become important during the end of the r-process when temperatures and neutron densities drop, causing equilibrium to break down, and in cold r-process models. Neutron emission following β-decay is also important as it can lead to late time production of neutrons and a significant modification of the final abundance pattern. Finally, fission is of potential importance in the

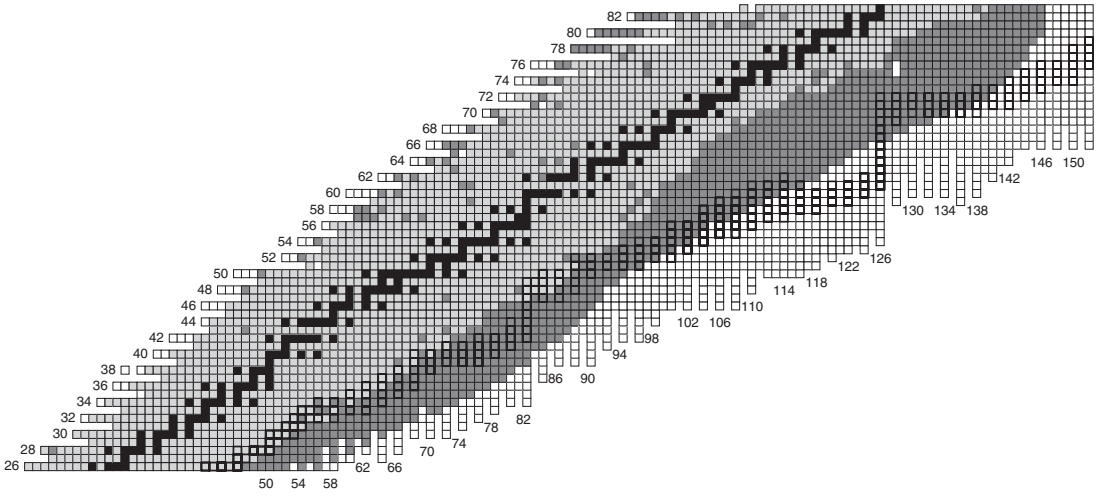


Figure 12.2 The thick squares mark the predicted r-process waiting points that contribute significantly to the final composition produced by the r-process [11]. Light gray nuclei have been studied in the laboratory and at least a half-life has been measured. Dark gray shading marks the reach of experiments at the FRIB facility in the United States.

r-process. In models that have particularly large neutron densities and low seed abundances, the r-process can reach the actinide region, where fission is thought to terminate the process. The resulting fission fragments can serve as seeds for a continued r-process, sometimes again reaching the fission region, repeating the process, the so-called fission cycling. Fission cycling has been proposed as a mechanism to create a very robust r-process abundance pattern regardless of the exact astrophysical conditions, possibly explaining the extraordinary event-to-event stability of the r-process inferred from observations. However, fission cycling models have difficulty in explaining the production of lighter r-process elements (lighter than the typical fission fragments, which have mass numbers around 130), which are observed to be coproduced in the main r-process.

This leaves the question of the mechanism of the ''weak r-process.'' A variety of models have been proposed. A particularly intriguing solution is the failed r-process in neutrino-driven wind models. When using realistic conditions and simply calculating the resulting nucleosynthesis instead of trying to adjust parameters to obtain an r-process, the resulting abundance pattern agrees reasonably well with the observed ''weak r-process'' [12]. As it turns out, the winds are actually proton rich, resulting in an rp-process as opposed to an r-process. The traditional rp-process (Section 12.2.5.1) requires long processing times as β-decays tend to be much slower than in the r-process. The reason is that the limit of particle stability is reached much more quickly when adding protons to a nucleus compared to adding neutrons. Half-life changes from nucleus to nucleus are therefore also larger, and chances are high that relatively long half-lives occur near the limit of proton stability where the process proceeds. However, in the case of the neutrino-driven wind, this limitation is overcome as neutrino interactions with protons produce a small level of neutrons via $p + \overline{\nu}_e \rightarrow n + e^+$ that allows the process to bridge slow β^+ rates with much faster (n,p) reactions resulting in the same final nucleus. This so-called νp-process can reach elements such as molybdenum or ruthenium within the short time available in a supernova explosion [13]. The critical nuclear physics of this recently discovered process is only now being identified but it is clear that proton capture rates and (n,p) rates on unstable neutron-deficient nuclei will be needed to predict the nucleosynthesis contribution of the νp-process.

12.2.4 p-Process

The p-process explains the origin of the 35 typically very rare neutron-deficient isotopes of elements between selenium and mercury that cannot be produced by neutron capture processes, the so-called p-isotopes. These isotopes can be produced in the outer layers of supernova explosions – either of the core-collapse type or of the thermonuclear type. Seed nuclei produced by an s-process before the explosion are heated and photodisintegrated by (γ,n), (γ,p), and (γ,α) reactions producing neutron-deficient isotopes. As the explosive heating travels through the star, peak temperatures decrease with distance. In a supernova, there is therefore a layer where it is hot enough for some photodisintegrations to produce neutron-deficient isotopes, but not hot enough to destroy all nuclei. The model explains the observed isotopic abundances of most p-isotopes reasonably well, with the exception of light p-nuclei such as ^{92,94}Mo and ^{96,98}Ru that are rather abundant in

nature compared to other p-isotopes. There are two possible solutions – either the seed for the p-process is enriched with nuclei in this mass region or a separate process such as the νp-process is responsible for the origin of these isotopes in nature. Reliable nuclear data are needed to calculate the exact nucleosynthesis contribution of the different proposed models and to possibly reveal deficiencies in models – either in terms of predicted conditions or in terms of the assumed seed composition.

12.2.5 Accreting Neutron Stars

Neutron stars are essentially large blobs of 1–2 solar masses of nuclear matter floating through space. Their radii are most likely in the range of 10–13 km. Neutron stars offer the unique opportunity to study bulk neutron-rich nuclear matter at low temperature and high densities. Neutron star studies can therefore address fundamental questions about matter, such as the maximum possible density, or the energy–pressure relationship. Neutron star properties represent a crucial data point toward the development of a comprehensive theory of quantum chromodynamics (QCD). Key observables are masses, radii, and cooling behavior.

Neutron stars are discussed in more detail in Chapter 14. Here the focus is on accreting neutron stars. These are neutron stars that are part of a stellar binary system with a regular star as companion. There is an evolutionary phase in such a system where the two stars are close enough for matter from the regular star to flow over, fill an accretion disk, and eventually accumulate on the surface of the neutron star. Mass transfer rates are typically 10^{-8} to 10^{-10} solar masses per year. The gravitational energy released when stellar matter falls onto the neutron star surface, about 200 MeV nucleon^{-1}, heats surface and accretion disk and leads to very bright X-ray emission that makes these systems some of the brightest X-ray sources in the sky. Accreting neutron stars offer unique opportunities to study neutron stars. For example, their brightness greatly expands observational possibilities. Because of the accretion process, they have a well-defined atmosphere, enabling reliable modeling of X-ray emission, a prerequisite for reliable radius determination. Thermonuclear bursts trigger an expansion of the atmosphere that depends on the surface gravity offering an observational handle on neutron star compactness. In addition, the response of the neutron star to accretion-induced processes in the crust can be studied.

12.2.5.1 X-Ray Bursts

On top of the bright, continuous X-ray emission of an accreting neutron star, X-ray bursts are frequently observed. These bursts are powered by nuclear reactions that explosively burn the accreted hydrogen and helium into heavier elements via α-capture reactions, and the αp- and rp-processes. The nature of the reaction sequence depends strongly on the accreted composition and on any changes in the composition due to stable nuclear burning before ignition of the burst. Bursts last typically 10–100 s, though bursts lasting many hours, so-called intermediate long bursts, or, at the extreme, super bursts, have been discovered in recent years. X-ray bursts can lead to a dramatic radiation pressure-driven expansion of the neutron star envelope but very little mass is thought to be ejected. X-ray bursts are therefore unlikely to constitute major sources of Galactic nucleosynthesis, though small amounts of material might be observable.

The main motivation to study X-ray bursts is that they offer the brightest possible view of a neutron star. In addition, they occur on neutron stars with low magnetic fields and a well-defined accreted atmosphere making spectral modeling easier than for isolated neutron stars. The ultimate goal for X-ray burst studies is therefore to obtain information about the compactness of the neutron star, in terms of mass, radius, gravity, or surface redshift. There are many connections between X-ray burst properties and neutron star compactness – the luminosity needed to trigger the frequently observed expansion of the burning atmosphere depends on surface gravity; the gravitational surface redshift stretches burst light curves; and gravity affects the burst spectrum to just name a few.

A second goal is to understand X-ray bursts itself – with about 100 systems in the Galaxy, each typically bursting several times a day, X-ray bursts are the most frequently observed thermonuclear explosions of stellar nature. There are many open questions related to X-ray bursts. For example, sometimes time intervals between two bursts are as short as 10 min – this is not enough time to accrete the nuclear fuel needed to power the second burst. Another problem is the burst behavior as a function of accretion rate. Theory predicts an increase of the burst rate with increasing accretion rate until the Eddington accretion rate is reached (the accretion rate where the gravitational energy release generates enough radiation pressure to expel material) at which point, coincidentally, a transition to stable burning without bursts should occur. Instead what is observed in systems where the accretion rate varies over time is a drastic reduction of burst frequency as the accretion rate increases beyond 10–30% of the Eddington accretion rate. A new phenomenon that is not well understood is that of the so-called super bursts, rare bursts with recurrence times of about a year that extend over many hours releasing more than 1000 times as much energy than regular bursts. While models assuming the ignition of a deep carbon layer accumulated over years explain the observations, its unclear what causes the carbon accumulation and what makes it ignite after a year.

The lack of understanding of various X-ray burst phenomena also hampers attempts to extract neutron star information. On the observational side, progress has been made with advanced X-ray observatories such as Beppo-SAX, RXTE, Chandra, XMM-Newton, and Swift providing data on thousands of bursts. An extensive multi-instrument burst database MINBAR is currently under development and will contain close to 5000 bursts. One-dimensional X-ray burst models that follow in detail the nuclear reaction sequences also exist. A promising approach to analyze the large amount of burst data is then to generate sets of theoretical burst templates where system parameters are varied, and compare this set of burst templates to the observational data. A recent example for this approach is the attempt to determine the surface redshift, which stretches burst profiles by comparing a set of burst model templates calculated with different redshifts to the observational data [14]. This is important because the surface redshift directly constrains compactness of the neutron star. Yet there are problems with this approach as the best fit of the burst tail does not provide the best fit to the burst peak. This is not surprising as the predicted differences in the burst light curve produced at different surface redshifts are smaller than the burst light curve

uncertainties induced by uncertain proton capture rates. Better nuclear physics is needed to fully analyze burst observations quantitatively.

Figure 12.3 shows the reaction sequence during an X-ray burst that burns a mix of hydrogen and helium. For lighter elements α-induced reactions dominate such as the 3α-reaction and the αp-process, a sequence of alternating (α,p) and (p,γ) reactions that uses protons (hydrogen) as a catalyst. For heavier elements, the larger Coulomb repulsion between α-particles and nuclei prevents α-induced reactions, and the rp-process takes over, a sequence of rapid proton captures and β^+ decays. For the most hydrogen-rich bursts, the rp-process ends in an Sn–Sb–Te cycle. Similar to the r-process, masses determine the degree of photodisintegration that counteracts proton capture and therefore the reaction path. β^+ decay half-lives determine then how fast the process proceeds along a given path. The slowest β-decaying nuclei, the so-called waiting points, are of particular importance as they set the timescale for the nuclear energy release and therefore shape the burst tail. Existing radioactive beam facilities have produced all the isotopes along the rp-process, and all half-lives have now been determined experimentally (though in some cases, they might be altered by the extreme stellar environment with densities up to $10^6\,\mathrm{g\,cm^{-3}}$ and temperatures up to 2 GK). Masses have been measured in most cases, though measurements still need to be extended to nuclei with 1–2 neutrons less to cover all of the rp-process. The main challenge for the future is measurements of α- and proton-induced reaction rates.

12.2.5.2 Accreted Neutron Star Crusts

The ashes of hydrogen- and helium-driven nuclear burning on the surface of accreting neutron stars, either in the form of X-ray bursts or steadily, remains largely on the surface of the neutron star. With typical accretion rates of 10^{-8} solar masses per year ($500\,\mathrm{tons\,s^{-1}\,m^{-2}}$), it takes only about 10 000 years to replace the entire original crust of the neutron star with ashes. While isolated neutron stars have a crust composition that is determined by nuclear statistical equilibrium at the extreme temperatures when the neutron star forms in a supernova explosion, the crust of an accreted neutron star is formed of burst ashes at low temperatures and is not in equilibrium.

To understand the composition profile of the crust of an accreted neutron star, one can follow the fate of an accreted fluid element that falls onto the neutron star surface, undergoes nuclear burning in an X-ray burst, and is then buried by the ongoing accretion and pushed deeper and deeper into the crust. The increasing density will raise the electron Fermi energy making the capture of electrons by nuclei energetically possible. This will convert the nuclei step by step into more neutron-rich species, while emitting energy in the form of neutrinos. At a depth corresponding to a density of around $5 \times 10^{11}\,\mathrm{g\,cm^{-2}}$, nuclei will be so neutron rich that converting more of their protons into neutrons will result in a neutron-unbound nucleus that will emit the newly formed neutron. At this depth, neutrons will therefore start to drip out of nuclei and the crust will be composed by a mixture of nuclei and free neutrons. At still greater depth, so-called pycnonuclear fusion reactions occur, fusion reactions induced by high density rather than high temperature. The result is a complex network of reactions that includes neutron capture, neutron release, electron capture, β-decay, and fusion reactions.

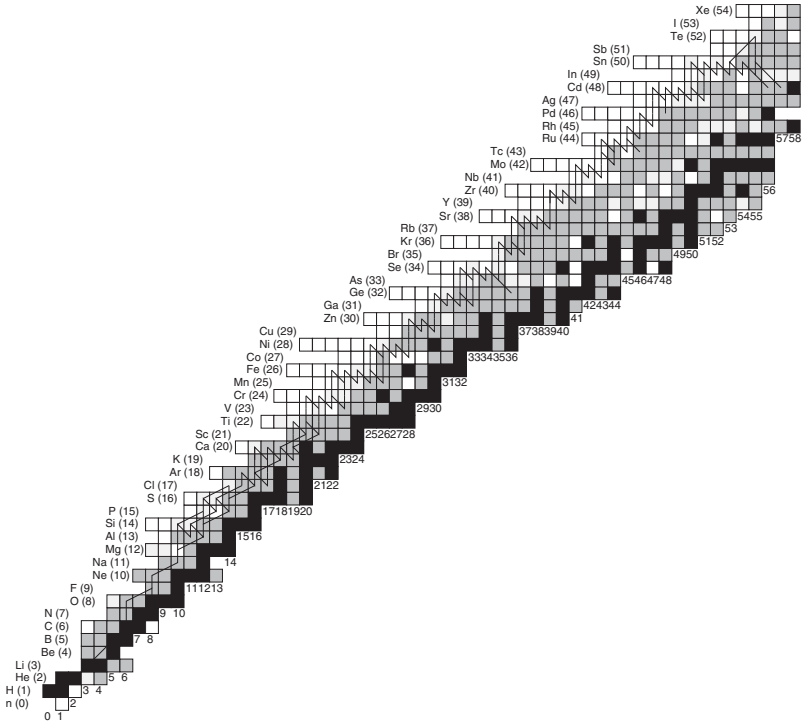


Figure 12.3 The main rp-process path on the chart of nuclides. Shaded in gray are nuclei for which the mass has been determined experimentally with an error of 10 keV or better.

The nuclear reactions in the crust of accreting neutron stars determine the crust composition and therefore, together with the actual spatial arrangements of the nuclei in the solid crust, mechanical, electrical, and thermal properties. During the accretion process, the nuclear reactions in the crust operate and release heat that leads to deep crustal heating. Because of its limited thermal conductivity, the crust of accreting neutron stars becomes significantly hotter than the neutrino-cooled neutron star core, with a characteristic temperature profile that is set by the location of the various nuclear-heating processes. This characteristic temperature profile can be directly observed in a class of transient systems, where accretion occurs for many years, only to turn off rather suddenly. During the subsequent period of quiescence, the luminosity of the system is drastically reduced, but a small residual luminosity can be observed with the largest X-ray observatories. This residual luminosity is found to decline on timescales of years and is thought to represent the cooling of the hot neutron star crust. With progressing time, this residual luminosity originates from deeper and deeper layers of the crust probing its temperature profile. For example, a kink in the power law describing the decline of luminosity has been interpreted as a signature of the appearance of free neutrons at neutron drip, which alters the heat capacity of the crust at that depth [15]. A handful of these cooling transients have been observed so far, with puzzling variations, and puzzling similarities despite vastly different heating histories. Explaining the observations also seems to require a very strong heat source in rather shallow regions of the crust that is not predicted by current nuclear reaction models.

To understand and interpret these observations and to extract information on the physics of neutron star crusts the nuclear physics needs to be understood. One important piece of nuclear physics is the location of neutron drip – the maximum number of neutrons that can be added to the nucleus of a given element. Neutron drip and nuclear heat release depend sensitively on nuclear masses. However, measurements with current radioactive beam facilities have not reached the neutron drip line, which is only known up to oxygen. Here, the new generation of rare isotope facilities will have a major impact enabling the mapping of neutron drip for elements as heavy as germanium ($Z = 32$) – the entire range needed for models of neutron star crusts (the heaviest nuclei that can be made in X-ray bursts have mass number $A = 106$ which according to current predictions corresponds to a neutron drip line around germanium). Electron capture rates, β-decay rates, and fusion rates are also important. These reactions are dramatically altered by the extreme density conditions inside the neutron star crust and cannot be measured directly in the laboratory. However, the underlying transition strengths for electron captures and β-decays can be probed by charge exchange reactions (Section 12.3.3). The dependence of fusion reactions on neutron richness may be probed by comparing fusion reaction measurements with more and less neutron-rich nuclei.

In addition, deep crustal heating depends sensitively on the initial composition at the neutron star surface created by X-ray bursts [16]. Therefore, X-ray bursts and their underlying nuclear reactions need to be well understood to model neutron star crusts. In particular, reliable X-ray burst models based on experimental

nuclear data would allow one to predict the initial composition for crust processes for a specific observed neutron star. This could be accomplished by comparing the X-ray burst light curves observed before the system's transition to quiescence with templates from theoretical X-ray burst models.

12.3 Nuclear Physics for Astrophysics

Nuclear physics and astrophysics are intimately connected. While some of the relevant nuclear physics is understood and the relevant nuclear data measured with sufficient precision, this is not the case in general – despite more than a century of nuclear physics research. This section gives an overview of how this problem is being addressed for the different types of nuclear data needed in astrophysical models using innovative approaches and new facilities.

12.3.1 Stellar Fusion and Capture Reactions Among Stable Nuclei

Capture or fusion reactions among stable nuclei are extremely slow at the low energies the nuclei have in stars. The reason is the electrostatic repulsion between the nuclei that requires tunneling for the reaction to occur. The tunneling probability decreases exponentially with energy. The challenge of measuring these rates is being addressed by dramatically increasing the sensitivity of experiments. This allows experimentalists to push measurements to lower energies that are closer to the astrophysical energy range resulting in a more reliable extrapolation. Several avenues are pursued to achieve this increase in sensitivity: one is to install accelerator facilities deep underground where cosmic-ray-induced backgrounds are much reduced. The pioneering facility for this approach is LUNA in the Gran Sasso Underground Laboratory in Italy. A new-generation underground accelerator laboratory DIANA is currently under development in the United States. DIANA will feature higher-intensity and higher-energy beams. A new-generation high-density gas jet target will be used at DIANA to avoid target degradation from the bombardment with intense beams. A second approach to increase the sensitivity of stable beam fusion rate measurements is the so-called inverse kinematics technique in connection with a recoil separator. This is currently pursued at the ERNA facility at CIRCE, Italy, and the St George facility at the University of Notre Dame in the United States for application to α-capture reactions. Instead of bombarding a heavy target with an ^{4}He-beam, a heavy ion beam bombards a ^{4}He gas jet target. This has the advantage that when the reaction product, the recoil, leaves the target, it can be separated from the intense unreacted beam with a dedicated and optimized recoil separator, and be detected together with the γ-ray emitted in the capture reaction. This coincidence requirement dramatically reduces background. Finally, detector improvements are being employed to increase sensitivity of experiments. Examples include new coincidence techniques using segmented γ-detectors that take advantage of the fact that specific multiple γ-rays are emitted after a typical fusion reaction.

Another approach is to measure the inverse reaction of a capture reaction – the splitting of a nucleus in a γ-ray beam. The time-reversal symmetry of the nuclear transition matrix elements allows to infer from

such measurements the capture reaction cross section. For specific reactions, this is a promising approach currently pursued, for example, at the HIgS facility in the United States.

Even with the most sophisticated experiments, a measurement of a fusion reaction rates at astrophysically relevant energies is only possible for a few cases. Nuclear theory is therefore needed to extrapolate from the measured energy region to the astrophysical energy region. A promising new development here is the combination of fully microscopic nuclear structure calculations with reaction theory that has enabled first calculations of astrophysical fusion reactions largely from first principles. This was demonstrated for example with a calculation of the ^{3}He+α fusion rate as a function of temperature [17]. However, such approaches are limited to nuclei lighter than about ^{12}C owing to limitations in computing power and therefore not of general applicability in the foreseeable future.

12.3.2 β-Decays at Moderate Astrophysical Conditions

β-decay rates are important in the r-process and the rp-process. β-decays are the slowest reactions in the reaction chain and determine therefore the timescale of the process. In addition, in any reaction chain, after some time at constant conditions, steady-flow equilibrium is achieved. In steady flow, $\lambda n = \text{const}$ for all nuclei, where n is the number density of the nucleus and λ its decay rate. Therefore, the abundance of a nucleus is inversely proportional to its decay rate. As the β-decays are by far the slowest reactions, it is only the β-decaying nuclei that have significant abundance. Among these β-decaying nuclei, the abundance depends linearly on the lifetime. It is therefore the β-decay rates that determine the abundance pattern created in the r-process and the rp-process for a given process path. On the basis of the experimental data available for some r-process waiting points and the r-process abundances deduced from the composition of the Solar System, it can be shown that $\lambda n = \text{const}$ does not hold globally. Clearly, r-process conditions change too rapidly for steady-flow equilibrium to be achieved. Nevertheless, the reaction chain will evolve toward steady flow, and therefore an individual abundance will still tend to be inversely proportional to the respective waiting point decay rate.

β-decay rates are therefore essential for predicting the abundance patterns created by an r-process model. The same holds true for the β^+ decay rates in the rp-process. In principle, these decay rates can be determined experimentally. Traditional methods employ implantation of a radioactive beam of the isotope of interest into a passive or active stopper. After some time, the stopper is moved in front of a detector, where the decay rate as a function of time is recorded. The process is repeated until the desired statistics is established. More recently, more efficient methods have been developed at radioactive beam facilities that produce a mixed radioactive beam by projectile fragmentation. The beam is implanted continuously into a highly segmented silicon detector array such that the average time between implantations in a detector pixel is long compared to the half-lives. The silicon detector array detects the implantation event and the decay event and both get time-stamped with a precision clock. Decay events are then correlated to preceding implantation events in the same pixel, and the time difference is the

decay time for the particular nucleus. The implanted nucleus can be identified event by event using upstream detection systems. With this method, decay half-lives have been determined with production rates as low as one per day for 10–20 types of nuclei simultaneously. This allows one to push these types of measurements to the most neutron- and proton-rich nuclei. While by now all decay rates in the rp-process have been measured, access to the r-process nuclei has been limited. β-decay half-lives have only been measured for 47 of the roughly 300 waiting points in the r-process that are relevant for the final composition (the exact numbers depend on the r-process model). In particular, no half-lives are known for the progenitors of r-process abundances beyond $A = 139$. This is a problem, as the main r-process pattern observed in metal poor stars spans the mass range from $A = 130$ to $A = 232$. This lack of nuclear data currently prevents a prediction of the r-process abundance patterns generated by specific r-process models and their comparison with observations.

β-decay half-lives measured in the laboratory must be corrected for effects induced by the stellar environment. Experimental β-decay half-lives characterize the decay of the ground state of the nucleus. In both, the r- and the rp-process temperatures of 1–2 GK are reached, leading to thermal excitations of nuclei with typical energies of $kT \approx 100$ keV. Provided the nucleus has low-lying excited states, the decay occurs then not only from the ground state but also from the various populated excited states. The decay rate from an excited state can be considerably different compared to the ground state. The population of excited states depends sensitively on the temperature, so this leads to a temperature-dependent decay rate.

For the β^+ decay rates in the rp-process, an additional complication is the competing electron-capture decay. In terrestrial experiments, the nucleus is typically neutral (though storage ring experiments can be carried out where this is not the case) and electron capture of atomic electrons adds to the β^+ decay rate. As a result, a combined electron-capture β^+ decay half-life is measured. In the stellar environment, nuclei are fully ionized and the electron-capture component therefore has to be subtracted. On the other hand, nuclei in the rp-process are embedded in an astrophysical plasma and they can capture plasma electrons. This leads to a density-dependent decay rate. The effect is small for most nuclei owing to the rather modest densities reached during an rp-process in X-ray bursts. Nevertheless, it does play a role in some cases, in particular for long-lived isotopes that have only a negligible β^+ component such as ^{56}Ni.

12.3.3 β-Decays and Electron Captures at Extreme Astrophysical Conditions

At the extreme temperatures and densities in the inner regions of core-collapse and thermonuclear supernovae, highly energetic electrons can induce electron captures. In such extreme environments, the electron-capturing nucleus can be in very highly excited states, and the high electron energies can populate highly excited states in the final nucleus. Similarly, β-decays occur on highly excited states. In the crust of accreting neutron stars, high densities lead to highly energetic electrons that induce electron captures on neutron-rich nuclei. All these types of decays do not occur in terrestrial environments and cannot be determined experimentally. For not too heavy nuclei in

the sd- or pf-shells, shell-model calculations can be used to predict these rates. The effective interactions employed in these models can be tested using experimental data. Ground-state β-decays and electron captures can be used for that purpose, but they tend to populate predominantly low-lying excited states within the energetically accessible excitation energy range.

A method to overcome this limitation is by way of charge-exchange reaction measurements. In these reactions, a proton in the target nucleus is exchanged with a neutron (or vice versa), while in the projectile, the opposite occurs. Typical reactions employed are (p,n), (n,p), (d,2p), (^{3}He,t), (t,^{3}He), and others. It has been shown that at energies above 100 MeV u^{-1}, the 0° cross sections of these reactions are proportional to the weak-interaction matrix elements for allowed Gamow–Teller transitions, which tend to be the most important ones for the electron captures and β-decays in supernova environments. The method can be calibrated using decay data. The major advantage of this approach is that the Gamow–Teller strength can be measured up to rather high excitation energies, including the Gamow–Teller giant resonance around 10–15 MeV, providing an excellent constraint for shell-model calculations in the excitation energy regime of relevance for the astrophysical rates. Efforts are underway to expand the approach to unstable nuclei. To that end, a first experiment with a radioactive ^{34}P beam using the (^{7}Li,^{7}Be) reaction has been performed recently providing information on electron capture on ^{34}P, producing ^{34}Si [18]. This is important as in core-collapse supernovae, and especially in neutron star crusts, electron capture rates of very unstable nuclei need to be known.

12.3.4 Masses

The story of nuclear masses is intertwined with the history of nuclear astrophysics since the beginning of the field, when the first measurements by Aston in 1927 revealed the mass defect in ^{4}He, which together with Einstein's mass–energy equivalence paved the way for identifying nuclear reactions as the elusive energy source of the Sun and the stars by Eddington and others. At present, the masses of the stable nuclei have been measured with sufficient precision to calculate the energy generation in the Sun and in stars (reaction rates are the big nuclear physics challenge in that context – see Section 12.3.1). However, for processes that involve very neutron deficient or very neutron rich nuclei this is not the case.

The challenge today is to measure the masses of the many short-lived nuclei in the rp-process, in the r-process, and in neutron star crusts. Nuclear masses enter the astrophysical model calculations in different ways. In the r-process (and similarly in the rp-process for protons instead of neutrons), rapid neutron captures are halted by inverse photodisintegration such that an equilibrium between rapid capture and rapid disintegration reactions is established in the reaction $(A, Z) + n \leftrightarrow (A+1, Z) + \gamma$ (with (A, Z) being a nucleus with mass number A and element number Z). The ratio of the number densities n of the two nuclei in such an equilibrium is simply determined by the chemical potentials of the participating particles and is given by the Saha equation:

$$\frac{n_{A+1}}{n_A} = n_n \frac{G_{A+1}}{2G_A} \left(\frac{A+1}{A} \frac{2\pi\hbar^2}{m_u kT} \right)^{\frac{3}{2}} \times \exp(S_n/kT) \qquad (12.1)$$

with J being the particle spin, G the partition functions accounting for the spins including of thermally excited states, n_n the neutron number density, T the temperature, and m_u the atomic mass unit. Clearly, the abundance ratio depends exponentially on the neutron separation energy S_n, the energy required to remove a neutron from a given nucleus, which can be calculated from the nuclear masses as $S_n = m_{(Z,A)} + m_n - m_{(Z,A+1)}$. With successive application of Eq. (12.1) one can show that one, or at most two nuclei along a neutron capture chain are favored by this equilibrium – this is the so-called waiting point nucleus that then has to β^- decay for the process to continue. According to Eq. (12.1), neutron separation energies and therefore masses need to be known with a precision of the order of kT, which for r- and rp-process conditions is around 50–100 keV. The typical difference in S_n between neighboring nuclei, and therefore the change in S_n that would result in a significant shift of the favored nucleus is of the order of 100 keV resulting in a similar precision requirement. For a nucleus with mass $A = 150$, this corresponds to a precision requirement of 1 : 1 000 000.

Nuclear mass models achieve at most 400–700 keV accuracy for known masses. However, predictions for unknown masses from different mass models differ often by more. For example, the average difference for predicted r-process masses from the finite-range droplet model (FRDM) and Hartree–Fock–Bogoliubov (HFB-21) mass models amounts to 2.8 MeV. This indicates much larger uncertainties when mass models are applied to more neutron-rich nuclei. In addition, unknown nuclear structure effects such as changes in shell structure for unstable nuclei induce potentially even larger systematic uncertainties that are difficult to quantify without data. Precision mass measurements are therefore the only way to provide reliable nuclear mass data for astrophysical models.

Nuclear masses also impact astrophysical model calculations in another way. Most of the β-decay, electron capture, neutron capture, or proton capture rates are not known experimentally and therefore need to be calculated with theoretical models. These calculations depend strongly on nuclear masses. This is most apparent in the rp-process where many reaction rates depend on a few narrow resonances. In this case, the statistical models typically used to predict astrophysical reaction rates cannot be applied. Reaction rates depend in many cases on the proton decay width Γ_p of the resonant state, which is governed by the proton tunneling through the Coulomb barrier and therefore depends exponentially on the proton separation energy and the nuclear masses. Typical mass model uncertainties can result in reaction rates that are uncertain by many orders of magnitude. A precision of typically 1–10 keV is needed to enable theoretical predictions of capture rates that are a few times better.

Measuring the mass of difficult-to-produce, unstable nuclei with a precision of 1 : 1–100 000 000 is a major challenge. Many masses in the rp-process (Figure 12.3) and the vast majority of masses in the r-process (Figure 12.2) or neutron star crusts are unknown because current radioactive beam facilities cannot produce these nuclei. This severely limits the predictive power of current astrophysical models. Nevertheless, many important measurements have been performed in recent years.

In this context, the development of Penning Trap mass measurements has revolutionized nuclear astrophysics. In that

technique, a radioactive beam is slowed and cooled to extremely low energies and is then trapped in a magnetic solenoid field that confines particles in a plane perpendicular to the beam axis. Forward and backward motion is inhibited by repelling electrodes of an electrical quadrupole field arrangement, effectively trapping the particle. The mass can then be determined through the measurement of the frequency of the particle's motion in the magnetic field. The main advantage of this method is that a precision of 1 : 1 000 000 can be achieved rather straightforwardly, and with some effort much higher precision can be reached. (Penning Trap mass measurements have reached precisions beyond 1 : 100 000 000 000 for stable nuclei where measurements can be performed over extended periods of time.) For the precision required for astrophysical applications, the method is sufficiently fast to carry out measurements with half-lives as low as a few tens of milliseconds.

However, Penning Trap measurements can only measure one mass at a time, and during the transfer of the beam into the trap, losses occur especially for very short-lived nuclei. Complementary time-of-flight methods have therefore been developed that are less precise, but can measure hundreds of nuclei simultaneously and push measurements toward the most exotic nuclei that can be produced. Often mass values obtained with Penning traps are used as reference points. Two techniques are commonly used. At GSI in Germany and Lanzhou in China, a mixed radioactive beam is injected into a storage ring with 108- or 130-m circumference, respectively. The circulation frequency of the nuclei is then measured, either by inserting a thin detector inducing minimal energy loss when a stored particle passes through, or by using a nondestructive pickup electrode. The former method registers a single particle during almost every turn and is therefore more sensitive, though the energy loss in the detector leads to changes in the particle orbit after many turns. The latter method is less prone to systematic errors but requires somewhat longer integration times and is therefore not suitable for the most short-lived nuclei. From the circulation frequency difference Δf between two ions, their mass-to-charge ratio difference $\Delta(m/q)$ can be obtained through

$$\frac{\Delta(m/q)}{m/q} = -\gamma_t^2 \frac{\Delta f}{f} + (\gamma_t^2 - \gamma^2)\frac{\Delta v}{v} \tag{12.2}$$

where v is the ion velocity and γ_t the relativistic $\gamma = 1/\sqrt{1 - v^2/c^2}$ for the transition velocity of the storage ring (for $\gamma < \gamma_t$ the circulation frequency increases with the momentum of the particle, while for $\gamma > \gamma_t$ it decreases with increasing momentum as changes in path length outweigh the gains from the increasing velocity). For a precision measurement to be possible, the velocity-dependent term on the right-hand side must be minimized, either by bringing the ions to the same velocity through cooling or, for short-lived nuclei where cooling takes too long, by injecting particles with $\gamma = \gamma_t$, thereby making the circulation frequency (to first order) independent of the ion velocity.

Another technique using a spectrometer has been developed at GANIL in France and is now also employed at MSU in the United States. In this method, a mixed radioactive beam that can contain hundreds of different nuclei is sent through a long beam line (82 m at GANIL and 59 m at MSU) that contains a spectrometer enabling a precise measurement of the particle's magnetic rigidity $B\rho = p/q$ with

p being the particle's momentum and q its charge, through a position measurement at a dispersive focus. In addition, detectors at the beginning and the end of the beam line section measure the time-of-flight and therefore the velocity v of each individual particle. The measurements of v and $B\rho$ can then be combined to obtain the mass-to-charge ratio of the nucleus

$$m/q = \frac{B\rho}{\gamma v} \tag{12.3}$$

In practice, it is impossible to measure $B\rho$ and v absolutely, and one measures differences in m/q to a set of well-known reference masses. Measurements can be performed continuously for hundreds of ions per second. The method can reach a mass resolution of about 1 : 10 000 and a precision of the order of 1 : 100 000 to 1 : 1 000 000, once a sufficiently large number of particles has been detected (of the order of a few thousands).

12.3.5 Capture Reaction Rates on Unstable Nuclei

The rates of proton- and α-induced reactions on unstable nuclei as they occur in the reaction sequence of the rp-process in novae and X-ray bursts can, in principle, be directly measured by using hydrogen or helium gas targets and bombarding them with a radioactive beam at low astrophysical energies – typically 0.2–1.5 MeV nucleon^{-1}. Such measurements, where the light projectile is the target and the heavy nucleus is the beam, are commonly referred to as *inverse kinematics* measurements, as opposed to the normal kinematics approach typically used to measure reactions among stable nuclei where a proton or helium beam hits a target containing the heavy nucleus. Measurements of (α,p) or (p,α) reactions can be performed with silicon detectors surrounding the gas target or using an active target. In the active target approach, the beam slows down in the target gas over a significant distance. The particles emitted in a reaction are detected and tracked, for example, in a time projection chamber. This enables identification of the particle type and reconstruction of the reaction vertex, which provides the reaction energy. The final state populated by the reaction can be determined from the measured energy of the outgoing particle. The active target approach has the advantage that the reaction cross section can be measured over a wide range of energy with a single measurement, making efficient use of radioactive beams that tend to be expensive to produce. A new promising concept that combines an active target with a silicon detection array for astrophysical reaction rate measurements is the Array for Nuclear Astrophysics Studies with Exotic Nuclei (ANASEN) detector system under development by an FSU/LSU collaboration in the United States.

The rates of capture reactions such as (p,γ) and (α,γ) are much harder to measure because γ-ray detection is much less efficient and is strongly limited by the ambient γ-ray background from natural radioactivity, cosmic-rays, and the radioactive beam itself. For heavy nuclei close to stability, for example, in the astrophysical p-process, the total energy released in multiple γ-rays for each reaction is very high. In this case, a highly efficient summing detector can be used to identify capture events by their large total energy that is beyond typical environmental backgrounds. The challenge is to overcome limitations from a high rate of lower energy background events, especially when used in inverse kinematics and with radioactive beams.

Nevertheless, this is a very promising approach that is currently being developed at MSU with the SUN detector.

However, the summing detection method is not applicable to proton and α-capture rates in the rp-process where reaction *Q*-values are lower and detection methods need to be more sensitive. This problem can be addressed using the recoil separator method that takes advantage of the fact that in inverse kinematics the nucleus produced in the capture reaction (the recoil) continues to move in the forward direction and leaves the target. While recoil and unreacted beam have similar momentum, there is a small difference in velocity owing to their slightly different mass. Therefore, a recoil separator system capable of separating nuclei by velocity can be installed after the target. The goal is to only transmit the recoil, or at least to suppress the unreacted beam sufficiently so recoils can be uniquely identified with a detector system behind the recoil separator. Detection of the recoil, if necessary in coincidence with a γ-ray detected at the target position, provides then a very clean signature of the reaction. Designing a successful recoil separator is not easy – because of the small mass difference between beam and recoil, very high mass separation is needed. In addition, the intensity ratio of recoil and beam is very low – ranging from $1 : 10^{12}$ to maybe $1 : 10^{16}$ at future facilities with very intense beams. So-called leaky beam background from small amounts of beam scattering off components or rest gas is unavoidable and sources cannot be modeled at the level of these very low recoil-to-beam ratios.

The first recoil separator used to directly measure an astrophysical proton capture reaction with a radioactive beam was the Astrophysics Recoil Separator (ARES) separator in Louvain La Neuve, Belgium where a strong resonance in the $^{13}N(p,\gamma)$ reaction was measured. The first separator built and optimized for astrophysical applications was the DRAGON separator in TRIUMF where recently a number of capture reactions could be measured. A new generation of recoil separators is now under development – examples are the St George project at Notre Dame or the SECAR project for FRIB. With these developments, separator technology has reached a stage where all reaction recoils can be transmitted and identified. The main limitation is now the beam intensity. Direct measurements require 10^7–10^{11} pps, depending on the reaction, to obtain at least a few reactions in a reasonable time. Such intensities are a major technical challenge. At so-called ISOL facilities, such as ISOLDE or TRIUMF/ISAC, radioactive beams are created in a thick target that stops an intense and energetic proton beam. While large amounts of radioactive nuclei are produced, chemical reactions make it difficult to extract the nuclei efficiently and fast from the target to reionize and reaccelerate them. Each element requires dedicated development of target and ion source technology that can take many years. There are a few particularly reactive elements that might be impossible to produce with this technique. An alternative approach is currently under development at MSU and FRIB in the United States, where radioactive beams are produced at higher energies by fragmenting a stable heavy ion beam in flight. The reaction products are stopped in a gas cell, or a solid stopper, and are then reionized and reaccelerated to produce high-quality beams at low astrophysical energies. The main advantage is that production and stopping with extraction are decoupled and can be optimized

separately, but there are complementary technical challenges.

Lower-intensity radioactive beams that are available more commonly can be used in the context of so-called indirect techniques. The goal is to measure some aspects of the reaction physics that when combined with theory allows one to derive a reaction rate that is much more reliable than a theoretical prediction. Such experiments are very important as they can provide information on very weak resonances where beam intensities are not sufficient for direct measurements. They are also required as a first step to guide direct measurements with information about the location and nature of resonances.

Reactions that transfer neutrons, such as (d,p), or protons, such as (^{3}He,d), can be used to constrain the transition matrix elements that determine the neutron and proton capture rates, respectively – at least in cases where the photon emission branching is large, so the particle channel determines the rate. A complication is that the transfer does not necessarily occur through the same parts of the wavefunction as the astrophysical capture reaction. This makes the results reaction model dependent. This can be alleviated in the case of proton captures. Owing to Coulomb repulsion these reactions occur in stars at a relatively large distance from the nucleus; they are peripheral. The transfer reaction measurement can be designed to be peripheral as well, and an analysis based on asymptotic normalization coefficients (ANCs) can be carried out that takes advantage of the fact that the peripheral wave function has a simple dependence on distance, requiring only a measurement of the overall normalization or the ANC. The ANC is then all that is needed to determine the astrophysical reaction rate. The advantage is that the analysis is independent of the more complicated structure in the interior of the nucleus where the less understood strong force acts.

Another indirect technique sometimes employed is Coulomb breakup, where instead of the capture reaction the inverse photodissociation rate is measured. The main advantage is that the measurement can be carried out with much less beam intensity. For unstable nuclei, it is difficult to use photon beams. Instead, the radioactive beam interacts with a target with a high element number, typically lead, and one uses the quasi-virtual photons in the electric field of the target nucleus to break up the incoming radioactive ion. Care must be taken to select only events that did not experience any nuclear interaction, and corrections must be applied for the different properties of the quasi-virtual photons compared to the photons emitted in the capture reaction.

12.3.6 Neutron Capture Rates

Neutron capture rates on stable nuclei are needed for s-process studies and have been measured with good precision for many decades. A number of facilities in the United States and in Europe have intense neutron beams available. Two basic techniques are employed: in the activation technique, the ^{6}Li(p,n) reaction is used to generate a neutron spectrum that resembles the thermal spectrum during the s-process for a temperature of $kT =$ 25 keV. A sample is then irradiated with a well-defined dose of neutrons. After the irradiation, the decay radiation from the radioactive nuclei produced by neutron captures is detected in a low-background counting station. The

advantage of this method is that it provides the stellar reaction rate with a single measurement with very good precision and that sample sizes can be very small. However, the method only works if neutron capture produces a radioactive isotope with a suitable half-life and suitable decay radiation. In addition, corrections need to be applied to obtain the reaction rate at different temperatures. A complementary approach that is widely used is the direct detection of the γ-radiation produced by a neutron capture event when a sample is bombarded with neutrons. This requires very efficient neutron detector arrays so multiple γ-rays emitted after a capture event can be summed. Typically, the incident neutrons have a wide energy spectrum, and the incident neutron energy is determined event by event using a pulsed neutron beam and time-of-flight measurements. Experiments of this type are performed at LANSCE in the United States and n-TOF at CERN, Switzerland.

A challenge for the future will be the expansion of this technique to more and more unstable nuclei, maybe into the 10–100 days half-life range. Radioactive samples are difficult to produce, and detectors need to be able to deal with the decay radiation. A possibility is to use isotope harvesting at new radioactive beam facilities to obtain the necessary sample material. Harvesting of beam-stop materials has already been used in the past, but more optimized methods are possible in facilities that produce radioactive beams by fragmentation. The fragment separator system used to select the desired radioactive beam can, in principle, be equipped with harvesting stations to capture specific isotopes. This approach is currently developed for the FRIB facility in the United States. The method has the advantage that very pure samples can be obtained.

However, r-process and neutron star models require neutron capture rates on nuclei with very short half-lives down to milliseconds. A direct measurement where a sample is irradiated with neutrons is then not possible. For such short-lived nuclei, neutron transfer reactions can be used to obtain data needed to calculate neutron capture rates. For more complex nuclei, the method can be combined with statistical model analysis. Indirect reaction techniques to infer neutron capture rates in combination with theory are commonly referred to as *surrogate techniques*. Such techniques are under intense development, including benchmarking with stable nuclei where results can be compared to direct measurements, and are of broad interest not only in astrophysics but also for applications such as the development of new types of nuclear reactors.

12.4 Outlook

Nuclear astrophysics can be expected to evolve significantly. A number of planned major accelerator facilities together with new techniques developed in nuclear theory promise to address the long-standing nuclear data challenges. New observational capabilities such as planned telescopes for advanced observations of time-variable astrophysical phenomena [e.g., the Large Synoptic Survey Telescope (LSST)], the prospect of larger optical telescopes that promise to reveal the chemical composition of fainter stars (e.g., the Giant Magellan Telescope), or the expected beginning of the era of gravitational wave astronomy will continue to drive the field in new

directions. In addition, advances in computing capabilities will enable to introduce significantly more realism in astrophysical model calculations of stars and stellar explosions.

These advances in capability will drive the field in a number of different directions. First, one can expect new discoveries. Nuclear astrophysics has a history of discovery of new nuclear-driven phenomena in the cosmos that open up completely new avenues of exploration. Recent examples include .Ia supernovae and the νp-process, phenomena discovered in astrophysical models, far more complex reaction sequences in the crust of accreting neutron stars discovered by nuclear physics, or the observational discovery of multiple r-processes. Another trend will be toward more precision. As the basic phenomena become understood and the reaction sequences identified, more and more precision, achievable with advances in experimental and theoretical nuclear physics and a quantitative understanding of uncertainties, will become a tool to extract the maximum amount of information from limited astronomical observations. Moreover, this precision promises to reveal what is going on deep inside neutron stars, supernovae, or other environments largely out of reach of direct observation. One example might be the r-process, where the detailed amounts of the produced elements contain information on astrophysical environments that can be unlocked once advanced rare isotope beam facilities have clarified the properties of extremely unstable nuclei.

Glossary

Ia supernova: Theoretically predicted explosion of a thick helium layer on a white dwarf that reaches about a tenth (hence .Ia) of the luminosity of a type Ia supernova.

Abundance: Amount of a specific isotope or element, typically expressed in mole per gram.

Accretion disk: The circular disk formed by matter with angular momentum flowing onto a compact object.

αp-process: Sequence of (α,p) and (p,γ) reactions occurring in X-ray bursts.

CNO cycle: Sequence of proton captures, β-decays, and (p,α) reactions that form a circular reaction path involving carbon, nitrogen, and oxygen isotopes. The CNO cycle generates energy in stars by burning hydrogen into helium.

Equation of state: Relation between pressure, temperature, and energy density.

FRDM: Finite-range droplet model, a model to predict nuclear masses.

HFB-21: Mass model based on Hartree–Fock–Bogoliubov theory.

Luminosity: Energy per second emitted by a star.

Metallicity: Fraction of elements heavier than helium.

νp-process: Sequence of rapid proton captures and (n,p) reactions possibly occurring in neutrino-driven winds in core-collapse supernovae.

Neutron star: Compact stellar object made of nuclear matter (mostly neutrons with small admixture of protons, electrons, muons, and other particles).

Neutron star crust: Thin layer of ordinary nuclei forming the outermost layer of a neutron star.

Neutrino-driven wind: Outflow of material from a neutron star forming in a core-collapse supernova, driven by heating from neutrino interactions.

Nova: Thermonuclear explosions of a fuel layer on an accreting white dwarf star.

Nuclear statistical equilibrium: Equilibrium composition created at high temperatures when all nuclear reaction channels are fast.

pp-chain: Reaction chain in stars burning hydrogen into helium.

Pycnonuclear reaction: Nuclear reaction where Coulomb repulsion is overcome by high density.

r-process: Rapid neutron capture process. A sequence of rapid neutron captures and β^- decays responsible for the origin of about half of the elements beyond iron.

Rare isotope: A radioactive, unstable isotope.

Reaction rate: Number of reactions occurring per second in a volume of material.

rp-process: Rapid proton capture process – a sequence of rapid proton captures and β^+ decays powering X-ray bursts.

s-process: Slow neutron capture process. A sequence of slow neutron captures and β^- decays responsible for the origin of about half of the elements beyond iron.

Seed: The nuclei where the reaction sequence of a nucleosynthesis process starts.

Solar abundances: Abundance of isotopes in the Solar System at the time of its formation, inferred from the surface composition of the Sun and primitive meteorites.

Supernova: Explosive end of stellar evolution. There are two types: thermonuclear supernovae (observationally, type Ia) triggered by the explosive thermonuclear burning of an entire white dwarf star, and core-collapse supernovae (observationally, type Ib, Ic, II, and others) triggered by the collapse of the iron core in a massive star.

Thermonuclear reaction: Nuclear reaction where Coulomb repulsion is overcome by high temperature.

X-ray burst: Thermonuclear explosion on the surface of an accreting neutron star.

References

1. Arnett, W.D. and Meakin, C. (2011) Turbulent cells in stars: fluctuations in kinetic energy and luminosity. *Astrophys. J.*, **741**, 33. doi: 10.1088/0004-637X/741/1/33.
2. Grevesse, N., Asplund, M., Sauval, A.J., and Scott, P. (2012) The New Solar Chemical Composition from Z = 0.02 to Z = 0.013, in *Progress in Solar/Stellar Physics with Helio- and Asteroseismology*, San Francisco: Astronomical Society of the Pacific Conference Series, 462 (eds H. Shibahashi, M. Takata, and A.E. Lynas-Gray), p 41.
3. Tur, C., Heger, A., and Austin, S.M. (2007) On the sensitivity of massive star nucleosynthesis and evolution to solar abundances and to uncertainties in helium-burning reaction rates. *Astrophys. J.*, **671**, 821–827. doi: 10.1086/523095.
4. Lodders, K. (2003) Solar system abundances and condensation temperatures of the elements. *Astrophys. J.*, **591**, 1220–1247. doi: 10.1086/375492.
5. Kappeler, F., Gallino, R., Busso, M., Picchio, G., and Raiteri, C. (1990) S-process nucleosynthesis – classical approach and asymptotic giant branch models for low-mass stars. *Astrophys. J.*, **354** (2, Part 1), 630–643.
6. Sneden, C., Cowan, J.J., Ivans, I.I., Fuller, G.M., Burles, S., Beers, T.C., and Lawler, J.E. (2000) Evidence of multiple R-process sites in the early galaxy: new observations of CS 22892-052. *Astrophys. J.*, **533**, L139–L142. doi: 10.1086/312631.
7. Travaglio, C., Gallino, R., Arnone, E., Cowan, J., Jordan, F., and Sneden, C. (2004) Galactic evolution of Sr, Y, And Zr: a multiplicity of nucleosynthetic processes. *Astrophys. J.*, **601**, 864–884. doi: 10.1086/380507.
8. Honda, S., Aoki, W., Ishimaru, Y., Wanajo, S., and Ryan, S.G. (2006) Neutron-capture

elements in the very metal poor star HD 122563. *Astrophys. J.*, **643**, 1180–1189. doi: 10.1086/503195.

9. Winteler, C., Käppeli, R., Perego, A., Arcones, A., Vasset, N., Nishimura, N., Liebendörfer, M., and Thielemann, F.K. (2012) Magnetorotationally driven supernovae as the origin of early galaxy r-process elements? *Astrophys. J.*, **750**, L22. doi: 10.1088/2041-8205/750/1/L22.
10. Korobkin, O., Rosswog, S., Arcones, A., and Winteler, C. (2012) On the astrophysical robustness of neutron star merger r-process, *ArXiv e-prints*.
11. Schatz, H., Toenjes, R., Pfeiffer, B., Beers, T.C., Cowan, J.J., Hill, V., and Kratz, K.L. (2002) Thorium and uranium chronometers applied to CS 31082-001. *Astrophys. J.*, **579**, 626–638. doi: 10.1086/342939.
12. Roberts, L.F., Woosley, S., and Hoffman, R. (2010) Integrated nucleosynthesis in neutrino-driven winds. *Astrophys. J.*, **722** (1), 954–967.
13. Fröhlich, C., Martínez-Pinedo, G., Liebendörfer, M., Thielemann, F.K., Bravo, E., Hix, W.R., Langanke, K., and Zinner, N.T. (2006) Neutrino-Induced Nucleosynthesis of A>64 Nuclei: The νp Process. *Phys. Rev. Lett.*, **96** (14), 142502. doi: 10.1103/PhysRevLett.96.142502.
14. Zamfir, M., Cumming, A., and Galloway, D.K. (2012) Constraints on neutron star mass and radius in GS 1826-24 from sub-eddington X-Ray bursts. *Astrophys. J.*, **749**, 69. doi: 10.1088/0004-637X/749/1/69.
15. Brown, E.F. and Cumming, A. (2009) Mapping crustal heating with the cooling light curves of quasi-persistent transients. *Astrophys. J.*, **698**, 1020–1032. doi: 10.1088/0004-637X/698/2/1020.
16. Gupta, S., Brown, E.F., Schatz, H., Möller, P., and Kratz, K.L. (2007) Heating in the accreted neutron star ocean: implications for superburst ignition. *Astrophys. J.*, **662**, 1188–1197. doi: 10.1086/517869.
17. Neff, T. (2011) Microscopic calculation of the ^{3}He(α,γ)^{7}Be and ^{3}H(α,γ)^{7}Li capture cross sections using realistic interactions. *Phys. Rev. Lett.*, **106** (4), 042502. doi: 10.1103/PhysRevLett.106.042502.
18. Zegers, R.G.T., Meharchand, R., Shimbara, Y., Austin, S.M., Bazin, D., Brown, B.A., Diget, C.A., Gade, A., Guess, C.J., Hausmann, M., Hitt, G.W., Howard, M.E., King, M., Miller, D., Noji, S., Signoracci, A., Starosta, K., Tur, C., Vaman, C., Voss, P., Weisshaar, D., and Yurkon, J. (2010) P34(^{7}Li,^{7}Be+γ) reaction at 100 AMeV in inverse kinematics. *Phys. Rev. Lett.*, **104** (21), 212504. doi: 10.1103/PhysRevLett.104.212504.

13
Relativistic Nucleus–Nucleus Collisions

Christoph Blume

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

13.1
Introduction

Relativistic nucleus–nucleus collisions provide a tool to study strongly interacting matter under extreme temperatures and densities in the laboratory. The most exciting aspect is the investigation of new states of matter, such as the quark–gluon plasma (QGP). In this phase, the confinement of quarks and gluons within hadrons is no longer effective and they exist as quasi-free particles. It is assumed that the early universe has passed through such a state of matter in its evolution a few microseconds after the Big Bang. In addition, deconfined matter might exist in nature in the core of superdense neutron stars.

In recent years, many interesting observations have been made, some of which are discussed in this review. They come from several experimental programs: at the Super Proton Synchrotron (SPS) and the Large Hadron Collider (LHC) at CERN, and at the Alternating Gradient Synchrotron (AGS) and Relativistic Heavy-Ion Collider (RHIC) at BNL. The combination of all these results leads to the commonly accepted notion that they are indeed best explained by the assumption of a QGP. Data from RHIC and LHC also provide much insight into the properties of this new state of matter (e.g., the behavior as a ''perfect liquid''). However, many questions are still unanswered, such as the exact structure of the quantum chromodynamics (QCD) phase diagram. For instance, it has been predicted that it might contain a critical point, whose discovery would be an extraordinary achievement. Currently ongoing and future programs (beam energy scan programs at RHIC and the SPS, planned facilities such as the Facility for Antiproton and Ion Research (FAIR) in Darmstadt and the Nuclotron-Based Ion Collider Facility (NICA) in Dubna) will certainly allow to expand our understanding of the properties of QCD matter under extreme conditions.

13.2
Phases of QCD Matter

The main features of the QCD phase diagram are summarized. The evolution of a heavy-ion reaction is discussed and an overview on the main experimental observables is given.

13.2.1
The QCD Phase Diagram

The main objective of ultrarelativistic heavy-ion physics is the exploration of the phase diagram of QCD. Figure 13.1

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

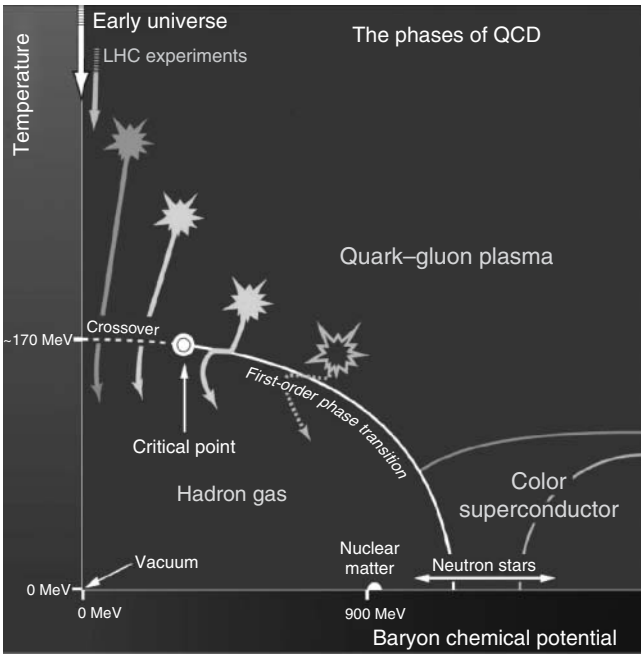


Figure 13.1 The QCD phase diagram. The arrows roughly indicate the different regions that can be explored by the current and future experimental programs. (Adapted from a figure in the NSAC Long Range Plan 2007.)

summarizes its basic features, the most important one being the phase boundary between a hadron gas (HG) and a new state of matter, the QGP. For a vanishing baryonic chemical potential $\mu_B = 0$, the location of the this phase boundary can nowadays be determined quite reliably by lattice-QCD calculations and is found to be around a temperature of $T \approx 155$MeV [1]. In addition, the various lattice-QCD results agree on the observation that the nature of the phase transition for $\mu_B = 0$ is of the type of a crossover (dashed line in Figure 13.1), that is, the relevant order parameters change fast but continuously. An example for a recent lattice-QCD result [1] on the temperature dependence of the energy density ϵ is shown in Figure 13.2. A rapid rise of the energy density can be observed around a transition temperature of $T_C \approx 155$MeV.

In the region of higher μ_B, the nature of the phase transition line is expected to change to a first-order type, indicated by the solid line in Figure 13.1. As a consequence, the phase diagram should also contain a critical point at the position where the crossover line turns into a first-order phase boundary. The search for this hypothetical critical point is the subject of several ongoing and planned experimental activities. It has been conjectured that additional phases of QCD matter, such as color superconductors, can exist in regions of very high values for the baryonic chemical potential. These phases might be relevant for the understanding of dense neutron stars, but most likely cannot be tested by accelerator-based experiments.

However, the QCD phase diagram can be probed experimentally in regions

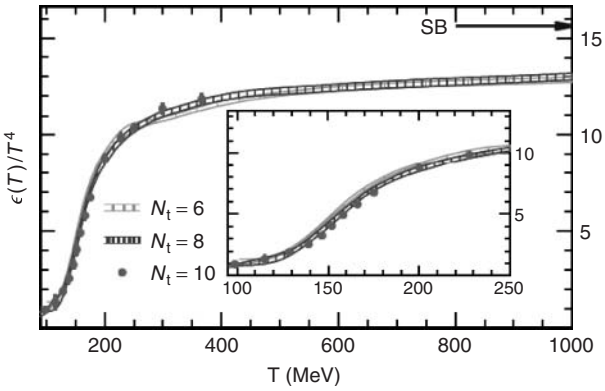


Figure 13.2 The energy density ϵ, divided by T^4, as a function of the temperature T [1]. Shown are results for lattices of different granularity in the time direction N_t. The arrow indicates the expectation for a gas of noninteracting particles in the Stefan–Boltzmann limit. (Courtesy of S.A. Bass, Duke University.)

of smaller μ_B by heavy-ion reactions. By varying the center-of-mass energies of the collisions, the reaction system is forced to follow different trajectories in the $T-\mu_B$ plane, as indicated by the arrows in Figure 13.1. At very high energies, such as provided by the LHC or RHIC, matter with almost zero μ_B is created, similar to the matter in the early universe. While cooling down, the system will pass the crossover region of the phase boundary. Heavy-ion reactions at these high energies are therefore well suited to study the properties of the QGP at high temperatures. Going toward lower energies, the initial temperature will decrease and μ_B will increase. This might allow to probe areas of the phase diagram where the phase boundary is of a first-order type, and, potentially, to locate the critical point. The corresponding energy regime is currently studied with the beam energy scan program at RHIC and at the CERN-SPS and will be further pursued at the future facilities FAIR and NICA.

13.2.2 Heavy-Ion Reactions

Figure 13.3 shows a schematic view of an ultrarelativistic heavy-ion collision. At the extreme left, the two incoming, Lorentz-contracted nuclei are depicted, followed by the pre-equilibrium state directly after the first binary collisions. The initial state of a high-energy heavy-ion reaction is a topic that is currently being discussed extensively. Following very basic considerations, the gluon densities within the parton distribution functions of the nucleons are expected to reach saturation values at very high energies. On the basis of this assumption, new theoretical concepts such as the color glass condensate (CGC) [2] have been developed, which might be used to describe this state of the reactions. In the middle of Figure 13.3, the expanding QGP

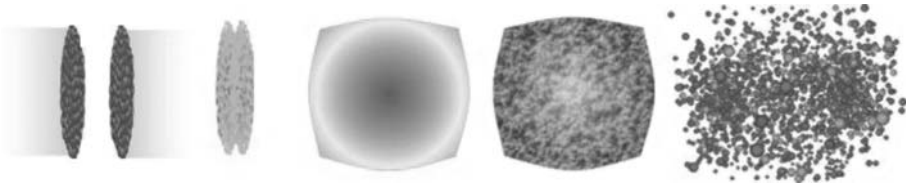


Figure 13.3 Sketch of a ultrarelativistic heavy-ion collision [3].

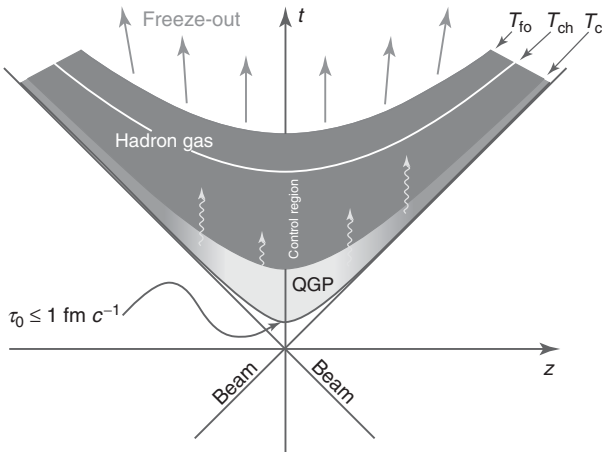


Figure 13.4 A space–time diagram of a heavy-ion collision. Shown is the evolution of the system versus time t and beam direction z. The hyperbolic lines denote the hadronization at T_C, the chemical freeze-out at T_{ch}, and the kinetic freeze-out at T_{fo}.

phase is shown, which, after hadronization, turns into a gas of interacting hadrons. These hadrons finally freeze-out, that is, they stop interacting when the system is dilute enough, and stream toward the detectors as free particles.

This evolution of a heavy-ion collision is also summarized in the idealized space–time diagram shown in Figure 13.4. After a formation time in the order of $\tau_0 = 1\text{fm}/c$, an equilibrated QGP phase is formed. Once this has expanded and cooled down to the critical temperature T_C hadrons form, which can interact further. The observable particle abundances are fixed as soon as all inelastic collisions stop, which defines the chemical freeze-out at the corresponding temperature T_{ch}. The phase-space distributions of the hadrons can still be altered until the kinetic freeze-out at T_{fo}, which is the temperature at which elastic collisions are not effective any more.

Different observables are sensitive to the various stages of a heavy-ion collision. Commonly, they are separated into soft and hard probes. The latter involve processes at a high momentum transfer, which therefore probe small distances and are mainly determined by the early stages of the collision. The following are examples for hard probes:

- Photons, which are emitted throughout the full evolution of the system and, because they only interact electromagnetically, are not modified by later stages of the reaction. Therefore, they can provide information also on the initial temperature of the QGP phase.
- Partons resulting from initial hard-scattering processes have to traverse the hot and dense medium. The free color charges in the QGP induce an additional energy loss of the partons via gluon bremsstrahlung. This effect can be exploited to study the properties of the QGP. The idea is to perform a tomography of the hot and dense matter by measuring the energy loss of the partons penetrating the medium.
- The production rates of quarkonia, that is, bound states of a heavy quark and antiquark such as the J/ψ, should be modified in a deconfined medium. As the $q\bar{q}$-potential will be screened in the presence of free color charges, a suppression of quarkonia production is expected at high temperatures. On the other hand, for J/ψ this mechanism

can potentially be counterbalanced at very high center-of-mass energies by the coalescence of $c\bar{c}$ pairs, which are relatively abundant at these energies. In this process, free charm and anticharm quarks group themselves together and form a quarkonium state.

Soft probes result from processes with small momentum transfers and are generally dominated by the late stages of the reaction. They provide information on the space–time evolution of the system, its kinetic and chemical freeze-out conditions, and its equation of state (EOS); in addition, they are potentially sensitive to the nature of the phase boundary and the presence of a critical point:

- Flow phenomena are driven by the pressure that is generated in the reaction system and are thus sensitive to the equation of state. Manifestations are radial flow that causes a blueshift of the transverse momentum spectra, and directed and elliptic flow. The latter provides information also on the early stages of the reaction and can be used to extract information about the viscosity of the system.
- Femtoscopy denominates the measurement of the space–time extent of particle-emitting sources via two-particle correlations, provides information on the freeze-out volume and the expansion velocity of the reaction system.
- The measurement of particle yields allows to determine the degree of chemical equilibration and to extract the chemical freeze-out parameters T_{ch} and μ_{B}. Strange particles play a particularly important role in this context, since an enhanced production of strangeness relative to elementary proton–proton collisions might be indicative of a fast equilibration due to partonic processes in a QGP.
- Event-by-event fluctuations of observables such as average transverse momenta, multiplicities, or particle ratios are generally sensitive to the susceptibilities of the studied matter. They can provide insight into the nature of the phase boundary where the medium properties change rapidly. They are also crucial observables for the experimental search for a critical point in the QCD phase diagram, as enhanced fluctuations are expected in the vicinity of a critical point.

An interpretation of heavy-ion reactions can of course not be based only on a single one of these measurements. Any of the above listed observables might to a certain extent be understood in a purely hadronic scenario. However, the combination of all observations indeed results in a picture that rather supports the point of view that the matter created at high energies is of partonic nature. Furthermore, performing systematic studies of any observable, such as a function of center-of-mass energy as well as system size, is important for the understanding of the complex physics of heavy-ion collisions. Especially for the search of the onset of deconfinement and the hypothetical critical point in the phase diagram, these systematic studies are mandatory.

13.3 Soft-Particle Production

In the following, selected topics concerning soft-particle production in heavy-ion collisions are discussed. An important aspect in this context is the determination of the volume of the reaction zone, which requires a well-defined procedure to deduce the

event centrality from measurements. Soft-particle production can to a remarkable extent be described by a statistical model approach, assuming chemical equilibration. The special role of strange particles is discussed, as strangeness enhancement as a consequence of production rates close to the equilibrium value has been proposed as a possible signature for a QGP formation.

13.3.1 Centrality Determination

The determination of the reaction volume in heavy-ion collisions is an important prerequisite for the comparison of experimental data to theoretical calculations. With decreasing impact parameter b, the number of participating nucleons (N_{part}) and also the number of binary nucleon–nucleon collisions (N_{coll}) increase. Since these quantities are not directly accessible by experiment, the connection to measurable quantities has to be made via model calculations. For this purpose, the Glauber model [5] is usually employed, which allows to relate b, N_{part}, and N_{coll} to experimental observables such as the multiplicity of charged particles seen by the detectors, or the total energy of the spectators that can be measured with calorimeters under very forward angles [4]. Using these measurements, events can be divided into classes of different centrality, usually defined by their fraction of the total inelastic cross section, and corresponding averaged values for b, N_{part}, and N_{coll} can be determined. Figure 13.6 shows, as an example, a schematic multiplicity distribution and its relation to the quantities derived from Glauber model calculations.

The relation between the number of participants and soft-particle production follows the assumption of the wounded nucleon model [6], where the number of produced particles is directly proportional to N_{part}. Even though this proportionality is not strictly valid in high-energy heavy-ion collisions, N_{part} is generally used as a scaling variable to characterize soft-particle production. Hard processes, on the other hand, should scale with the number of binary collisions. Therefore, N_{coll} is the natural scaling variable to determine nuclear modification factors for rare probes such as jets or quarkonia.

The Glauber model can also be used to determine the initial conditions of the reactions. As the distribution of participating nucleons is not uniform in configuration space for a given event (Figure 13.5a) and will also fluctuate event by event, this has to be taken into account when analyzing higher-order flow phenomena. The initial conditions determined from CGC models differ somewhat from the ones calculated with the Glauber model. Currently, many attempts are being made to distinguish between the two scenarios with the help of flow measurements, for example.

13.3.2 Stopping and Multiplicities

The kinetic energy of the participating nuclei is converted into compression and heating of the resulting fireball during the collision process. The amount of energy loss that they experience in this process therefore determines the energy density that can be achieved in heavy-ion collisions and, consequently, whether the conditions for a QGP formation are met. Since the baryon number is a conserved quantity, the measurement of net-baryon (i.e., baryon–antibaryon) rapidity distributions allows to extract information on the energy loss of the nucleons and thus to determine the stopping power of heavy-ion

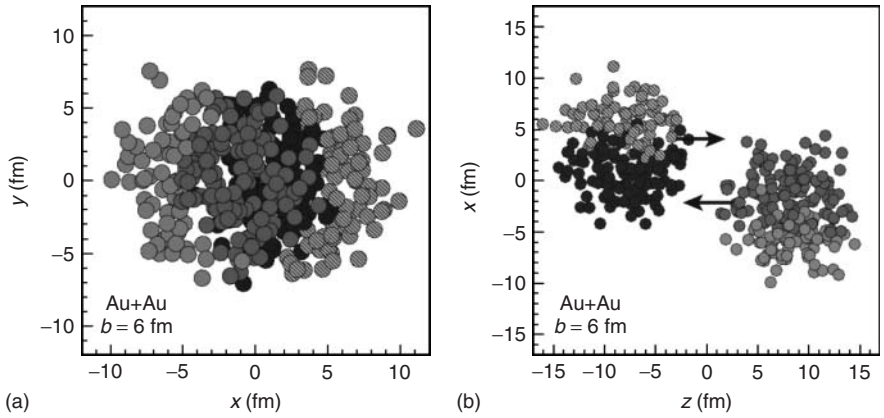


Figure 13.5 Result of a Glauber model calculation for an Au–Au collision at $\sqrt{s_{NN}} = 200\text{GeV}$ with an impact parameter of $b = 6$ fm [4]. Shown is a view in the transverse plane (a) and that along the beam axis (b). The participating nucleons are shown in dark gray.

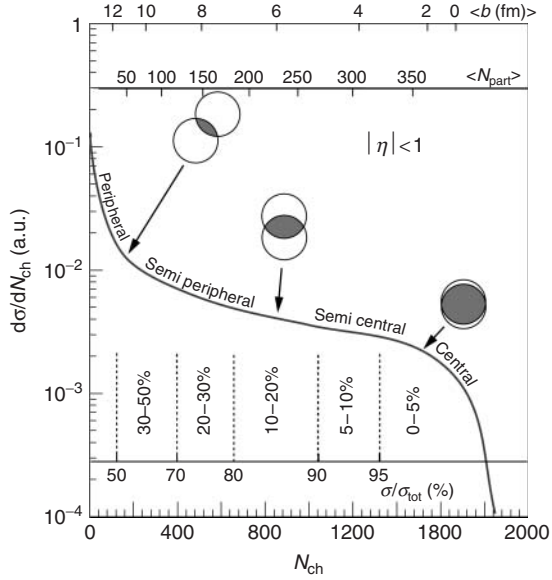


Figure 13.6 Illustration of the centrality determination via the number of charged particles N_{ch}, together with the impact parameter b, the number of participants N_{part}, and the corresponding fractions of the total cross section σ/σ_{tot} of the selected centrality classes [4]. The shown distribution is idealized and not a real measurement.

collisions. Because in most experiments not all baryons can be measured, net-protons are commonly used as the relevant observable. Figure 13.7a shows a compilation of net-proton rapidity distributions, as measured at the AGS, the SPS, and at RHIC. A strong energy dependence of their shapes is observed. While at low energies the net-protons peak at midrapidity, a dip in the middle of the distribution is seen at SPS energies, which develops further into a broad minimum over several units of rapidity at RHIC. The midrapidity region is therefore almost net-baryon free at high

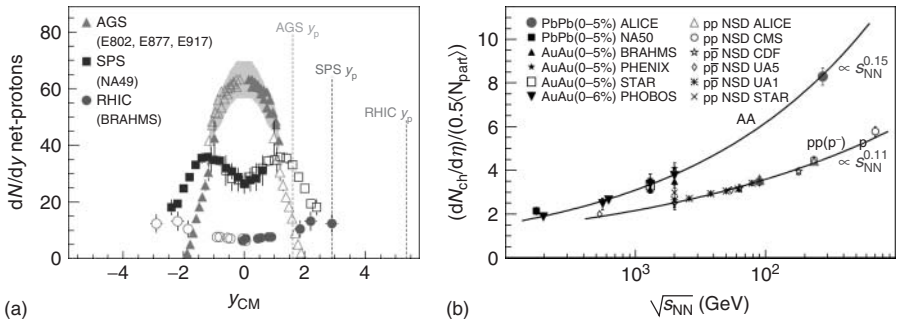


Figure 13.7 (a) The net-proton rapidity distributions for central Au–Au/Pb–Pb collisions for different center-of-mass energies (AGS: $\sqrt{s_{NN}} = 5$ GeV, SPS: $\sqrt{s_{NN}} = 17.3$ GeV, and RHIC: $\sqrt{s_{NN}} = 200$ GeV) [7]. (b) The charged-particle pseudorapidity density per participant pair for central nucleus–nucleus and nonsingle diffractive pp ($p\bar{p}$) collisions as a function of $\sqrt{s_{NN}}$ [8].

center-of-mass energies. The distributions shown in Figure 13.7a can be used to calculate the average rapidity shift of the baryons ($\langle\delta y\rangle = y_{proj} - \langle y\rangle$), where y_{proj} is rapidity of the incoming projectile and $\langle y\rangle$ the mean net-baryon rapidity after the collision. This quantity can vary between two extreme scenarios: full stopping, where the baryons lose all of their kinetic energy ($\langle\delta y\rangle = y_{proj}$), and full transparency, with no loss of kinetic energy ($\langle\delta y\rangle = 0$). It is found that the average rapidity loss increases roughly linearly as $\langle\delta y\rangle = 0.58 y_{proj}$ up to SPS energies, while the further increase toward RHIC is significantly smaller [7]. Using the net-baryon distributions and the average transverse mass $\langle m_t\rangle$ one can also estimate the total energy that the participants lose in the collision and that is available for particle production and collective effects. It turns out that this amounts to 70–80% of the initial kinetic energy for center-of-mass energies above $\sqrt{s_{NN}} > 6$ GeV [9]. This value is clearly higher than what is observed for pp collisions in the same energy region ($\sim 50\%$).

Part of the available energy is transformed into particle production. The midrapidity pseudorapidity density of charged particles is usually among the first observables studied as soon as a new energy regime can be accessed experimentally. Figure 13.7b shows a recent compilation of the data collected for heavy-ion and for pp ($p\bar{p}$) collisions as a function of $\sqrt{s_{NN}}$. The values for $dN_{ch}/d\eta$ per participant are clearly higher in AA than in pp from SPS energies onward and also rise more rapidly (AA: $\propto s_{NN}^{0.15}$, pp: $\propto s_{NN}^{0.11}$). Using the measured $dN_{ch}/d\eta$ values, the initial energy density ϵ_0 in the fireball can be estimated via the formula by Bjørken [10]:

$$\epsilon_0 = \frac{\langle m_t\rangle}{\tau_0 \mathcal{A}} \frac{dN}{dy}\bigg|_{y=0} \tag{13.1}$$

Here, $\langle m_t\rangle$ is the average transverse mass of the produced particles, τ_0 the formation time (typically, $\tau_0 \approx 1\text{fm}/c$), and $\mathcal{A}$ the transverse area of the overlap zone of the nuclei. Following this simplified ansatz, energy densities larger than $\epsilon_0 \sim 18\text{GeVfm}^{-3}$ are achieved for Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Already at the SPS it is about $\epsilon_0 \sim 3\text{GeVfm}^{-3}$ and thus clearly higher than the critical energy

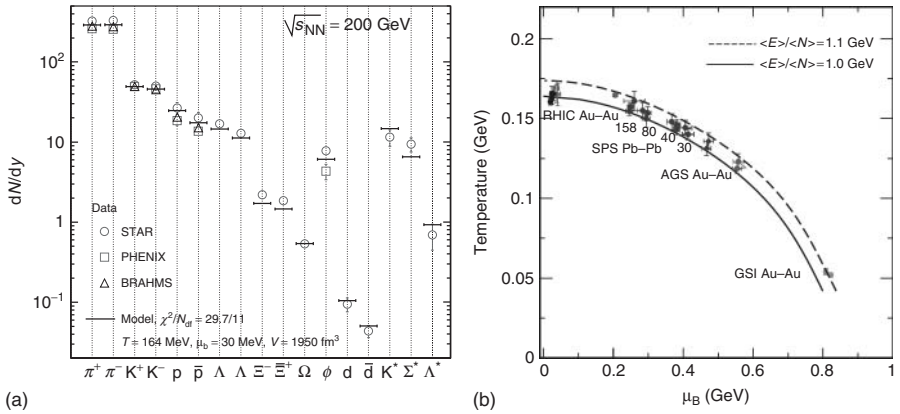


Figure 13.8 (a) An example of a statistical model result. Shown is a fit to measured hadron yields for central Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV [12]. (b) The chemical freeze-out parameters T_{ch} and μ_B for different energies, as obtained from fits to data from RHIC, SPS, AGS, and SIS. The lines correspond to a freeze-out condition of a constant average energy per particle $\langle E\rangle/\langle N\rangle$ [13].

density required for a QGP formation ($\epsilon_C \sim 0.7\,\mathrm{GeVfm}^{-3}$ [11]).

13.3.3 Statistical Model

Statistical models have generally been very successful in describing the yields of a large variety of particles up to multistrange hyperons as measured in heavy-ion collisions. These models are based on the assumption that particle yields correspond to their chemical equilibrium expectation and can thus be described by a partition function. In the case of large systems, as usually appropriate for central heavy-ion reactions, the grand-canonical partition function can be used:

$$\langle N_i\rangle = (2J_i+1)\frac{V}{(2\pi)^3} \times \int \mathrm{d}^3p \frac{1}{\gamma_s^{-s_i}\exp\left[(E_i-\mu\cdot\mathbf{q}_i)/T_{ch}\right]+1} \tag{13.2}$$

The essential parameters of this model are the volume V, the chemical freeze-out temperature T_{ch}, and the chemical potentials μ. Figure 13.8a, showing an example of a statistical model fit to particles yields measured at RHIC, demonstrates that a reasonable fit to all particle yields, up to the rare Ω^-, is possible with this small set of parameters. The success of the statistical approach is, however, not limited to heavy-ion reactions, but also provides reasonable fits to proton–proton collision data, if the model is based on a canonical ensemble instead [14]. Therefore, it is still an open debate whether the success of the statistical model is due to a dynamical equilibration or rather a general property of the hadronization process [15].

By performing these fits for central collisions at different center-of-mass energies, the energy dependence of the chemical freeze-out points can be extracted. It turns

out that the resulting values for T_{ch} and μ_B lie on a single curve, as shown in Figure 13.8b. An important feature of this freeze-out curve is the observation that it approaches the phase transition line, as determined by lattice-QCD, for high energies, where μ_B goes to 0. However, toward low energies (large μ_B) the freeze-out curve does not coincide any more with the deconfinement phase transition line. It has been conjectured that a different phase of matter, the so-called quarkyonic matter, might exist in the area between the two [16]. This hypothetical phase should still be confined in hadrons but have a restored chiral symmetry, which is broken in normal hadronic matter [17]. An experimental signature for this kind of phase has still to be established.

13.3.4 Strangeness

It has been suggested very early on that a fast dynamical equilibration of strangeness production in a QGP can lead to an enhancement of rare multistrange (anti-)baryons (Ξ, Ω) in heavy-ion relative to pp collisions [19] and thus be interpreted as a signature for a QGP formation. As strangeness production in hadronic reactions is strongly suppressed due to the corresponding high threshold energies, a dynamical equilibration in a HG should take a long time compared to that in a QGP, where partonic interactions with low thresholds can take place. In fact, such an enhancement, defined as the yield in A+A normalized by N_{part} and divided by the yield in pp, has been established experimentally at SPS and RHIC (Figure 13.9). It depends

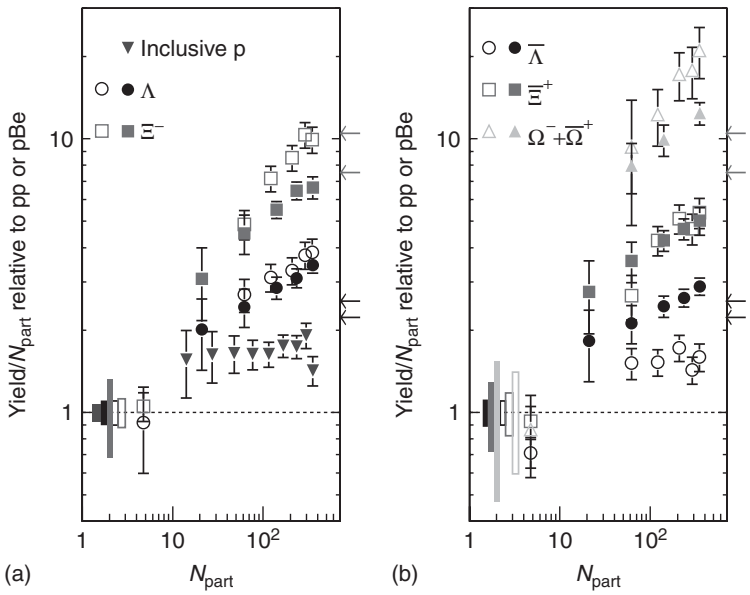


Figure 13.9 The enhancement factor for different baryons as measured at RHIC ($\sqrt{s_{NN}} = 200$ GeV) and at the SPS ($\sqrt{s_{NN}} = 17.3$ TeV) [18].

on the strangeness content of the particle and can reach values up to ~ 20 for the Ω. Models that involve only hadronic degrees of freedom are not able to describe this enhancement (for a recent review on this subject see [20]). However, so far no clear evidence for an onset of strangeness enhancement as a function of center-of-mass energy has been seen, as one might expect if it were due to the presence of a QGP. In fact, the enhancement rather increases toward lower energies, as can also be seen in Figure 13.9.

Another interesting feature related to strangeness measurements is the observation that the relative strangeness production is maximal around $\sqrt{s_{\mathrm{NN}}} = 7 - 8$ GeV. The most prominent observation of this effect is a sharp maximum in the energy dependence of the K^+/π^+ ratio [21]. Within a specific model approach, this structure has also been proposed as a signature for first-order deconfinement phase transition [22]. However, there are also alternative explanations of this feature within a normal statistical model approach [12].

13.4 Spectra and Flow

The analysis of transverse momentum spectra and of elliptic flow effects allows to extract information on the pressure of the system and thus on the EOS that describes it. Therefore, these observables might allow to determine whether the fireball produced in heavy-ion reactions contains a deconfined phase of quarks and gluons or whether it can be described throughout its whole evolution by hadronic degrees of freedom only. Elliptic flow measurements also yield insight into the fluid properties of the medium, namely, the shear viscosity over entropy density ratio.

13.4.1 Transverse Momentum Spectra

The pressure generated in heavy-ion collisions causes a collective expansion of the produced fireball. This radial flow manifests itself in a modification of the transverse momentum spectra of the measured hadrons. Owing to the additional outward boost, the spectra are blueshifted toward flatter distributions. This effect is stronger the larger the mass of the hadrons is. Thus, the observed increase of the inverse slope parameters, as extracted from fits with an exponential function to the p_{t} spectra, with the mass of the hadron is an indication of the presence of radial flow.

A proper description of the flow effects and the resulting transverse momentum spectra can be achieved within hydrodynamic model calculations. These, however, require as input the initial conditions of the reaction, a knowledge of the EOS, and a proper modeling of the freeze-out stage as well as viscous effects and are therefore technically very involved. A simpler approach that is widely used to determine the kinetic freeze-out temperature T_{kin} and the average transverse expansion velocity of the fireball $\langle\beta_{\mathrm{t}}\rangle$ by a simultaneous fit to the p_{t} spectra of several hadron species employs the so-called blast wave formula [25]

$$\frac{\mathrm{d}N}{m_{\mathrm{t}}\mathrm{d}m_{\mathrm{t}}\mathrm{d}y} \propto \int_0^{R_{\mathrm{s}}} r\mathrm{d}r m_{\mathrm{t}} I_0\left(\frac{p_{\mathrm{t}}\sinh\rho}{T_{\mathrm{kin}}}\right) \times K_1\left(\frac{m_{\mathrm{t}}\cosh\rho}{T_{\mathrm{kin}}}\right) \tag{13.3}$$

where R_{s} is the radius of the source and $\rho = \tan h^{-1}\beta_{\mathrm{t}}$ the boost angle. Figure 13.10 summarizes the results of these fits to

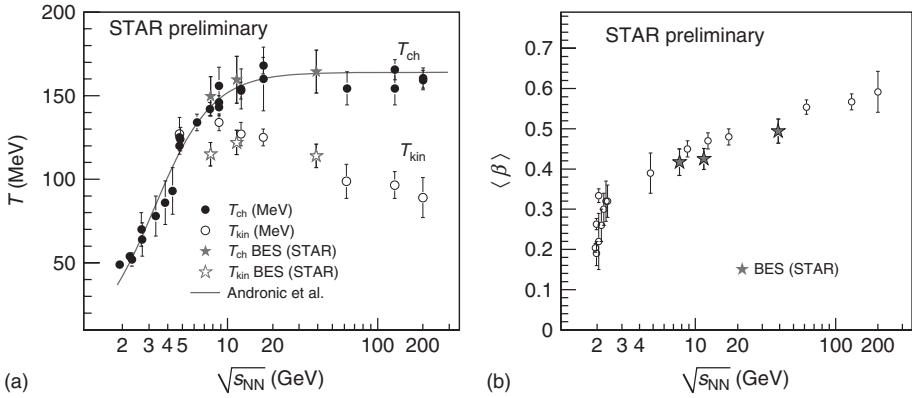


Figure 13.10 (a) The chemical T_{ch} (filled symbols) and kinetic T_{kin} (open symbols) freeze-out temperatures as a function of $\sqrt{s_{NN}}$ [23]. The solid line shows a parameterization of T_{ch} [24]. (b) The mean transverse expansion velocity $\langle\beta_t\rangle$ as a function of $\sqrt{s_{NN}}$ [23].

particle spectra at many different center-of-mass energies $\sqrt{s_{NN}}$. As Figure 13.10a shows, the values extracted for the kinetic freeze-out temperature T_{kin} are close to the ones of the chemical freeze-out T_{ch} at low $\sqrt{s_{NN}}$, but are clearly lower than these at higher $\sqrt{s_{NN}}$. While T_{ch} saturates at high energies are close to the critical temperature of the deconfinement phase transition, T_{kin} decreases toward higher $\sqrt{s_{NN}}$. This might indicate the presence of an extended hadronic phase after chemical freeze-out that allows for further cooling down of the system. The average transverse expansion velocity $\langle\beta_t\rangle$ (Figure 13.10b) exhibits a relatively steep rise with energy at low $\sqrt{s_{NN}}$, which then turns into a moderate increase at higher energies. This change of the slope around $\sqrt{s_{NN}} =$ 4–8 GeV is perhaps caused a change of the EOS due to the release of new degrees of freedom in a deconfined state of matter. Such a behavior was first seen in a study of the average transverse mass $\langle m_t\rangle$ of pion, kaons, and protons as a function of $\sqrt{s_{NN}}$ [21].

13.4.2 Elliptic Flow

An observable of special importance in heavy-ion physics is elliptic flow. It is defined as the second harmonic v_2 of the particle distribution in the azimuthal angle ϕ relative to the reaction plane orientation Ψ_R (Figure 13.11a):

$$E\frac{d^3N}{d^3p} = \frac{1}{2\pi}\frac{d^2N}{p_t dp_t dy} \times\left\{1+2\sum_{n=1}^{\infty} v_n(p_t, y)\cos[n(\phi-\Psi_R)]\right\} \tag{13.4}$$

The event plane method requires that the reaction plane orientation is determined from the data (in this case, usually referred to as the event plane, which is associated with a sizable error. Modern methods, such as cumulants, do not rely on an estimate of Ψ_R but rather use multiparticle correlations. Their advantage is a reduced sensitivity to nonflow effects, such as resonance decays [26].

The elliptic flow measured in the final particle distributions at higher center-of-mass energies ($\sqrt{s_{NN}} > 3$ GeV) is caused by the different pressure gradients present in the almond-shaped overlap zone. It is larger in the direction of the reaction plane than orthogonal to it. Thus, the particles will experience a stronger boost in-plane than out-of-plane, translating the initial anisotropy in configuration space into an anisotropy in momentum space, yielding a positive v_2. As the azimuthal difference of the pressure gradients of the system is largest in the early stage of the evolution, later stages will not contribute as much to the measured v_2. Therefore, elliptic flow is an observable that is mainly sensitive to the initial properties of the fireball.

Figure 13.11b summarizes the energy dependence of the p_t integrated v_2. While at low energies ($\sqrt{s_{NN}} < 3$ GeV), it has a complicated structure due to other physical effects (e.g., shadowing by spectators or angular momentum conservation), v_2 is positive at higher energies and rises slowly with $\sqrt{s_{NN}}$. It should be noted that this rise is mainly caused by the overall increase of the average transverse momenta and to a lesser extent by a change of the p_t dependence of $v_2(p_t)$.

A hydrodynamic calculation without viscous corrections does not provide a good description of the measurements at lower center-of-mass energies. However, at higher energies ($\sqrt{s_{NN}} = 200$ GeV), the data get very close to the ideal scenario (Figure 13.12a), indicating that the value for the shear viscosity η, divided by entropy density s, of the medium is very low. Indeed, hydrodynamic calculations that include viscous corrections find that the data can be described very well if η/s is small (see, e.g., the calculation shown in Figure 13.12b, which would suggest $\eta/s = 0.08$). This observation has dramatically changed the theoretical concepts of the QGP. Originally, it was mainly thought of as a gas of weakly interacting quarks and gluons. The extremely low η/s value of the observed matter suggests that it

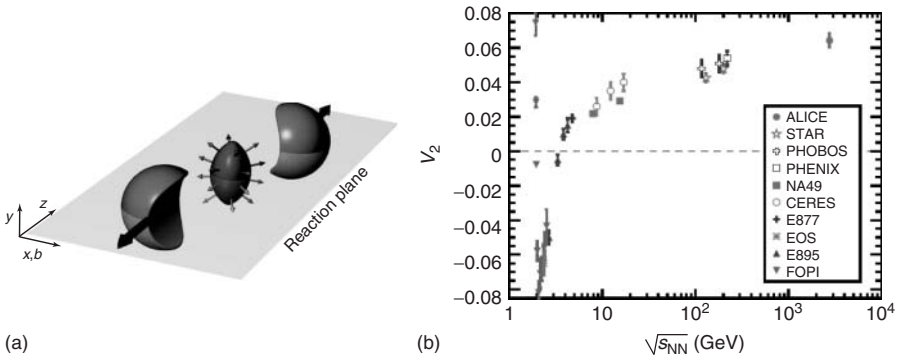


Figure 13.11 (a) A sketch of a semicentral heavy-ion collision. Shown is the almond-shaped fireball, the forward- and backward-moving spectators, and the reaction plane determined by the beam direction z and the impact parameter vector $\vec{b}$. (b) The p_t integrated elliptic flow v_2 for semicentral nucleus–nucleus collisions as a function of the center-of-mass energy [26].

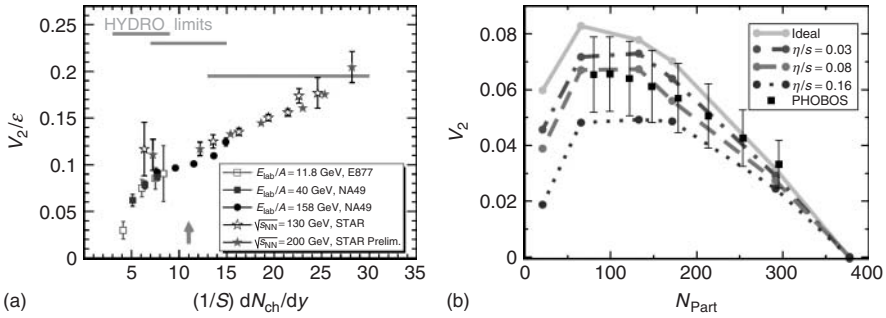


Figure 13.12 (a) The elliptic flow v_2, divided by the initial spatial eccentricity ϵ, as a function of the charged particle multiplicity density $1/S dN_{ch}/dy$, where S is the area of the overlap region [27]. Shown is a compilation of data with different centralities and center-of-mass energies. Also included are expectations from an ideal hydrodynamic model calculation, depicted as horizontal lines. (b) The centrality dependence of v_2 as measured for Au + Au collisions at $\sqrt{s_{NN}} = 200$ GeV in comparison with a hydrodynamic model calculation including viscosity for different values of η/s [28].

rather behaves like an ideal or, as it is frequently termed, ''perfect'' liquid [29]. In fact, it appears that the matter produced in high-energy heavy-ion collisions is the most ideal liquid realized in nature.

Apart from elliptic flow v_2 there are other harmonics that are of importance. Directed flow v_1 can provide information on the EOS as well and the study of its energy dependence might help to localize the deconfinement phase boundary [30, 31]. Recently, higher harmonics v_n with n up to 10 have also been measured at RHIC and LHC. Their analysis within hydrodynamic models promises to further improve our understanding of the properties of the QCD matter.

Flow studies are also performed with heavy particles, such as ϕ mesons, Ω^-, and D-mesons. Owing to their large mass and their low hadronic cross section, they, in principle, should not participate in any collective expansion if it were just driven by hadronic phases. The fact that a significant v_2 is measurable for these heavy particles also illustrates that most of the collectivity develops during a partonic phase.

13.5 Hard Probes

In the following, some examples of hard probes of a QGP are discussed. Hard probes are characterized by either a high momentum transfer in the interaction or a large mass of the involved quarks (charm or bottom quarks). At these scales, perturbative QCD can be applied to describe the corresponding processes in the vacuum, which serve as a thus well understood reference for the same observables in heavy-ion collisions. Two important ways of studying the QGP with hard probes are by measuring the medium modifications of high p_t particles (''jet quenching'') and of quarkonia.

13.5.1 Jet Quenching

In the initial phase of a heavy-ion collision, many hard-scattering processes occur and as a result the scattered partons have to traverse the hot and dense medium that is subsequently formed. Since these partons will experience an energy loss in

the medium, both via elastic scattering and gluon-bremsstrahlung induced by the free color charges in the deconfined matter, they provide a useful probe of the properties of a QGP [33]. The amount of energy loss will be proportional to the gluon density in the medium and is in some theoretical approaches characterized by a transport coefficient $\hat{q}$, which quantifies the amount of momentum transferred to the medium per unit of path length traversed by the parton (for a recent review of the theoretical concepts and experimental results see [34]).

Figure 13.13a illustrates the basic principle: in a peripheral collision, the volume of the fireball is small and both scattered partons will suffer only a moderate energy loss due to the short path length inside the medium. In a central collision, the partons have to traverse a longer path length and loose significantly more energy. If the initial hard scattering happens at the periphery of the reaction zone, it can therefore happen that one parton escapes the medium relatively unmodified, while its counterpart is stopped inside the fireball. In fact, such a scenario can already be observed in a simple event display as shown in Figure 13.13b. What is shown here is a calorimetric measurement of the energies of a 2-jet configuration in a heavy-ion event. As can easily be seen, the energy of the downward jet is not counterbalanced by the upward jet, wich can be understood within the scenario described earlier.

Jet reconstruction in heavy-ion physics, which in comparison to the pp case is complicated by the large underlying event of soft particles, is currently an important field of research and developing rapidly. However, for a quantitative comparison to theoretical calculations up to now mostly the spectra of the leading particles from jet fragmentation are used instead of fully reconstructed jets. The medium-induced energy loss will deplete the high- p_t part of the particle spectra. This is best seen when comparing the p_t-spectra measured for heavy-ions to the ones for pp collisions, scaled by the number of collisions N_{coll}, as it is quantified in the nuclear modification

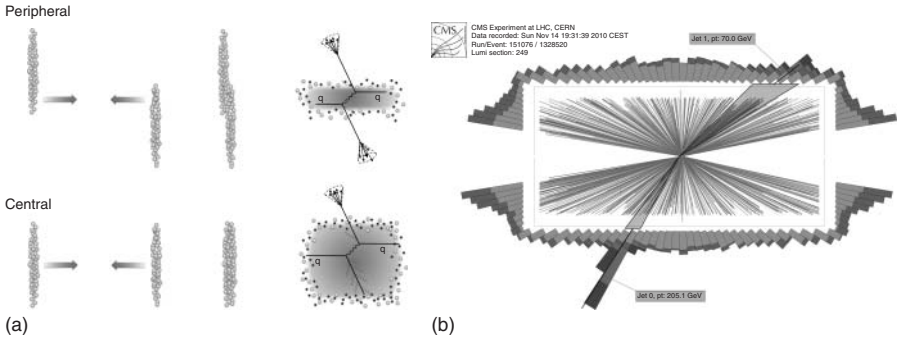


Figure 13.13 (a) Schematic view of hard-scattering processes in heavy-ion reactions. Shown is a peripheral (upper figure) and central (lower figure). (b) Event display of a Pb–Pb collisions at the LHC ($\sqrt{s_{NN}} = 2.76$ TeV) as measured by the CMS collaboration [32]. In this cross section view along the beam axis, a clear asymmetric 2-jet structure is visible.

factor R_{AA}:

$$R_{AA} = \frac{\sigma_{inel}^{pp}}{N_{coll}} \frac{d^2N^{AA}/dp_t}{d^2\sigma^{pp}/dp_t} \qquad (13.5)$$

A suppression of high-p_t particle production has for the first time been established experimentally at RHIC for neutral pions and charged hadrons [36, 37]. It was found that R_{AA} is significantly below one for $p_t > 2 - 3$ GeV/c. On the other hand, for high p_t photons, that do not interact strongly, $R_{AA} \approx 1$ was measured, indicating that the scaling approach with the number of binary collisions in fact is valid and that the suppression is due to strong interaction in the medium. A measurement of R_{AA} at the LHC is shown in Figure 13.14. While for peripheral collisions only a weak suppression is observed, in central collisions a R_{AA} much smaller than unity is measured, implying a strong energy loss of high p_t particles. The magnitude and the complex shape of the p_t dependence of R_{AA} allows to discriminate between various model approaches and to extract quantitative results on the medium properties.

Additional information is obtained by performing R_{AA} measurements of identified particles. Of particular interest in this context are hadrons with open charm (e.g., D-mesons), due to their very high mass and the possible reduction of their energy loss due to the deadcone effect. First measurements at the LHC indicate that, contrary to first expectations, also charmed hadrons experience an energy loss on the same level than lighter particles, such as pions and protons.

13.5.2 Quarkonia

Another important hard probe of a QGP are quarkonia. These are bound states of heavy quark–antiquark pairs (e.g., the J/ψ

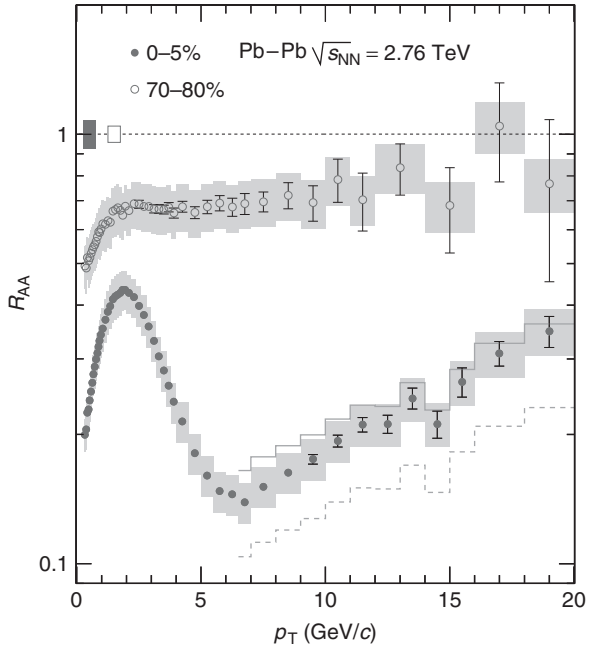


Figure 13.14 The nuclear modification factor R_{AA} of unidentified charged particles for peripheral (70–80%, open symbols) and central (0–5%, filled symbols) Pb–Pb reactions at the LHC ($\sqrt{s_{NN}} = 2.76$TeV) [35].

as a bound $c\bar{c}$ pair or the Υ as a bound $b\bar{b}$ pair). The free color charges in a QGP are expected to reduce the attractive potential between the $q\bar{q}$ pairs, similar to Debye screening in an electromagnetic plasma. This temperature-dependent effect should thus reduce the number of quarkonia that can survive in the medium produced in heavy-ion collisions and a suppression of quarkonia production relative to N_{coll} scaled pp yields should be the consequence. J/ψ suppression has been suggested as a signature for a QGP formation already very early [39].

Figure 13.15 schematically displays the potential fate of $c\bar{c}$ pairs that are produced in the initial hard-scattering processes in a heavy-ion collision. Owing to the screening of the $c\bar{c}$ potential, the initially correlated pairs are dissociated (a), will subsequently diffuse through the QGP (b), and finally hadronize in mesons with open charm (c). As has been pointed out more recently at high center-of-mass energies (e.g. at LHC), the number of initially produced $c\bar{c}$ pairs can be high enough that J/ψ can be formed via the coalescence of previously uncorrelated $c\bar{c}$ pairs [40, 41]. This effect could to some extent counteract the suppression due to dissociation.

On top of the effects induced by a deconfined medium, cold nuclear matter (CNM) effects also need to be considered. These include all modifications to quarkonia production rates that are already present in normal nuclear matter, such as nuclear modifications of the parton distribution functions and absorption of quarkonia by surrounding nucleons. Experimentally, they can be addressed by studying p–A collisions.

A significant suppression of J/ψ relative to CNM effects, as measured via p-Pb collisions, has been seen experimentally at the SPS already [43], even though the effect turned out to be smaller than initially thought after a more recent assessment of the p-Pb reference and nuclear shadowing is taken into account [44]. The nuclear modification factor of J/ψ measured at RHIC was found to be of the same magnitude as the one seen at the SPS [45]. This is remarkable, since the energy density achieved in heavy-ion collisions at RHIC is considerably higher and thus the J/ψ suppression should also be stronger. A significant difference of the R_{AA} for J/ψ in comparison to lower energies is observed only when going to LHC energy [42]. As shown in Figure 13.16, the R_{AA} measured at forward

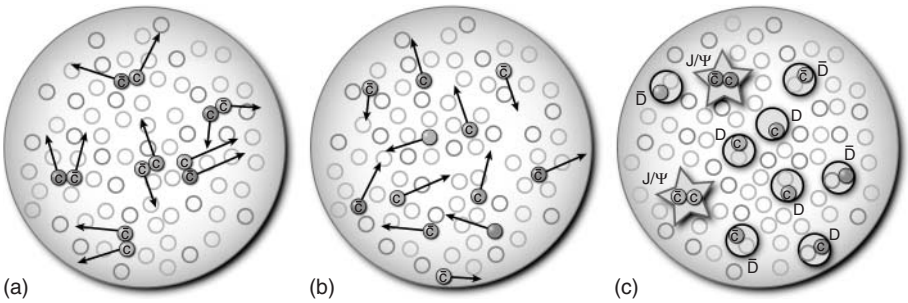


Figure 13.15 Sketch of the possible effects of a QGP on correlated $c\bar{c}$ pairs: (a) dissociation due to color screening, (b) diffusion in the medium, and (c) formation of J/ψ by quark coalescence. (Courtesy of F. Kramer, Goethe-University of Frankfurt.)

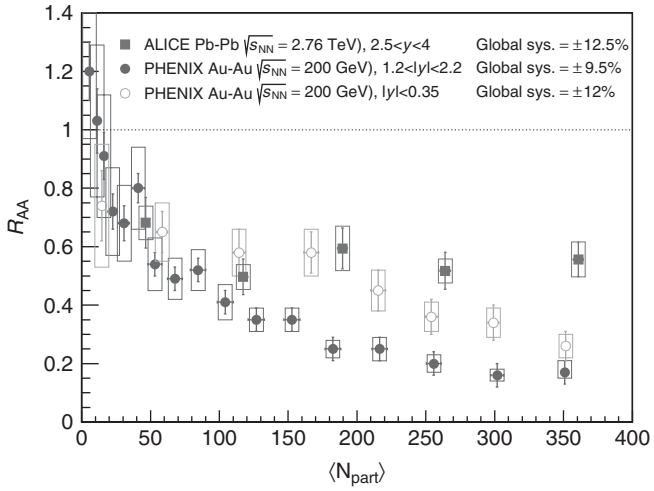


Figure 13.16 The nuclear suppression factor R_{AA} for inclusive J/ψ as a function of the number of participants N_{part}, measured at RHIC ($\sqrt{s_{NN}} = 200$ GeV) and at the LHC ($\sqrt{s_{NN}} = 2.76$ TeV) [42].

rapidities is clearly higher than the one measured at RHIC. This might indicate that at these energies new mechanisms, such as the suggested quark coalescence, contribute to the J/ψ yield in heavy-ion collisions. Further information about the mechanisms that modify the J/ψ production in heavy-ion collisions will come from the investigation of the J/ψ elliptic flow (v_2). This can be nonzero if there is a contribution from quark coalescence.

Because for the excited states of quarkonia the average distance between the bound (anti-)quarks is larger than for the ground states, they will also be more easily dissolved in a QGP. Therefore, the comparison of different quarkonia states, in particular the Υ states, will offer additional information on the properties of the medium and might even serve as a "QGP thermometer." At the LHC, the suppression of Υ was measured recently [46]. As can already be seen from a comparison of the invariant mass spectra as measured in pp and Pb–Pb (Figure 13.17), the excited states [$\Upsilon(2S)$ and $\Upsilon(3S)$] are in fact significantly more suppressed than in the ground state [$\Upsilon(1S)$]. Future investigations in this direction will certainly also provide more quantitative insights into the properties of the hot and dense medium.

13.6 Outlook

The field of relativistic heavy-ion physics is currently evolving in two directions, the high-energy and the high-density frontier. The first is mainly driven by the experimental program at the LHC, whose high center-of-mass energy (up to $\sqrt{s_{NN}} = 5.5$ TeV) and luminosity allows to access many rare probes in an unprecedented manner. Especially promising is the possibility to measure γ-jet correlations. Here hard-scattering processes are exploited in which one outgoing particle is a parton

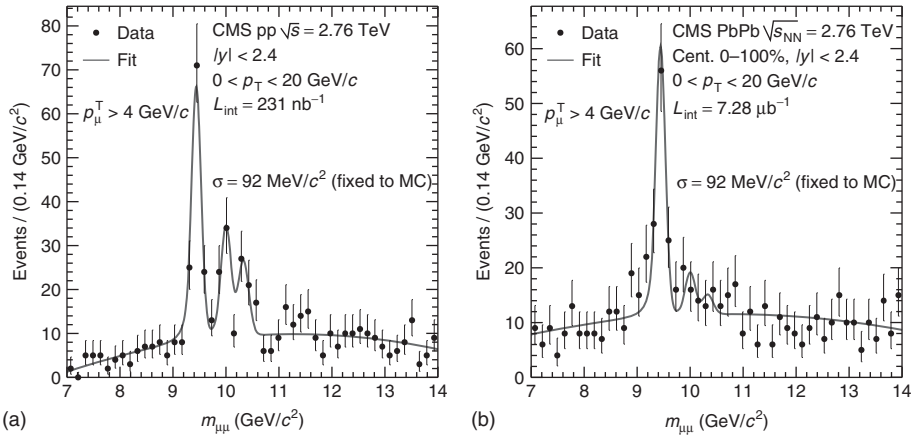


Figure 13.17 The invariant mass distributions of $\mu^+\mu^-$ pairs measured in (a) pp and (b) Pb–Pb collisions at the LHC ($\sqrt{s_{NN}} = 2.76 TeV$) [46].

and the other one a γ. As the γ does not interact strongly with the matter, it can provide unmodified information on the original energy of the scattered parton and therefore will allow to measure its energy loss on an absolute scale. In addition, more results on the high p_t suppression and the elliptic flow of heavy-flavor particles (e.g., D-mesons) will provide further insight into the properties of the QGP. More information on the bulk properties, such as higher harmonics of collective flow will be collected. These will help to understand also the initial state of the collisions in greater detail. The question whether it behaves like a CGC can then perhaps be addressed in an unambiguous way. The LHC program of the coming years thus promises to deliver many more exciting results and experiments are currently being prepared for the measurement campaigns at high luminosities.

At the high-density frontier, the current programs at the CERN-SPS and BNL-RHIC, as well as the planned experiments CBM at FAIR and MPD at NICA, will try to gather information on the nature of the phase boundary (first-order versus crossover) in different regions of the QCD phase diagram and will search for a critical point. In order to achieve the high net-baryon densities, corresponding to the high baryonic chemical potentials that are needed to probe the relevant areas of the phase diagram, these experiments measure at rather low center-of-mass energies ($2 < \sqrt{s_{NN}} < 20$ GeV). The focus of all these programs is a systematic study of the $\sqrt{s_{NN}}$ dependencies of many observables. For the critical point search event-by-event fluctuation observables are relevant. Here it seems to be promising to measure fluctuations of conserved quantities, such as the net-baryon number, which have a higher probability to remain unchanged by the late hadronization stage. Also, higher moments of event-by-event observables should be more sensitive to divergencies caused by the vicinity of a critical point. This subject is currently undergoing rapid theoretical and experimental development and promises many important discoveries for the future.

Glossary

Baryons: Hadrons containing three valence quarks, such as protons, neutrons, and hyperons, with half-integer spin (i.e., fermions).

Baryonic chemical potential: Parameter in the grand canonical partition function used in statistical model approaches that is sensitive to the net-baryon number in the reaction zone.

Chiral symmetry: Basic symmetry of QCD that is strictly valid only in the limit of massless quarks. It is, in addition, spontaneously broken in the QCD vacuum, which results in the known hadron masses.

Collective flow: Denotes the observation that soft-particle emission in heavy-ion collisions cannot be understood by a simple superposition of independent nucleon–nucleon collisions, but rather exhibits anisotropies that are due to the pressure-driven expansion of the reaction system. Flow phenomena are theoretically addressed within hydrodynamic models.

Color glass condensate (CGC): Theory on the initial state of a heavy-ion collision, which is based on the saturation of the gluon fields expected at very high energies.

Confinement: Property of QCD due to which partons cannot exist as free particles in vacuum but only bound within hadrons. In the hot and dense matter generated by heavy-ion collisions, this confinement should be no longer effective (deconfinement) and quarks and gluons could exist as quasi-free particles within a QGP.

Critical point: Position in a phase diagram where a first-order phase transition line ends. At a critical point, susceptibilities diverge and the phenomenon of critical opalescence can be observed.

Equation of state (EOS): Describes the connection between pressure and temperature of a given medium and depends on the microscopic number of degrees of freedom. The EOS of strongly interacting matter should change at the deconfinement phase transition, due to the much larger number of degrees of freedom of a QGP.

Fragmentation: Hadronization mechanism in which hadrons are generated by the breaking of a so-called string (i.e., color flux tube) spanned between leading partons.

Glauber model: Model, originally invented by Roy Glauber, that treats the interaction of two nuclei as a superposition of single nucleon–nucleon collisions, which move on straight-line trajectories. It is generally used to calculate the number of participants and spectators, as well as the number of binary nucleon–nucleon collisions.

Hadrons: Elementary particles that interact via the strong interaction. Hadrons are further divided into baryons and mesons.

Hadronization: Process in which free partons are converted into hadrons. Possible hadronization mechanisms are fragmentation or quark coalescence.

Hydrodynamic models: Class of theoretical models that describes the matter generated in high-energy nucleus–nucleus collisions as a relativistic fluid. The EOS is an important ingredient for these models. While in the ideal case, viscous corrections are ignored, the state-of-the-art versions of these models include the effects of bulk and shear viscosity.

Hyperons: Group of baryons that contain one or more strange quarks, for example, Λ, Ξ^-, or Ω^-.

Jets: Collimated bundle of hadrons that originates from the fragmentation of partons which have undergone a scattering process involving a large momentum transfer (so-called hard scattering).

Lattice-QCD: Theoretical approach in which QCD problems are solved by Monte Carlo methods on a discrete space–time lattice. It also allows to treat soft processes (i.e., small momentum transfers), relevant for confinement and the properties of the QCD phase diagram, which are not accessible to perturbative methods.

Mesons: Hadrons formed by a bound quark–antiquark pair, such as pions or quarkonia. They have an integer spin and thus are bosons.

Participants: Nucleons that fall within the overlap area of the two colliding nuclei and thus participate in the reaction. Their number depends on the impact parameter of the collision and is usually determined via the Glauber model from observables such as the number of produced charged particles.

Partons: Elementary constituents of hadrons, namely, quarks and gluons.

Pseudorapidity: Defined as $\eta = 1/2 \ln\left((|\vec{p}| + p_z)/(|\vec{p}| - p_z)\right)$ and used to describe longitudinal (i.e. along the beam axis z) particle distributions in the case when particle identification is not performed. For identified particles the rapidity is used.

Quantum chromodynamics (QCD): Field theory describing the strong interaction. It is a non-Abelian gauge theory based on the SU(3) group. The corresponding eight field quanta are called gluons and couple to the color charges carried by quarks and the gluons themselves.

Quark coalescence: Hadronization mechanism in which hadrons are formed by grouping valence quarks into hadrons. It requires the presence of a reservoir of free quarks as realized in a QGP.

Quark–gluon plasma (QGP): State of strongly interacting matter in which confinement is not effective and quarks and gluons can exist as quasi-free particles.

Quarkonia: Bound states of heavy (charm or beauty) quark–antiquark pairs. The most prominent ones are the vector mesons J/ψ ($c\bar{c}$) and Υ ($b\bar{b}$).

Quarkyonic matter: A hypothetical state of QCD matter in which hadrons are still confined, but chiral symmetry is restored.

Rapidity: Defined as $y = 1/2 \ln\left[(E + p_z)/(E - p_z)\right]$ and used to describe longitudinal (i.e. along the beam axis z) particle distributions. The rapidity is constructed such that the shape of y spectra is Lorentz invariant, in contrast to the case of pseudorapidity distributions.

Shear viscosity: Fluid property that describes how effectively momentum is transferred between adjacent fluid layers. Even though the QGP matter has a very high shear viscosity, the ratio of shear viscosity to entropy density is found to be extremely small, thus allowing its description as an ideal fluid.

Spectators: Nucleons that fall outside the overlap area of the two colliding nuclei and thus do not participate in the reaction. The number of projectile spectators can be determined experimentally by measuring the remaining beam energy.

Strangeness: Quantum number of hadrons, indicating the number of strange valence quarks ($S = -1$).

References

1. Borsányi, S. *et al.* (2010) *JHEP*, **11**, 077.
2. (a) McLarren, L. and Venugopalan R., (1994) *Phys. Rev.*, **D49**, 2233; (b) Iancu, E., Leonidov, A., and McLarren, L. (2001) *Nucl. Phys.*, **A692**, 583.
3. Bass, S.A., Müller B., and Srivastava D.K., (2003) *Phys. Lett.*, **B551**, 277.
4. Miller, M.L., Reygers, K., Sanders, S.J., and Steinberg, P. (2007) *Ann. Rev. Nucl. Part. Sci.*, **57**, 205.
5. Glauber, R.J. (1955) *Phys. Rev.*, **100**, 242.
6. Bialas, A., Bleszyński, M., and Czyź, W. (1976) *Nucl. Phys.*, **B 111**, 461.
7. Bearden, I.G. *et al.*, (BRAHMS Collaboration) (2004) *Phys. Rev. Lett.*, **93**, 102301.
8. Aamodt, K. *et al.*, (ALICE Collaboration) (2010) *Phys. Rev. Lett.*, **105**, 252301.
9. Blume, C. *et al.*, (for the NA49 Collaboration) (2007) *J. Phys.*, **G 34**, S951.
10. Bjørken, J.D. (1983) *Phys. Rev.*, **D 27**, 140.
11. Karsch, F. (2002) *Nucl. Phys.*, **A 698**, 199.
12. Andronic, A., Braun-Munzinger, P., and Stachel, J. (2009) *Phys. Lett.*, **B 673**, 142.
13. Cleymans, J., Oeschler, H., Redlich, K., and Wheaton, S. (2006) *J. Phys.*, **G 32**, S165.
14. Becattini, F. and Heinz, U. (1997) *Z. Phys.*, **C 76**, 269.
15. Stock, R. (1999) *Phys. Lett.*, **B 456**, 277.
16. Andronic, A. *et al.* (2010) *Nucl. Phys.*, **A 837**, 65.
17. McLarren, L. and Pisarski, R. (2007) *Nucl. Phys.*, **A 796**, 83.
18. Abelev, B.I. *et al.*, (STAR Collaboration) (2008) *Phys. Rev.*, **C 77**, 044908.
19. Rafelski, J. and Müller, B. (1982) *Phys. Rev. Lett.*, **48**, 1066.
20. Blume, C. and Markert, C. (2011) *Prog. Part. Nucl. Phys.*, **66**, 834.
21. Alt, C. *et al.*, (NA49 Collaboration) (2008) *Phys. Rev.*, **C 77**, 024903.
22. Gaździcki, M. and Gorenstein, M.I. (1999) *Acta Phys. Polon.*, **B 30**, 2705.
23. Kumar, L. *et al.*, (for the STAR Collaboration) (2011) *J. Phys.*, **G38**, 124145.
24. Andronic, A., Braun-Munzinger, P., and Stachel, J. (2006) *Nucl. Phys.*, **A772**, 167.
25. Schnedermann, E., Sollfrank, J., and Heinz, U. (1993) *Phys. Rev.*, **C 48**, 2462.
26. Snellings, R. (2011) *New J. Phys.*, **13**, 055008.
27. Alt, C. *et al.*, (NA49 Collaboration) (2003) *Phys. Rev.*, **C 68**, 034903.
28. Romatschke, P. and Romatschke, U. (2007) *Phys. Rev. Lett.*, **99**, 172301.
29. (a) Arsene, I. *et al.*, (BRAHMS Collaboration) (2005) *Nucl. Phys.*, **A 757**, 1; (b) Back, B.B. *et al.*, (PHOBOS Collaboration) (2005) *Nucl. Phys.*, **A 757**, 28; (c) Adams, J. *et al.*, (STAR Collaboration) (2005) *Nucl. Phys.*, **A 757**, 102; (d) Adcox, K. *et al.*, (PHENIX Collaboration) (2005) *Nucl. Phys.*, **A 757**, 184.
30. Stöcker, H. and Greiner, W. (1986) *Phys. Rep.*, **147**, 277.
31. Rischke, D.H. *et al.* (1995) *Heavy Ion Phys.*, **1**, 309.
32. CMS Collaboration, *http://cdsweb.cern.ch/record/1309898*.
33. Wang, X.-N. and Gyulassy, M. (1992) *Phys. Rev. Lett.*, **68**, 1480.
34. Majumder, A. and van Leeuwen, M. (2011) *Prog. Part. Nucl. Phys.*, **66**, 41.
35. Aamodt, K. *et al.*, (ALICE Collaboration) (2011) *Phys. Lett.*, **B 696**, 30.
36. Adcox, K. *et al.*, (PHENIX Collaboration) (2002) *Phys. Rev. Lett.*, **88**, 022301.
37. Adler, C. *et al.*, (STAR Collaboration) (2002) *Phys. Rev. Lett.*, **89**, 202301.
38. Kramer, F. (2012) PhD Thesis, Goethe-University of Frankfurt.
39. Matsui, T. and Satz, H. (1986) *Phys. Lett.*, **B 178**, 416.
40. Thews, R.L., Schroedter, M., and Rafelski, J. (2001) *Phys. Rev.*, **C 63**, 054905.
41. Braun-Munzinger, P. and Stachel, J. (2001) *Nucl. Phys.*, **A 690**, 119.
42. Abelev, B. *et al.* (ALICE Collaboration), (2012) *Phys. Rev. Lett.*, **109**, 072301.
43. Alessandro, B. *et al.*, (NA50 Collaboration) (2005) *Eur. Phys. J.*, **C 39**, 335.
44. Arnaldi, R. *et al.*, (for the NA60 Collaboration) (2009) *Nucl. Phys.*, **A 830**, 345c.
45. Adare, A. *et al.*, (PHENIX Collaboration) (2007) *Phys. Rev. Lett.*, **98**, 232301.
46. Chatrchyan, S. *et al.*, (CMS Collaboration) (2012) *JHEP*, **05**, 063.

Part B
Applied Nuclear Physics

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

14
Neutron Stars

Jürgen Schaffner-Bielich

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

14.1
Introduction: Supernovae and Neutron Stars

Neutron stars are born in the aftermath of a cataclysmic cosmic event, in a core-collapse supernova. Historically, supernovae are classified spectroscopically by the measurement of hydrogen, helium, and silicon lines. A supernova of type I has no dominant hydrogen line in its spectrum, while a supernova of type II shows a significant presence of hydrogen. The present interpretation of the different types of supernovae is grossly that supernovae of type I emerge from binary stars, where the exploding star has largely lost its hydrogen envelope in the evolution of the binary system. Supernovae of type II are the ''classic'' textbook cases, where a massive star has reached the final point in its stellar evolution. The degenerate core collapses as its degeneracy pressure cannot withstand the overwhelming attractive gravitational pull. This trigger mechanism for a supernova defines the class of a so-called core-collapse supernova. The current understanding is that not only supernovae of type II but also some of type I, the types Ib and Ic, belong to the class of core-collapse supernovae. The progenitor stars of supernovae of type Ib and Ic are supposed to be massive stars in binary systems, where the companion has stripped off their hydrogen envelope (type Ib) or also their helium envelope (type Ic), so that corresponding lines are not visible in the supernova spectra. Contrary to these core-collapse supernovae, supernovae of type Ia show strong silicon lines in their spectra. The underlying explosion mechanism of supernovae of type Ia is a completely different one and is associated with the presence of white dwarfs. The common explanation is that the accretion from a companion star causes a chain of nuclear runaway reactions in the white dwarf leading to a complete dissolution of the white dwarf. This explanation has been questioned, recently, as it has been realized in detailed numerical simulations that there is another plausible mechanism for type Ia supernova by the merger of two white dwarfs. The sum of initial masses is above the mass limit of white dwarfs, the Chandrasekhar mass limit of about 1.4 solar masses, so that the final outcome of the white dwarf merger is a black hole. So in both cases discussed so far in the literature, there is no compact remnant formed in type Ia supernova, contrary to the case of core-collapse supernovae.

The detailed outline of the supernova classification scheme and the underlying explosion mechanism might be considered to be baroque by the reader, but points to crucial differences in the evolution of

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

neutron stars, the compact remnants of core-collapse supernovae.

First, there should be at least two different classes of neutron stars, the ones formed in type II supernovae with an isolated progenitor star, and the ones formed in type Ib and Ic, that is, in progenitor stars with a companion. So we expect that there should be isolated neutron stars and, if the binary system is not disrupted by the supernova explosion, also neutron stars with a companion. Binary-evolution studies indeed suggests that a neutron star can be formed in a binary system with a white dwarf, a neutron star, or even a black hole. While the former two kinds of system have been found, the latter one is considered to be the ''holy grail'' of relativistic astrophysics as it allows to map out the gravitational field around a black hole by studying the orbiting neutron star. Future large-scale radio surveys as with the square kilometre array SKA envision to find such an elusive relativistic system.

Second, neutron stars formed in a binary system can accrete matter from the companion star and can spin up by the transfer of angular momentum from the accreting disk. Rotation-powered neutron stars in binary systems are found to be the oldest known pulsars, where the age is estimated from the characteristic age. Among them, we find the ones with the fastest known rotation period, the so-called millisecond pulsars. The general scheme, that the older pulsar should be spinning faster than the younger one, has been confirmed with the discovery of the double pulsar (PSR J0737-3039), the only system where the radio pulses of both neutron stars in the binary system have been detected. Indeed, pulsar A is a millisecond pulsar with a period of 23 ms and an age of 250 Myr, while the companion, pulsar B, has a period of 2.8 s with an age of only 50 Myr and is obviously not a recycled pulsar.

Third, accreting neutron stars will increase their mass depending on the companion star and the length of the accretion period of the binary system. The natural expectation is that in binary systems the pulsars with the longest accretion period should exhibit the highest masses. The evolution timescale for the formation of a white dwarf is much larger than that for a neutron stars so that more-massive neutron stars are expected for those having a white dwarf companion. In fact, the most-massive, known neutron star as of this writing is the pulsar PSR 1614-2230, which was discovered in a binary system with a white dwarf in 2010 [1]. With a mass of $M_{\odot} = (1.97 \pm 0.04) \times M_{\odot}$, it is considerably more massive compared to previous pulsar mass measurements of double-neutron-star systems, which cluster around 1.4 $M_{\odot}$. The new neutron star mass limit provides strong constraints on the properties of neutron star matter, in particular, on the nuclear equations of state at high densities, and is therefore of enormous interest to the nuclear physics community.

In addition, isolated neutron stars are also particular interesting in determining the properties of neutron stars. Isolated neutron stars are particular clean probes for direct thermal emission from their surfaces, which are unspoiled by the presence of an accretion disk. Observations with X-ray satellites were able to measure the spectra of certain close-by isolated neutron stars, the so-called magnificent seven. The surprisingly featureless blackbody spectrum is still puzzling and points to physics on the surface of the neutron stars that is not well understood. One explanation is that the high magnetic field transforms the atmosphere to a metal, so that no spectral lines can be observed.

Fits to the overall spectrum of the isolated neutron star RX J1856-3754 points to a large radius and a correspondingly high mass. The determination of the radius of a neutron star is also of utmost importance for exploring the nuclear equation of state at supranuclear densities.

From simulations studies, one infers that core-collapse supernovae have progenitor masses of more than eight solar masses, which is the dividing mass scale for producing a white dwarf or a neutron star as the final endpoint in stellar evolution. The crucial criterion is the onset of carbon burning in the core to produce a degenerate core for stars more massive than eight solar masses. Note, that the initial mass at stable hydrogen burning on the main sequence, the so-called zero-age main sequence mass, is quoted here and not the mass shortly before the collapse, which can be considerably different owing to mass loss by stellar winds or accretion. Supernova simulations indicate that progenitor masses of more than about 25 solar masses will cause the proto-neutron star to collapse to a black hole. It has been realized in recent years that rotation is also a deciding factor for the final outcome of core-collapse supernovae, as well as the initial data from the progenitor, an input from stellar evolution calculations. Core-collapse supernovae with progenitor masses of about 8–10 solar masses start with a degenerate oxygen–neon–magnesium core. The common lore that core-collapse supernovae have a degenerate iron core only applies to progenitor masses of more than about 10 solar masses. A large fraction of observed neutron stars originates therefore from the collapse of a degenerate O–Ne–Mg core while there is another sample of, possibly more massive, neutron stars coming from textbook iron core-collapse supernovae.

The connection between the nuclear properties of neutron stars and their astrophysical observations are twofold. First, the bulk properties of the nuclear equation of state determine the global features of a neutron star, that is, the mass–radius relation. There is a one-to-one relation between the nuclear equation of state and the mass–radius relation. The determining factor for the mass and the radius of a neutron star is the high-density behavior of the nuclear force; in more general terms, the properties of strong interactions, quantum chromodynamics, under extreme baryon densities. This feature is not present for white dwarfs as their global properties are fixed mostly by the degeneracy pressure of electrons. There are small corrections to the mass–radius relation from the composition of the white dwarf, which could harbor a degenerate helium or carbon core, but iron cores are also discussed as a possibility in the literature. Contrary to neutron stars, white dwarfs with different compositions can coexist in nature as matter in white dwarfs is not in nuclear equilibrium. Nuclear reactions are suppressed by Coulomb repulsion at the usual densities and temperatures encountered in white dwarfs. On the other hand, weak and strong interactions are in equilibrium in bulk neutron star matter, so that there should be only one unique composition of a neutron star realized in nature, meaning that there is only one unique mass–radius relation solely determined by the bulk properties, that is, the nuclear equation of state. We need to add here that the transition from ordinary nuclear matter to more exotic forms of dense matter could be happening on such large timescales that it would allow for the existence of another second class of compact stars. One example is that of so-called strange stars made of

absolutely stable strange quark matter. It is a speculative scenario but still discussed in the community as it cannot be ruled out at present, either from observations or from our poor understanding of quantum chromodynamics (QCD) at high baryon densities. The more refined statement would be that there is only one unique branch of ordinary neutron stars, while there might exist another one associated with a new phase at high densities.

Second, details in the composition and the overall structure of a neutron star have an impact on transport properties such as their cooling and rotational mode (r-mode) instabilities. The late-time cooling of neutron stars is controlled by the emission of neutrinos, as they can escape freely from the whole neutron star. Their interaction cross section is low even for the densities encountered in the core of (cold) neutron stars. The emissivities of neutrinos depends on the nuclear properties of the medium, in particular, on the pairing energy of baryons in the medium. The r-mode instability leads to the emission of gravitational waves and is itself determined by the viscosity of the dense nuclear medium. Therefore, in addition to using electromagnetic probes, the observation of cooling neutron stars, as well as gravitational waves, can also yield information about the properties of dense matter in compact stars. In principle, the direct observation of neutrinos from a galactic core-collapse supernovae could also be utilized to study the nuclear equation of state and the properties of hot proto-neutron stars, within the first minute after collapse. The advantage is that neutrinos and gravitons can propagate freely from the core of the compact star, making them ideal tools for probing nuclear matter at extreme densities. Present and future detectors, such as the neutrino detectors IceCube and Super-K, and the new generation of gravitational wave detectors, such as the advanced large interferometer gravitational-wave observatory (LIGO), will open new windows for the exploration of neutron stars. The prospect of learning more about nuclear physics at supranuclear densities by the observation of neutron stars either by their global properties or by multimessenger astronomy of their transport properties is striking.

In the following, the structure of a neutron star is outlined starting from the surface to the inner core. The focus will be on the nuclear physics aspects of cold neutron stars.

14.2 The Outer Layers of a Neutron Star

The outermost shells of a neutron star can be divided into three different domains. The skin of the neutron star is the atmosphere, a gas of atoms that is present up to densities of about 10 $\mathrm{g\,cm^{-3}}$. The composition of the atmosphere could be determined by the fallback from a core-collapse supernova or by the accretion from a companion star so that it would be hydrogen dominated. The enormous gravitational field present at the surface of a neutron star, which is of the order of 2×10^{11} times the field strength compared to the one on earth, crushes the atmospheric layer down to a thickness of just a few centimeters. As the mean free path of photons is considerably smaller, the spectrum of neutron stars is determined by the properties of the atmosphere and cannot provide a direct probe of the shells below. Observations of neutron stars with optical telescopes and X-ray satellites provide, however, important constraints on their mass and radius, by measuring the total emitting surface area.

The other two outermost shells are commonly denoted as the outer and the inner crusts. The outer crust ends and the inner crust begins at the so-called neutron-drip density, which is about $(4–5) \times 10^{11}$ g cm^{-3}. The inner crust terminates around a density of about one-half of the saturation density, corresponding to a mass density of 1×10^{14} g cm^{-3}. The density range of the whole neutron star crust spans a colossal 14 orders of magnitude, but the immense gravity shrinks the crust to just 1 km in total for typical neutron star masses of around 1.4 $M_\odot$. An extensive review on the physics of the neutron star crust can be found in [2].

14.2.1 The Outer Crust: A Lattice of Nuclei

The outer crust starts when the electrons become unbound and build a degenerate Fermi gas. For cold, nonaccreting neutron stars, the nuclei form a lattice, which is immersed in the background electron gas, ensuring overall charge neutrality of the matter. For accreting neutron stars, an intermediate layer appears as the lattice melts from the heat generated by the accreting matter. As the electron pressure increases, it becomes more and more favorable to form neutron-rich nuclei. In nuclear equilibrium, the nucleus with the lowest total energy per baryon will be present for a given baryon density. At the lowest baryon densities, the nucleus ^{56}Fe will be populated, which is the most stable nucleus with the lowest energy per baryon. Note, that ^{58}Fe and ^{62}Ni are more tightly bound than ^{56}Fe, that is, they have the highest binding energy per nucleon known. However, the higher number of neutrons results in a higher energy per nucleon in total. At a mass density of about 10^6 g cm^{-3}, the electrons become relativistic. At a slightly larger mass density of about 10^7 g cm^{-3}, nuclei other than ^{56}Fe begin to be populated as more neutron-rich nuclei are favored in the background of the relativistic electron gas. The charge-to-mass ratio of the nuclei present in the matter will decrease with increasing density. For a given density, the most stable nucleus with a given charge-to-mass ratio will appear in the dense medium, provided that β-equilibrium has been established, that is, that all nuclear reactions are in equilibrium. The sequence of nuclei following ^{56}Fe will be dictated by the relative binding energy of nuclei. Up to a density of about 6×10^{10} g cm^{-3} the masses of nuclei are experimentally known, so the sequence of nuclei is fixed and given by ^{62}Ni, ^{64}Ni, ^{66}Ni, ^{86}Kr, ^{84}Se, ^{82}Ge, and ^{80}Zn. It is noteworthy that these nuclei have magic numbers. The nickel sequence with the magic proton number 28 is followed by a series of nuclei with the magic neutron number 50 (Table 14.1).

The sequence beyond the last experimentally known nucleus can then only be addressed with nuclear models. Recent calculations with several nuclear models, such as the finite range droplet model, Skyrme–Hartree–Fock, and the relativistic mean-field model, with modern parameterizations, indicate that the sequence of nuclei continues along the next neutron magic number 82, that is, along the nuclei ^{124}Mo, ^{122}Zr, and ^{120}Sr, likely ending at ^{118}Kr, the last nucleus before the neutron dripline. The final nucleus is as uncertain as the knowledge of the location of the neutron dripline in the nuclear chart. The corresponding neutron-drip density in the neutron star marks the end of the outer crust and the beginning of the inner crust. Model predictions range between a neutron-drip density of $(4–5) \times 10^{11}$ g cm^{-3}. The sequence of

Table 14.1 Sequence of nuclei in the outer crust of nonaccreting cold neutron stars, calculated by using the experimental nuclear data from the atomic mass table (upper part) and the theoretical mass table of the relativistic nuclear model parameter set TMA (lower part). The last line corresponds to the neutron-drip point. Table taken from [3]. Copyright 2006 by the American Physical Society.

μ [MeV]	μ_e [MeV]	ρ_{max} [g cm^{-3}]	P [dyne cm^{-2}]	n_b [cm^{-3}]	Element	Z	N
930.60	0.95	8.02×10^{6}	5.22×10^{23}	4.83×10^{30}	^{56}Fe	26	30
931.32	2.61	2.71×10^{8}	6.98×10^{25}	1.63×10^{32}	^{62}Ni	28	34
932.04	4.34	1.33×10^{9}	5.72×10^{26}	8.03×10^{32}	^{64}Ni	28	36
932.09	4.46	1.50×10^{9}	6.44×10^{26}	9.04×10^{32}	^{66}Ni	28	38
932.56	5.64	3.09×10^{9}	1.65×10^{27}	1.86×10^{33}	^{86}Kr	36	50
933.62	8.38	1.06×10^{10}	8.19×10^{27}	6.37×10^{33}	^{84}Se	34	50
934.75	11.43	2.79×10^{10}	2.85×10^{28}	1.68×10^{34}	^{82}Ge	32	50
935.93	14.71	6.21×10^{10}	7.86×10^{28}	3.73×10^{34}	^{80}Zn	30	50
937.28	18.64	1.32×10^{11}	2.03×10^{29}	7.92×10^{34}	^{78}Ni	28	50
937.63	19.80	1.68×10^{11}	2.55×10^{29}	1.01×10^{35}	^{124}Mo	42	82
938.13	21.38	2.18×10^{11}	3.48×10^{29}	1.31×10^{35}	^{122}Zr	40	82
938.67	23.19	2.89×10^{11}	4.82×10^{29}	1.73×10^{35}	^{120}Sr	38	82
939.18	24.94	3.73×10^{11}	6.47×10^{29}	2.23×10^{35}	^{118}Kr	36	82
939.57	26.29	4.55×10^{11}	8.00×10^{29}	2.72×10^{35}	^{116}Se	34	82

nuclei found in these modern nuclear models is surprisingly close to the first calculation done by Baym, Pethick, and Sutherland in their classic work from 1971 [4] (Figure 14.1).

Some models predict the presence of other nuclei between those two sequences with magic neutron numbers such as ^{82}Zn. Neutron-rich nuclei close to ^{80}Ni are in experimental reach and some of them have been studied recently. At ISOLTRAP, ^{82}Zn nuclei have been observed in 2012 and the mass measurements settled that the nucleus can not be present in the outer crust of neutron stars. The sequence of nuclei at the magic neutron number 82 will be explored at the Facility for Antiproton and Ion Research (FAIR) at GSI, Darmstadt, and at the future Facility for Rare Isotope Beams (FRIB) at Michigan State University. Note that the nuclear models predict the sequence of nuclei to be at the magic neutron number 50 as they are the nuclei with the highest binding energy for a given charge-to-mass ratio. The magic numbers could change for heavy and neutron-rich isotopes by, for example, shell-quenching effects or a strong dependence of the spin–orbit splitting for nuclei with a high-isospin asymmetry, which are not taken into account in the above-mentioned nuclear models.

At this point, a final remark is in order: it is interesting to state that the stable nuclei realized in nature are not only found along the valley of stability but also along a unique sequence beyond ^{56}Fe up to the neutron dripline, which is hidden inside a neutron star. These nuclei are unstable in vacuum but stabilized in the environment of the outer crust of a neutron star by the background of a degenerate electron

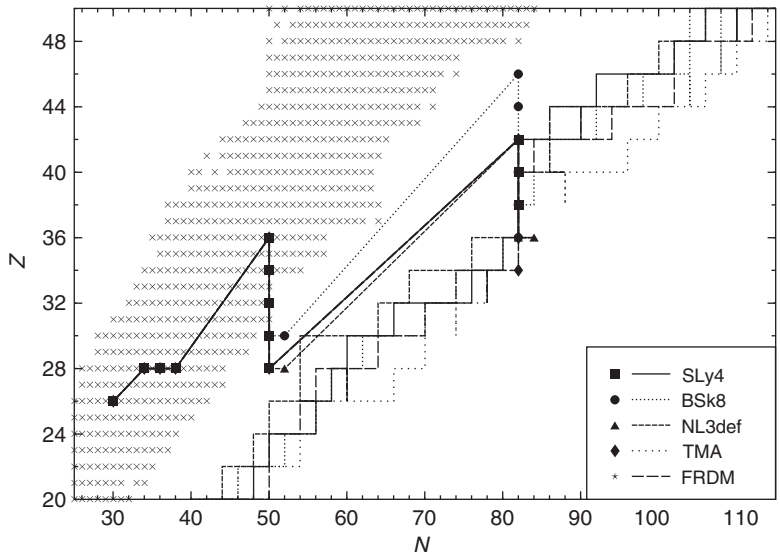


Figure 14.1 Nuclear chart of the sequence of nuclei in the outer crust of neutron stars for some selected modern nuclear models assuming β-equilibrium. The crosses mark nuclei taken from the atomic mass table. The thin lines show the neutron dripline for the various nuclear models (taken from [3]). Copyright 2006 by the American Physical Society.

gas. As the neutron star with a crust in β-equilibrium is in its final state and cannot decay to, for example, a black hole, these nuclei are indeed stable forever, at least for the lifetime of the universe.

14.2.2 The Inner Crust: Superfluid Neutrons

A relativistic Fermi gas can not provide sufficient repulsion to prevent a gravitational collapse as shown by Chandrasekhar for the case of white dwarfs. The counterforce to balance gravity has to originate then from nonrelativistic fermions; here, from neutrons. The inner crusts starts with the appearance of a neutron fluid in addition to the background of electrons in the lattice of neutron-rich nuclei at the neutron-drip density of $(4-5) \times 10^{11}\ \mathrm{g\,cm^{-3}}$. The lattice of nuclei disappears at a density of about $10^{14}\ \mathrm{g\,cm^{-3}}$, which is the critical density for the onset of the liquid core of the neutron star. In their original paper of 1971, Baym, Pethick, and Sutherland found the minimum mass of a neutron star of $0.09\ M_{\odot}$, which was entirely composed of the inner and outer crust, so a completely solid neutron star. Modern nuclear models with improved description of the neutron–neutron interactions arrive at the same value of the minimum mass, $M_{\mathrm{min}} = 0.09\ M_{\odot}$, with a corresponding radius of 200–300 km depending on the nuclear interaction used. The neutron star with the minimum mass exhibits a small liquid core, which is mainly composed of neutrons, of a few kilometers so that the constellation is close to being a fully solid neutron star. In any way, there is no

standard astrophysical scenario that would produce such small masses of neutron stars. Standard proto-neutron stars with conditions found at their birth in core-collapse supernovae have to have masses of 0.8 $M_\odot$ or more. There is no stable proto-neutron star configuration possible below that mass scale according to present models for proto-neutron star matter.

The appearance of a separate liquid of neutrons allows for physically new phenomena. The neutrons form a superfluid in the inner crust, at lower densities in the spin-singlet $^1S_0$ channel and for larger density in the spin-triplet $^3P_2$ channel, as the nucleons feel an attraction for the latter channel at higher energies as seen in nucleon–nucleon phase-shift analysis. For a rotation-powered neutron star, a pulsar, vortices can form in the superfluid. These vortices can pin and unpin with the rigid structure of the crust, the lattice of nuclei, changing the total moment of inertia of the whole neutron star. This sudden change of the moment of inertia will leave an imprint on the time of arrival of the radio pulses of the pulsar. This phenomenon has been observed and is known to be a so-called pulsar glitch. Unfortunately, the data are quite scarce as the observation of pulsar glitches need a constant monitoring and the repetition rate of pulsar glitches can be of the order of years.

There is a strong impact of a neutron superfluid on the cooling behavior of a neutron star. A neutron star cools by emitting neutrinos via the modified *Urca* process

$$N + n \rightarrow N + p + e^- + \overline{\nu}_e \qquad (14.1)$$

with a bystander nucleon in order to accomplish energy and momentum conservation for the degenerate fermions participating in the process. The neutrino emissivity is controlled by the small, but finite temperature encountered below the surface of the neutron star. Observations of isolated neutron stars reveal surface temperature in the range of 10^5–10^6 K, which increases to about 10^7–10^8 K deeper inside the neutron star according to cooling models. The presence of neutrons building Cooper pairs has a twofold effect on the cooling rates. On one hand, there will be a suppression of the emissivities as neutron pairs have to be broken up before the Urca process can take place. On the other hand, the breaking of Cooper pairs will induce an additional cooling process, the pair-breaking mechanism, releasing pairs of neutrinos via the reaction:

$$n + n \rightarrow (nn) + \nu_e + \overline{\nu}_e \qquad (14.2)$$

The latter reactions dominate, so that effectively there will be an additional cooling and a corresponding drop in the surface temperature of a neutron star. Such a temperature drop has been observed in 2011 for the neutron star in the supernova remnant Cassiopeia A (Cas A). With an estimated age of about 330 years, it is one of the youngest supernova remnants known in our galaxy. The young age and the drastic decrease in the surface temperature of the supernova remnant is the first clear indication that the cooling evolution of neutron stars is controlled by pairing processes in the interior of the neutron star and that there exists a neutron superfluid inside neutron stars. It serves as a striking example that we can extract features of the nuclear medium at high density from the observation of neutron stars. Furthermore, we can use the cooling curve to constrain the neutron-pairing gap in neutron star matter. Detailed numerical simulations of the cooling pattern of neutron stars show that the critical temperature for the onset of neutron superfluidity due to spin-triplet pairing should be about 5×10^8 K [5, 6].

The transition from the crust to the liquid core can produce phases with unusual geometric features in the bottom layer of the inner crust of a neutron star. Global charge neutrality for neutron star matter still allows for the possibility of having locally charged regions with opposite charges. In the case of interest here, positively charged regions are surrounded by electrons so that the total charge for a given cell vanishes. The balance of surface energy and Coulomb energy can produce phases with varying geometries. The different phases have been termed the *pasta phases* or simply *nuclear pasta* because of their resemblance to the shapes of certain types of pasta and are now a common terminology in the literature. With increasing density, first bubbles of the new phase will appear (gnocchi), then rods (spaghetti), and then slabs (lasagna). Then the whole geometry turns around and there will be rods of the old phase, then bubbles which appear as holes in the liquid phase. The energy difference between the pasta phases and the homogeneous structure is quite small, so that the presence of the phases and the corresponding density range cannot be determined precisely. Recent advances by using more sophisticated approaches, such as quantum molecular dynamics or 3D Skyrme–Hartree–Fock calculations, demonstrate that the pasta phase is present with rather intricate structures at the edge of the outer core. The typical density range of $(0.2–0.5) \times n_0$ found for the nuclear pasta phases, where $n_0 = 0.15–0.17$ fm^{-3} is the number density of normal nuclear matter, seems small. However, the nuclear pasta can constitute a substantial fraction for core-collapse supernova material as the density in the core at bounce reaches densities of just a little bit above n_0. Many new geometrical structures have been identified in recent 3D simulations, which can not be simply quantified in terms of the five conventional pasta phases listed above. These structures will prevail also for typical temperatures and isospin asymmetries of nuclear matter in core-collapse supernovae. The mesoscopic sizes of the nuclear pasta will have an impact on the transport properties and elasticities of the crust with corresponding implications, for example, on neutrino cooling and neutrino transport in core-collapse supernovae.

14.3 The Inner Structure of a Neutron Star

The liquid core of a neutron star can be divided into two regions: the outer core, which consists of neutrons, protons, electrons, and myons, and the inner core. The dividing line between the inner and the outer core is not really known, neither is there a detailed knowledge about the equation of state nor about the composition of the inner core. Current estimates place the critical density for the onset of a new phase to be around two to three times normal nuclear matter density. The core of low-mass neutron stars, below $(1.2–1.4) \times M_\odot$, will then consist of the outer core only, as their central density is lower than this critical density. The inner core could either consist of new particles, which emerge owing to weak equilibrium in the high-density environment, as hyperons or kaons, where kaons can form a Bose condensate. Alternatively, a completely new phase could appear, which would be more appropriately described in terms of quarks as quasi-particles, that is, strange quark matter, with an associated phase transition due to chiral symmetry restoration. In any case, the inner core will harbor strange material. Remarkably, the matter in the inner core of a neutron star

can be explored to a certain extent by astrophysical observations despite the fact that it is hidden deep inside the compact star.

14.3.1 The Outer Core: Nucleonic Matter

The typical density range of the outer core, from about 0.5 to $(2-3)n_0$, seems small in comparison to the crust, but can constitute the major part of the whole structure of a low-mass neutron star. The major components of the matter are neutrons with a small admixture of protons and leptons, which are mainly electrons with a small portion of muons. The amount of protons can reach values of more than 10% in some nuclear models, depending on the density dependence of the asymmetry energy. If the proton fraction reaches a critical value of about 11%, cooling of the neutron star can proceed via the so-called direct Urca process

$$p + e^- \rightarrow n + \nu_e, \quad n \rightarrow p + e^- + \overline{\nu}_e, \tag{14.3}$$

which does not need a bystander nucleon to guarantee energy–momentum conservation. The cooling rate via the direct Urca process is orders of magnitude faster than via the modified Urca process of the crust. Indeed, enhanced cooling has been observed for some supernovae remnants. The most prominent example is the case of the supernova remnant of 1181, where X-ray observations reveal an extremely low surface temperature of the radio source 3C58, which is the pulsar associated with the historical supernova. Note, that the corresponding NASA press release claimed evidence for exotic matter, implying the existence of quark matter in the core of the neutron star. But, as stated above, a sufficient amount of protons would also be a possibility. However, the neutrino emission rate will be suppressed by proton pairing, which can effectively reduce the emissivity such that there is no enhanced cooling originating from the nucleon direct Urca process. Reviews on cooling of neutron stars can be found in [7, 8].

The nuclear equation of state is the key input for determining the properties of the core of the neutron star. As the crust constitutes only a tiny fraction of the total mass and a small correction to the radius for typical neutron star masses, the nuclear equation of state controls the global properties of neutron stars, that is, the mass–radius relation. As the dominant contribution of the outer crust comes from neutrons, we can utilize constraints for pure neutron matter from nuclear data with a controlled extrapolation to bulk matter (see [9] for the relation of isospin asymmetry in nuclei and neutron star matter).

It became customary to classify isospin-asymmetric bulk matter in terms of an expansion in baryon number density n and isospin asymmetry. The binding energy of the system is expanded in terms of the isospin asymmetry parameter

$$\delta = \frac{n_\mathrm{n} - n_\mathrm{p}}{n}, \tag{14.4}$$

where n_p, n_n are the proton and neutron number densities, respectively. The expansion of the binding energy per baryon of bulk matter reads

$$E/A(n,\delta) = E_0(n) + \delta^2 S(n) + \mathcal{O}(\delta^2), \tag{14.5}$$

where $E_0(n)$ is the binding energy of isospin-symmetric nuclear matter and $S(n)$ is the density-dependent symmetry energy. The density dependence of the asymmetry energy could be highly nonlinear though.

An expansion of the symmetry energy around saturation density via

$$S(n) = S_0 + L \times \frac{n - n_0}{3n_0} + K_{\rm iso}\frac{(n-n_0)^2}{6n_0^2} + \cdots \quad (14.6)$$

introduces the symmetry energy $S_0 = S(n_0)$, the slope parameter L, and the curvature parameter of the symmetry energy $K_{\rm iso}$ at saturation density. Note, that the expansion is really in the nucleon Fermi momentum $k_{\rm F}$, similar to the definition of the "curvature parameter," the compression modulus K_0, for isospin-symmetric nuclear matter at saturation density

$$K_0 = k_{\rm F,0}^2 \cdot \left.\frac{d^2(E/A)}{dk_{\rm F}^2}\right|_{k_{\rm F,0}} = 9n_0^2 \times \left.\frac{d^2(E/A)}{dn^2}\right|_{n_0} \quad (14.7)$$

The slope parameter can be constrained considerably by the nuclear data from the neutron skin thickness of Sn isotopes, isospin diffusion from intermediate energy heavy-ion collisions, giant dipole resonances, and the dipole polarizability of ^{208}Pb. In general, the density dependence of the symmetry energy will involve also terms of the nucleon Fermi momentum of different orders. A direct comparison of the experimentally measured symmetry energy and compression modulus in terrestrial laboratories to properties of neutron stars involving densities well beyond saturation density seems to be not justified. However, detailed investigations have established certain cross-relations between properties of nuclei and properties of neutron stars. The pressure of neutron star matter around saturation density is fixed by the symmetry energy, as the contribution from isospin-saturated nuclear matter vanishes by definition. Hence, the pressure of neutron matter in the density range just above saturation density is controlled by the density dependence of the symmetry energy, that is, by the parameter L. In addition, it has been found that there is a strong correlation of the pressure just above n_0 and the neutron star radius R for a variety of nuclear equations of state. A rather constant radius with varying neutron star mass is a feature seen for many different equations of state for neutron stars. This feature can be explained by an equation of state, where the pressure increases with the density squared, which is the case for an interaction-dominated equation of state. Such an equation of state corresponds to a polytrope with an index of $\Gamma = 2$, which is known to give a constant radius independent of the mass as a solution of the Lane–Emden equation for compact stars. In this case, the radius is determined by the onset of the sudden stiffening of the neutron matter equation of state around saturation density. So, at least for low-mass neutron stars with a mass of 1.4 $M_\odot$ and less, the radius of the neutron star could indeed be related to the density dependence of the symmetry energy under these circumstances. A correlation between the neutron radius of ^{208}Pb and neutron star radius has been pointed out by Horowitz and Piekarewicz with regard to the ^{208}Pb Radius Experiment (PREX) experiment at Jefferson Laboratory measuring parity violation in electron scattering on ^{208}Pb [10]. First results of the PREX collaboration have been published recently [11].

Neutron star equations of state based on relativistic mean-field models usually result in rather large radii for low-mass neutron stars of around 15 km. The density dependence of the symmetry energy originates from ρ-meson exchange, which

has a rather stiff dependence on density. In its simplest version, relativistic mean-field models with parameterizations fitted to properties of nuclei result in typical values of $L = 110–140$. Extensions of the relativistic mean-field model, by introducing self-interaction terms of the vector fields of the ρ- and ω-meson or density-dependent coupling constants, can reduce the value of the slope parameter L substantially, that is, by a factor of two, so that the values of L are more compatible with the ones extracted from Skyrme–Hartree–Fock models. Correspondingly, the radii for a 1.4 $M_\odot$ neutron star decreases to values around 13 km. Skyrme models usually result in a moderate density dependence of the symmetry energy and therefore in somewhat smaller neutron star radii, so that values down to about 10 km for a 1.4 $M_\odot$ neutron star can be reached.

Microscopic approaches can give important insights into the equation of state of pure neutron matter. The widely used equation of state of Akmal, Pandharipande, and Ravenhall (APR) is based on the Argonne V_{18} two-nucleon interaction fitted to nucleon–nucleon scattering data [12]. In addition, boost corrections and a three-nucleon force derived from the Urbana model (UIX) are included. Neutron star matter in the outer core can be approximated by pure neutron matter, as there is likely only a small admixture of protons and electrons present. Then the neutron star matter equation of state is mainly determined by the neutron–neutron interaction and can be fairly well constrained by using the results of the APR equation of state. Recent calculations of pure neutron matter based on chiral-effective field theory arrive at a rather narrow band up to saturation density by varying the theoretical uncertainties of the model as the cutoff and the coefficients of the three-body force [13]. The chirally motivated neutron matter equation of state is compatible with the one of APR. Extrapolations using polytropes result in neutron star radii between about 10–14 km for a 1.4 $M_\odot$ neutron star [14]. The calculations of pure neutron matter from chiral-effective field theory and quantum Monte Carlo methods [15] give a narrow range of $L = 40–60$, which is compatible with the values extracted from nuclear data.

Conventional nuclear data can not probe the properties of neutron star matter above saturation density. For that purpose, heavy-ion data provide an outstanding opportunity to study hot and baryon-rich dense matter in the laboratory with important cross-relation to astrophysics, here in particular to core-collapse supernovae and neutron stars (see e.g., [16] for a review). As mentioned above, at intermediate colliding energies, isospin diffusion data and isotopic particle ratios can be used to derive constraints on the density dependence of the symmetry energy, although on densities at or below saturation density. High baryon densities can be achieved in relativistic bombarding energies. Transport simulations indicate that at energies of about 1–2 GeV/nucleon central baryon densities of $2–3n_0$ are reached. This density region turns out to be of particular interest to derive constraints for neutron star properties, as it is still within the region of the outer crust, meaning that exotic matter with strangeness is likely not present. The conditions in heavy-ion collision are such that because of the extremely short timescale involved, strangeness can only appear by associated production. Hence, the strangeness number is conserved and the matter is not in weak (strangeness changing) equilibrium, contrary to the situation in neutron star matter at the highest density. The comparison between

heavy-ion data and neutron star data can then be made more easily for the less-extreme densities found in the outer core. Still, heavy-ion data can not reach the extreme isospin asymmetry encountered in neutron star matter, but can constrain the bulk properties of nuclear matter at high densities.

For probing the high-density region produced in central heavy-ion collisions we need a messenger, which is produced at the highest densities and can escape freely through the low-density matter, thereby reaching the detector unattenuated. Luckily, there is a particle fulfilling these constraints at the collision energy of interest here, that is, around 1–2 GeV/nucleon. The kaon is produced via associated production either with a hyperon or with an antikaon. The threshold energy in the elementary proton–proton collision for producing a kaon is higher than 2 GeV, so that kaons cannot be produced directly within a single interaction in heavy-ion collisions. Multiple scattering processes are needed to produce kaons at sub-threshold energies in heavy-ion collisions. The denser the medium produced in the central region of the heavy-ion collisions, the larger will be number of produced kaons. By varying the incident energy of the heavy-ions, we can test the nuclear medium at different central densities. The kaons can escape nearly freely from the high-density region, where they are most favorably produced, as they have a mean free path larger than the typical size of the heavy-ion reaction zone. Hence, kaons serve as a messenger of and a tool for probing the conditions prevailing in the densest region produced in the heavy-ion collision.

The input from nuclear physics at high densities in transport simulations is not directly the equation of state; it is the self-energy of the nucleons at given momentum and surrounding local baryon density. The extrapolation from the self-energy to the equation of state has to be provided then by the nuclear models used. Of course, there will be other effects or unknowns that can affect the extraction of the nuclear equation of state, as the in-medium cross sections or the in-medium potential for kaons. Therefore, collisions of light nuclei are used to cancel the unwanted trivial in-medium effects of lower densities from the heavy-ion data. The ratio of produced K^+ per baryon is measured for the collision systems of C + C and Ni + Ni at the same incident energy per nucleon. The double ratio of the produced K^+ per baryon for the heavy-ion and the light-ion collisions appears to be useful to extract the nuclear equation of state at high densities. Different input to the transport simulations have been tested with this double ratio, showing that the equation of state is the decisive factor in controlling the energy dependence of the double ratio. The measurements were carried out by the KaoS collaboration at the SIS machine at GSI, Darmstadt . The energy dependence of the double ratio of produced K^+ can only be explained by transport simulations, if one adopts a soft nuclear equation as input. Soft equation of state means here that the Skyrme model used has a compression modulus of $K = 200$ MeV. Nuclear equations of state can now be tested if they are soft enough at the density range probed by the KaoS heavy-ion data: at $(2-3)n_0$ the nuclear model must provide a nuclear potential that lies below the result from the Skyrme model used with the parameter $K = 200$ MeV. It is important to note, that the nuclear potential below that density range is not relevant here for the comparison. This analysis is so far the only one that can be considered to make a fair statement about the nuclear

equation of state well above saturation density from heavy-ion data, see [17] for a recent review.

The central densities reached in low-mass neutron stars is about two to three times saturation density, so that a determination of the radius will probe the neutron star matter equation of state around that density. A comparison with the analysis of the heavy-ion data will then provide a measure of the symmetry energy at such densities of purely nucleonic dense matter.

14.3.2 The Inner Core: Matter with Strangeness

For massive neutron stars, it is likely that strangeness-changing processes can be present in the inner core, producing strangeness-rich matter. The new phase could be just characterized by the appearance of new strangeness-carrying hadrons, as hyperons or kaons, in addition to nucleons. It might equally well be that the new phase is more aptly described in the form of a completely new phase and that the relevant degrees of freedom to be considered are quasi-particles of up-, down-, and strange quarks, that is, strange quark matter.

The equation of state of strangeness-bearing dense matter can be probed by observing massive neutron stars. In the last years, massive pulsars have indeed been found, well above the canonical mass of 1.44 $M_\odot$ of the Hulse–Taylor pulsar. The mass of the pulsar PSR J1903-0327 was determined to be $M = (1.667 \pm 0.021) \times M_\odot$ by measuring the Shapiro delay parameters and the periastron advance [18]. In 2010, an extremely strong signal for the Shapiro delay was found for the PSR J1614-2230 resulting in a mass of $M = (1.97 \pm 0.04) \times M_\odot$ [1]. This new mass limit is so much larger than the previous one of the Hulse–Taylor pulsar that it imposes considerable constraints on the nuclear equation of state at high densities. The equation of state for the inner core, above two to three times saturation density, has to be sufficiently stiff to provide enough pressure to support a two-solar-mass neutron star (Figure 14.2).

With this new mass limit, it became considerably more difficult to construct equations of state with hyperons appearing in the inner core of the neutron star. The emergence of new particles as another degree of freedom will reduce the overall Fermi pressure, as high-momentum neutrons will be transformed to Λ hyperons, for example, and the equation of state will soften correspondingly. Only for a sufficiently stiff equation of state of pure nucleonic matter at high densities and a sufficiently large repulsive interaction between hyperons and nucleons and between hyperons themselves is it possible to reach the two-solar-mass benchmark. It implies that the amount of hyperons present in the inner core of a neutron star is limited and likely does not reach more than 15% or that strange quark matter appears at the highest densities, stiffening the equation of state again. Still the presence of just a small layer of hyperons can be important as hyperons can help cooling down a neutron star by the hyperon direct Urca process $\Lambda \to p + e^- + \overline{\nu}_e$. Data from hypernuclear experiments can provide information about the hyperon–nucleon interaction, which can be used to calculate the critical density for the onset of hyperons in neutron star matter. Calculations for various different nuclear and hypernuclear models demonstrated that hyperons are likely to appear at two to three times saturation density. The study of the hyperon–hyperon interactions is experimentally accessible via the production of

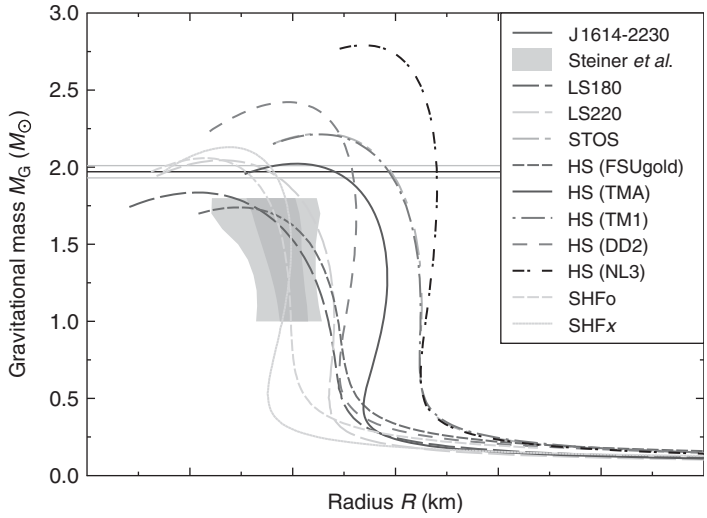


Figure 14.2 The mass–radius relation of neutron stars for equations of state used in supernova simulations. The new pulsar mass limit of $(1.97 \pm 0.04) \times M_\odot$ is shown as a vertical line. They gray shaded area stands for limits on the mass–radius relation derived from binary X-ray data. (figure courtesy Matthias Hempel, based on [19, 20]).

double hypernuclei or the measurement of hyperon–hyperon correlations in relativistic heavy-ion collisions. The strength of the hyperon–hyperon interaction will give information about hyperon pairing gaps in dense matter and its implications for suppressing the hyperon direct Urca process, similar to the case for nucleons. The role of hyperons in neutron stars has been reviewed in [21].

If strange quark matter is formed in the inner core of a neutron star, its properties can be assessed by studying astrophysical systems in several ways (for a review on strange quark matter in compact stars see [22]). There are two different types of modification from high-density strange quark matter.

- The equation of state will change at high densities. It is possible that a first-order phase transition is present, which is related to chiral symmetry restoration.

The quark matter equation of state can be probed by

– determining the mass–radius relation, which is solely fixed by the equation of state and the Tolman–Oppenheimer–Volkoff (TOV) equation of general relativity. A strong first-order phase transition can result in a new stable solution of the TOV equation, the so-called third family of compact stars.
– detecting a galactic core-collapse supernovae. An onset of a strong first-order phase transition at low densities can result in a second shock wave, producing a second peak of antineutrinos, which is observable in the present-day neutrino detectors IceCube and Super-K.

  - measuring the proto-neutron star evolution in the aftermath of a galactic supernova. The appearance of a new phase at high density can modify the spectrum and amplitude of the emitted neutrinos. In the extreme case, it leads to a delayed collapse to a black hole so that the neutrino signal shuts off suddenly.
  - finding a neutron star merger by the measurement of a gravitational-wave signal. In just a couple of years, advanced LIGO has the prospect to detect a few events per year. The spectrum of emitted gravitational waves is an indicator of the compactness of the neutron star and can provide a measurement of its radius. So far, implications of a strong first-order phase transition have not been investigated in detailed simulations, but should also be visible in the emitted gravitational-wave signal.
- The transport properties of high-density matter will be considerably different from that of ordinary matter.
  - The bulk and shear viscosity of dense matter are crucial in providing stability against the emission of gravitational waves. Certain combinations of the core temperature and the rotation frequency can be ruled out owing to the presence of the so-called r-mode instability, which is distinctly different for strange quark matter and ordinary nuclear matter.
  - The emissivities and the mean-free path of neutrinos will change in strange quark matter, which has an impact on the cooling pattern of proto-neutron stars and cold neutron stars. Observations of the surface temperature of neutron stars can provide information about the processes occurring in the very core of the neutron star.

The properties of high-density QCD matter will be also explored by studying relativistic heavy-ion collisions. At ultra-relativistic bombarding energies, the temperature generated in the fireball will be high and many baryon–antibaryon pairs will be produced. However, the net baryon density will be small. The highest possible net baryon density will be achieved at an energy range provided by the upcoming Nuclotron-based Ion Collider facility (NICA) facility at Dubna and FAIR at GSI, Darmstadt. According to transport simulations, the extreme baryon densities encountered in the central region of massive neutron stars can be reached. A multimessenger analysis of the heavy-ion data is poised to give telltale signatures of the properties of QCD matter at extreme baryon densities. Certainly, there will be a growing interplay between the nuclear astrophysics of neutron stars and relativistic heavy-ion collisions in the coming years.

14.4 Summary and Outlook

The physics of neutron stars covers many different modern research fields. Already in the crust of neutron stars, we encounter what we need to know about the

properties of nuclei far from stability, up to the neutron dripline, to understand its composition. Neutrons can be present in a separate fluid forming superfluid states. Combined with the rotation of the neutron star, vortices will form in the neutron star crust. The whole setup is present in a lattice of nuclei immersed in a background sea of electrons. Furthermore, protons can form superconducting states that will result in flux tubes influencing the magnetic field of a neutron star in a nontrivial way. It is likely that there are nontrivial geometric structures at the phase boundary from the crust to the core of a neutron star due to a first-order phase transition, the so-called nuclear pasta.

These effects have strong ties to modern solid-state physics, now present in a strongly interacting system. There is mutual interplay between the properties of neutron matter and the experimental observation of the phenomenon of Bose–Einstein condensation and the crossover to a superconducting Bardeen–Cooper–Schrieffer (BCS)-type state in ultracold atoms.

In addition, weak interactions play a decisive role in neutron stars. Neutrino reactions, which are the main processes for cooling a neutron star, are present in dense nuclear matter as the neutron β-decay. Strong interactions at high densities control the global properties of neutron stars, such as mass and radius, via the nuclear equation of state. There might be exotic phases present in the core of the neutron star, as nuclear matter with hyperons or even a phase, where the constituents of nucleons and hyperons, the quarks, form so-called strange quark matter. The theoretical modeling of the properties of the high-density matter is hampered by the fact that the underlying theory of strong interactions, QCD, cannot be solved even with the fastest supercomputers at present, so that one has to rely on effective models mimicking the features of QCD.

There is again a mutual interplay of neutron star physics with modern nuclear physics and astrophysics, relevant for the new large-scale research facilities, such as FAIR and FRIB, and their ambitious program in studying nuclei far from stability. The exploration of nuclear matter under extreme conditions, at high isospin and high densities, will open new vistas with the upcoming experiments in relativistic heavy-ion collisions study of compressed baryonic matter, the low-energy run at the relativistic heavy ion collider (RHIC) in Brookhaven National Laboratory's (BNL), the compressed baryonic matter (CBM) experiment at FAIR, and the NICA experiment in Dubna. Supplemented with the astronomical facilities such as the future James Webb space telescope, the radio telescopes of the SKA, and the gravitational wave observatory LIGO, the field of the study of neutron star is poised for discoveries and surprises in the coming years.

During the completion of this work, the mass of the pulsar PSR 0348+0432 was reported to be 2.01 ± 0.04 solar masses, setting a new mass record and supporting the existence of two solar mass pulsars [1].

Glossary

Binary Pulsar: A pulsar with a companion star, which could also be a white dwarf or another neutron star

Compression Modulus: An energy scale quantifying the amount of energy needed to compress bulk nuclear matter

Core-Collapse Supernovae: Explosion of a star more massive than eight solar masses due to the collapse of its degenerate core

Double Pulsar: The binary pulsar system PSR J0737-3039, where both neutron stars have been observed in radio waves as pulsars, allows for stringent tests of general relativity in strong gravitational fields

Equation of State: The relation between the pressure and the energy density, used here to describe the properties of neutron star matter

FAIR: Acronym for Facility for Antiproton and Ion Research; this is a large-scale research facility constructed at GSI, Darmstadt, Germany; it consists of a large ring accelerator and which conducts experiments in basic research of antiproton physics, atomic and plasma physics, relativistic heavy-ion physics, and nuclear astrophysics

FRIB: Acronym for Facility for Rare Isotope Beams, a large-scale research facility constructed at the site of the Michigan State University, East Lansing, MI, USA; it is a linear accelerator for basic research in nuclear physics and astrophysics

LIGO: Short form of large interferometer gravitational-wave observatory, a facility for measuring gravitational waves generated by neutron star mergers, consists of two 4 km long L-shaped arms with two installations near Livingston, Louisiana, and at Hanford, Washington, in the USA

Magic Numbers: Certain numbers of neutrons or protons in a nucleus, which make the nucleus particularly stable (2, 8, 20, 28, 50, 82, 126, ...)

Mass–Radius Relation: The unique relation of a neutron star mass and its radius as given by the solution of the Tolman–Oppenheimer–Volkoff equations of general relativity

Neutron-Drip Density: The density inside a neutron star, where neutrons start to become unbound so that they are dripping out of the nucleus and form a separate (neutron) fluid

Neutron Star: Compact stellar remnant of stars more massive than eight solar masses produced in core-collapse supernovae with typical masses of one to two solar masses and radii of 10–15 km

Nuclear Pasta: Geometrical structures appearing in the crust of a neutron star with shapes similar to gnocchi (bubbles), spaghetti (rods), lasagna (plates), maccheroni (tubes), and swiss cheese (holes)

Progenitor Mass: Commonly used for describing the initial mass of the star before exploding in a core-collapse supernova, the mass value at the start of stable hydrogen burning is usually adopted

Proto-Neutron Star: Hot remnant formed within the first minute after a core-collapse supernova, which usually cools down to a cold neutron star, therefore the name, but can also collapse to a black hole

Pulsars: Rotation-powered neutron stars observed by their regular pulses usually found in radio waves due to the misalignment of the rotation axis with the magnetic axis

QCD: Short form of quantum chromodynamics, the standard quantum field theory for the strong interactions between quarks by exchanging gluons of the Standard Model in particle physics, the underlying basic theory for describing the nuclear force

SKA: Short form of square kilometre array, a future radio detector array spanning a receiving area of about a square kilometer to be built in southern Africa and Australia

Strange Quark Matter: Elementary bulk matter consisting of up, Down, and strange quarks, produced in relativistic heavy-ion

collisions, likely to be present in the core of neutron stars

Symmetry Energy: Energy difference between symmetric and antisymmetric nuclear matter with regard to the relative amounts of neutrons and protons

Urca Process: Reactions that effectively cool down a neutron star by emitting neutrinos, coined by Gamow and Schoenberg with regard to a casino in Urca, a neighborhood in Rio de Janeiro, where money was effectively drawn from the gambler's pockets, also a dated Russian slang expression for thief

White Dwarf: Compact stellar remnant of stars of up to eight solar masses with masses of up to 1.4 solar masses and typical radii of about 10 000 km

References

1. Demorest, P.B., Pennucci, T., Ransom, S.M., Roberts, M.S.E., and Hessels, J.W.T. (2010) A two-solar-mass neutron star measured using Shapiro delay. *Nature*, **467**, 1081–1083.
2. Chamel, N. and Haensel, P. (2008) Physics of neutron star crusts. *Living Rev. Rel.*, **11**, 10.
3. Rüster, S.B., Hempel, M., and Schaffner-Bielich, J. (2006) The outer crust of non-accreting cold neutron stars. *Phys. Rev. C*, **73**, 035–804.
4. Baym, G., Pethick, C., and Sutherland, P. (1971) The ground state of matter at high densities: equation of state and stellar models. *Astrophys. J.*, **170**, 299–317.
5. Page, D., Prakash, M., Lattimer, J.M., and Steiner, A.W. (2011) Rapid cooling of the neutron star in cassiopeia a triggered by neutron superfluidity in dense matter. *Phys. Rev. Lett.*, **106**, 081–101. doi: 10.1103/PhysRevLett.106.081101.
6. Shternin, P.S., Yakovlev, D.G., Heinke, C.O., Ho, W.C., and Patnaude, D.J. (2011) Cooling neutron star in the Cassiopeia A supernova remnant: evidence for superfluidity in the core. *Mon. Not. Roy. Astron. Soc.*, **412**, L108–L112.
7. Yakovlev, D.G. and Pethick, C.J. (2004) Neutron star cooling. *Annu. Rev. Astron. Astrophys.*, **42**, 169–210.
8. Page, D., Lattimer, J.M., Prakash, M., and Steiner, A.W. (2004) Minimal cooling of neutron stars: a New paradigm. *Astrophys.J. Suppl.*, **155**, 623–650. doi: 10.1086/424844.
9. Steiner, A.W., Prakash, M., Lattimer, J.M., and Ellis, P.J. (2005) Isospin asymmetry in nuclei and neutron stars. *Phys. Rep.*, **411**, 325–375.
10. Horowitz, C.J. and Piekarewicz, J. (2001) Neutron star structure and the neutron radius of 208pb. *Phys. Rev. Lett.*, **86**, 5647.
11. Abrahamyan, S., Ahmed, Z., Albataineh, H., Aniol, K., Armstrong, D. *et al.* (2012) Measurement of the neutron radius of ^{208}Pb through parity-violation in electron scattering. *Phys. Rev. Lett.*, **108**, 112–502. doi: 10.1103/PhysRevLett.108.112502.
12. Akmal, A., Pandharipande, V.R., and Ravenhall, D.G. (1998) The equation of state for nucleon matter and neutron star structure. *Phys. Rev. C*, **58**, 1804–1828.
13. Hebeler, K. and Schwenk, A. (2010) Chiral three-nucleon forces and neutron matter. *Phys. Rev.*, **C82**, 014–314. doi: 10.1103/PhysRevC.82.014314.
14. Hebeler, K., Bogner, S., Furnstahl, R., Nogga, A., and Schwenk, A. (2011) Improved nuclear matter calculations from chiral low-momentum interactions. *Phys. Rev.*, **C83**, 031–301. doi: 10.1103/PhysRevC.83.031301.
15. Gandolfi, S., Carlson, J., and Reddy, S. (2012) The maximum mass and radius of neutron stars and the nuclear symmetry energy. *Phys. Rev.*, **C85**, 032–801. doi: 10.1103/PhysRevC.85.032801.
16. Stock, R. (1989) The physics of dense nuclear matter from supernovae to quark gluon plasma. *Nature*, **337**, 319–324.
17. Aichelin, J. and Schaffner-Bielich, J. (2010) *Relativistic Heavy-Ion Collisions*, New Series, Vol. I/23, Springer, Landolt-Börnstein, Chapter 3.1.
18. Freire, P.C.C., Bassa, C.G., Wex, N., Stairs, I.H., Champion, D.J., Ransom, S.M., Lazarus, P., Kaspi, V.M., Hessels, J.W.T., Kramer, M., Cordes, J.M., Verbiest, J.P.W., Podsiadlowski, P., Nice, D.J., Deneva, J.S., Lorimer, D.R., Stappers, B.W., McLaughlin, M.A., and Camilo, F. (2011) On the nature and evolution of the unique binary pulsar

J1903+0327. *Mon. Not. Roy. Astron. Soc.*, **412**, 2763–2780. doi: 10.1111/j.1365-2966.2010.18109.x.

19. Hempel, M., Fischer, T., Schaffner-Bielich, J., and Liebendorfer, M. (2012) New equations of state in simulations of core-collapse supernovae. *Astrophys. J.*, **748**, 70. doi: 10.1088/0004-637X/748/1/70.
20. Steiner, A.W., Hempel, M., and Fischer, T. (2012) Core-collapsesupernova equations of state based on neutron star observations; eprint arxiv:1207.2184
21. Schaffner-Bielich, J. (2008) Hypernuclear physics for neutron stars. *Nucl. Phys.*, **A804**, 309–321.
22. Weber, F. (2005) Strange quark matter and compact stars. *Prog. Part. Nucl. Phys.*, **54**, 193–288.

15
Supernovae and Their Nucleosynthesis

Friedrich-Karl Thielemann and Matthias Liebendörfer

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

15.1 Introduction

Supernovae belong to the most powerful explosions in galaxies (only γ-ray bursts seem to be more energetic). They are also the dominant sources of the elements in the Universe, as the Big Bang led only to the formation of H, He, and Li. Heavier elements (in astronomical terms, metals) are to some extent due to stellar winds, which contribute to C, N, O, and some fraction of the heavy elements from Fe to Bi. Supernovae are responsible for all other elements and come in two different kinds, both similar in their output of kinetic energy close to 10^{51} erg. Type Ia supernovae (SNe Ia) are related to the explosion and complete disruption of a white dwarf in a binary stellar system, which – due to mass transfer from the companion star – exceeds its maximum stable mass (the Chandrasekhar mass, $1.4,M_{\odot}$) and contracts. This leads to the ignition of C- and O-burning in an explosive manner, causing a nuclear burning front that disrupts the entire star and ejects about $0.6,M_{\odot}$ of Ni/Fe, smaller amounts of intermediate mass elements from Si through Ca, and some unburned C and O into the interstellar medium. As these objects start from very similar initial conditions (a $1.4,M_{\odot}$ white dwarf), they emit close to identical light outbursts, which turns them into standard light candles to measure distances in the Universe (this property was the basis for the 2011 Nobel Prize in Physics). While stars with an original mass of less than $8,M_{\odot}$ end their stellar evolution as white dwarfs (after having finished H- and He-burning and losing significant amounts of mass in stellar winds), more massive stars pass through advanced nuclear burning stages, also encountering C-, Ne-, O-, and Si-burning and end with a central Fe-core more massive than the Sun. This core consists of matter with the highest binding energy per nucleon and no further nuclear burning can prevent the core collapse up to nuclear densities, causing the formation of a hot so-called protoneutron star. The release of the gained gravitational binding energy of about 10^{53} erg in the form of neutrinos of all types, plus possibly the winding of magnetic fields due to rotation, permits the ejection of the outer layers with a total kinetic energy of about 10^{51} erg, similar to SNe Ia.

Both supernova explosion mechanisms depend highly on the nuclear physics in such hot, high density plasmas. While nuclear burning through fusion is one important ingredient, weak interactions also play an important role, as well as

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

the nuclear equation of state (EoS). For the conditions in both events, electrons behave close to degenerate fermions with sufficiently high Fermi energies to initiate electron capture on protons and nuclei. This is essential for the overall neutron/proton ratio in matter, affecting the nuclear EoS for core-collapse supernovae, and the composition of the ejecta in both types of events. In core-collapse supernovae, neutrinos play a major role. First, they are essential for the explosion mechanism, if a sufficient fraction of their energy is absorbed in the outer part of the Fe-core. Second, neutrino and antineutrino absorption on neutrons and protons dominates the overall neutron/proton ratio in the innermost ejected layers. This is important for the composition of the Fe-group (from Ti to Zn), but determines also whether proton- or neutron-rich isotopes of elements heavier than Fe are produced in these events. Recent investigations showed that the early phases of this so-called neutrino wind can produce, in what is called the *νp-process*, elements up to Sr on the proton-rich side of nuclear stability. We are still searching for the exact conditions that permit production of highly neutron-rich conditions, leading to the so-called rapid neutron capture process (r-process), responsible for the production of the heaviest elements up to Th, U, and Pu, where the understanding of nuclear structure far from stability is essential.

The discussion of the astrophysical conditions, the essential nuclear physics aspects for these conditions, plus a description of how this leads to the abundances of ejected elements is the focus of this chapter.

15.2 Nuclear Burning in Stars

Thermonuclear energy generation shapes the interior structure of stars, their evolutionary timescales, and the generation of new chemical elements and isotopes. The burning and resulting nuclear abundances are determined by thermonuclear reactions and weak interactions. In the following, we discuss the reactions involved in the evolution of stars and their catastrophic end stages. They differ insofar that (i) stellar burning stages proceed on timescales dictated by stellar energy loss, while (ii) explosive burning is determined by hydrodynamics of the event. The important ingredients for describing nuclear burning and the resulting composition changes (i.e., nucleosynthesis) are (i) strong interaction fusion cross sections and photodisintegrations; (ii) weak interactions related to decay, electron, or positron Captures; and finally (iii) neutrino-induced reactions.

15.2.1 Burning During Stellar Evolution

Burning stages in stellar evolution are characterized by temperature thresholds, permitting thermal Maxwell–Boltzmann distributions of (charged) particles (nuclei) to penetrate increasingly larger Coulomb barriers of electrostatic repulsion. Abundances of nuclei Y_i in a hot plasma of mixed composition are defined by their mass fraction X_i via $Y_i = X_i/A_i$ with A_i being the nuclear mass number. H-burning converts ^{1}H into ^{4}He via pp-chains or the CNO (carbon–nitrogen–oxygen)-cycles. The simplest pp-chain is initiated by $^1\mathrm{H(p,e^+\nu)^2H(p,\gamma)^3He}$ and completed by $^3\mathrm{He(^3He,}2p)^4\mathrm{He}$. The dominant CNO-cycle chain $^{12}\mathrm{C(p,\gamma)^{13}N(e^+\nu)^{13}C(p,\gamma)}$

$^{14}N(p,\gamma)^{15}O(e^{+}\nu)^{15}N(p,\alpha)^{12}C$ is controlled by the slowest reaction $^{14}N(p,\gamma)^{15}O$. The major reactions in He-burning are the triple-alpha reactions $^{4}He(2\alpha,\gamma)^{12}C$ and $^{12}C(\alpha,\gamma)^{16}O$. The triple-alpha reaction, being essentially a sequence of two two-body reactions with an extremely short-lived intermediate nucleus ^{8}Be, can formally be written as if it acts like a three-body reaction (Eq. (15.1)). The H- and He-burning stages are the only ones encountered in low- and intermediate-mass stars, leaving white dwarfs as central objects. Massive stars undergo further burning stages up to those involving the production of Fe-group nuclei. Table 15.1 lists these burning stages with typical central densities and temperatures, the duration of such burning stages, and the typical luminosity in photons (from Woosley *et al.* [1]).

In C-burning, the reaction $^{12}C(^{12}C,\alpha)^{20}Ne$ dominates, in O-burning, it is $^{16}O(^{16}O,\alpha)^{28}Si$. The nuclear reaction rates (resulting from an integration over a Maxwell–Boltzmann distribution of targets and projectiles) are discussed in the chapter on nuclear astrophysics; for more details see [2–4].

The alternative to fusion reactions are photodisintegrations, which start to play a role at sufficiently high temperatures T when $30\ kT \approx Q$ (the Q-value or energy release of the inverse capture reaction). This ensures the existence of photons with energies $> Q$ in the Planck distribution and leads to Ne-Burning [$^{20}Ne(\gamma,\alpha)^{16}O$,$^{20}Ne(\alpha,\gamma)^{24}Mg$] at $T > 1.5 \times 10^{9}$K (preceding O-burning) due to a small Q-value of ≈ 4 MeV and Si-burning at temperatures in excess of 3×10^{9}K (initiated like Ne-burning by photodisintegrations). Such photodisintegration rates (determined by integrating over a thermal (Planck) distribution of photons at temperature T) act similarly to decays but with a temperature-dependent decay rate. For a more detailed set of reactions in all burning stages, see [5].

High-density stellar cores possess a Fermi gas of degenerate electrons. The Fermi energy of electrons is $E_F = \hbar^2/2m_e(3\pi^2)^{2/3}n_e^{2/3}$, where n_e is the density of electrons $n_e = \rho N_A Y_e$, ρ denotes the matter density, and N_A Avogadro's number. In late stages of O-burning, in Si-burning (and during the later collapse stage), this Fermi energy of electrons increases to the level of nuclear energies (megaelectronvolts).

In a neutral, completely ionized plasma, the electron abundance Y_e is equal to the total proton abundance $Y_e = \sum_i Z_i Y_i$ (summing over all abundances of nuclei,

Table 15.1 Burning stages and their duration for a 20 $M_\odot$ star.

Fuel	ρ_c (g cm^{-3})	T_c (10^9K)	τ (yr)	L_{phot} (erg s^{-1})
Hydrogen	4.5(0)	0.04	8.1(6)	2.4(38)
Helium	9.7(2)	0.19	1.2(6)	3.9(38)
Carbon	1.7(5)	0.87	9.8(2)	5.5(38)
Neon	3.1(6)	1.57	6.0(−1)	5.6(38)
Oxygen	5.6(6)	1.98	1.3(0)	5.6(38)
Silicon	4.3(7)	3.34	11.5 days	5.6(38)

Table 15.2 Electron capture.

$p + e^- \rightarrow \nu_e + n$ or $p(e^-, \nu_e)n$
$(A, Z) + e^- \rightarrow \nu_e + (A, Z-1)$ or ${}^A Z(e^-, \nu_e)^A Z-1$
$E_F(\rho Y_e = 10^7 \mathrm{g\,cm^{-3}}) = 0.75$ MeV
$E_F(\rho Y_e = 10^9 \mathrm{g\,cm^{-3}}) = 4.70$ MeV

Table 15.3 Neutrino capture reactions and inelastic scattering.

$\nu_e + n \leftrightarrow p + e^-$ or $n(\nu_e, e^-)p$
$\overline{\nu}_e + p \leftrightarrow n + e^+$ or $p(\overline{\nu}_e, e^+)n$
$\nu_e + (Z, A) \leftrightarrow (Z+1, A) + e^-$ or ${}^A Z(\nu_e, e^-)^A Z+1$
$\overline{\nu}_e + (Z, A) \leftrightarrow (Z-1, A) + e^+$ or ${}^A Z(\overline{\nu}_e, e^+)^A Z-1$
$(Z, A) + \nu \leftrightarrow \nu + (Z, A)^*$

including free protons/hydrogen) and limited by the extreme values 0 (only neutrons) and 1 (only protons) with typical values during stellar evolution close to 0.5 or slightly below. If the negative Q-value of an electron capture reaction can be overcome by the electron (Fermi) energy, electron captures on protons and/or nuclei are possible. The general features for typical conditions are presented in Table 15.2.

At sufficiently high densities, electron captures can lead to an enhanced neutronization of the astrophysical plasma, in addition to the role of β-decays. In degenerate Ne–O–Mg cores (after core C-burning of stars with $8 < M/\mathrm{M}_\odot < 10$), electron captures on ^{20}Ne and ^{24}Mg cause the loss of degeneracy pressure support and introduce a collapse rather than only a contraction, which combines all further burning stages on a short collapse timescale.

In Si-burning of more massive stars, electron capture on intermediate-mass and Fe-group nuclei becomes highly important and determines the neutronization (Y_e) of the central core. These capture rates (resulting from integrating over the electron Fermi distribution) behave like decay rates, being a function of T and $n_e = \rho N_A Y_e$, the electron number density [6].

Cross sections for neutrino interaction with nucleons, nuclei, and electrons are minute, compared to those for the above reactions. It therefore requires high densities of the order $\rho > 10^{12} \mathrm{g\,cm^{-3}}$ that also neutrino capture and neutrino scattering can occur on relevant timescales.

Such neutrino-induced reactions can be expressed in a similar way as photodisintegrations and electron captures, integrating now over the corresponding neutrino distribution. The latter is, however, not necessarily in thermal equilibrium and therefore not just a function of local temperature and neutrino density. Neutrino distributions are rather determined from (neutrino) radiation transport calculations.

All the reactions discussed above, occurring at different times in the sequence of burning stages, contribute to three types of terms in reaction networks (see the chapter on nuclear astrophysics) Table 15.3. Changes in nuclear abundances Y_i are determined by a set of differential equations shown below

$$\frac{\mathrm{d}Y_i}{\mathrm{d}t} = \sum_j P^i_j \lambda_j Y_j + \sum_{j,k} P^i_{j,k} \rho N_A < j,k > Y_j Y_k + \sum_{j,k,l} P^i_{j,k,l} \rho^2 N_A^2 < j,k,l > Y_j Y_k Y_l \quad (15.1)$$

The Ps stand for the number of nuclei i created (+) or destroyed (−) in these reactions, λs stand for decay rates from pure decays, photodisintegrations, electron and neutrino capture, $< j,k >$ for Maxwellian averaged cross sections $< \sigma v >$ of reactions between nuclei j and k, while $< j,k,l >$ includes a similar expression for three-body reactions. In

general, such nuclear networks have to be solved numerically. (The abundances Y_i occurring in Eq. (15.1) are related – such as electron abundances Y_e – to number densities $n_i = \rho N_A Y_i$ and mass fractions $X_i = A_i Y_i$, with $\sum X_i = 1$.)

Core Si-burning, the final burning stage during stellar evolution, which is initiated by the photodisintegration $^{28}Si(\gamma,\alpha)^{24}Mg$ and followed by a large number of fusion and photodisintegration reactions, ends with nuclear reactions in a complete chemical equilibrium (nuclear statistical equilibrium, NSE; all compositon-changing reactions are equilibrized) and an abundance distribution around Fe. These temperatures permit photodisintegrations with typical Q-values of 8–10 MeV as well as the penetration of Coulomb barriers in capture reactions. In such an NSE, the abundance of each nucleus Y_i is only dependent on temperature T, density ρ, its nuclear binding energy B_i, and, via charge conservation, on $\sum_i Z_i Y_i = Y_e$ as discussed above. Y_e is altered by weak interactions on longer timescales. Quasi-equilibrium can occur, if localized nuclear mass regions are in equilibrium with the background of free neutrons, protons, and alphas, but offset from other regions of nuclei and thus their NSE values. Different quasi-equilibrium regions are usually separated from each other by slow reactions with typically small Q-values. Such boundaries between QSE (quasi-equilibrium)-groups, due to slow reactions, can be related to neutron or proton shell closures, such as $Z = N = 20$, separating the Si- and Fe-groups in early phases of Si-burning.

All reactions discussed above, occurring during all stellar burning stages, involve essentially nuclei from H to the Fe-group, and not much beyond. Through neutron capture reactions, also during regular stellar evolution, there is a chance to produce heavier nuclei. During core and shell He-burning specific α-induced reactions can liberate neutrons that are responsible for the slow neutron capture process (s-process). A major neutron source is the reaction $^{22}Ne(\alpha,n)^{25}Mg$, with ^{22}Ne being produced via successive α-captures on the H-burning CNO product $^{14}N(\alpha,\gamma)^{18}F(\beta^+)^{18}O(\alpha,\gamma)^{22}Ne$. If it occurs, the mixing of ^{12}C into H-burning shells can produce an even stronger neutron source $^{13}C(\alpha,n)^{16}O$ via $^{12}C(p,\gamma)^{13}N(\beta^+)^{13}C$. In massive, rotating, low-metallicity stars, mixing can lead to the production of "primary" ^{14}N and ^{22}Ne, that is, a neutron source which does not reflect the initial metallicity of ^{14}N in the CNO-cycle, and can thus be much stronger. Without such mixing processes, only secondary (metallicity-dependent) ^{22}Ne is available for $^{22}Ne(\alpha,n)^{25}Mg$ during core He-burning and shell C-burning and leads to a weak s-process. The s-process can, in principle, lead to the buildup of elements up to Pb and Bi through a series of neutron captures and β^--decays, starting from existing heavy nuclei around Fe [7]. Weak s-processing, based on secondary ^{22}Ne, does not go beyond mass numbers $A = 80 - 90$. The production of heavier nuclei is possible if primary ^{14}N and ^{22}Ne are available.

15.2.2 Burning During Explosions

Many of the nuclear-burning processes, discussed above, occur also under explosive conditions at higher temperatures and on shorter timescales. The fuels for explosive nucleosynthesis consist mainly of $N = Z$ nuclei such as ^{12}C,^{16}O,^{20}Ne,^{24}Mg, or ^{28}Si (the ashes of prior stellar burning), producing also unstable nuclei for which cross sections and also nuclear masses have to be

often predicted by theory [8–10]). In addition, at high densities, substantial electron captures on nuclei $e^- + {}^A Z \rightarrow {}^A Z - 1 + \nu$ can occur owing to energetic, degenerate electrons, when Fermi energies are high, as already discussed for late burning stages.

Explosive Si-burning differs strongly from its counterpart in stellar evolution and can be divided into three different regimes: (i) incomplete Si-burning and complete Si-burning with either (ii) a normal (high-density, low-entropy) or (iii) an α-rich (low-density, high-entropy) freeze-out of charged-particle reactions during cooling from NSE. At high temperatures or during a normal freeze-out, the abundances remain in a full NSE. The full NSE can break up into smaller equilibrium clusters (QSE). An example for such QSE-bevavior is an α-rich freeze-out, caused by the inability of the triple-alpha reaction $^4\mathrm{He}(2\alpha,\gamma)^{12}\mathrm{C}$, and the $^4\mathrm{He}(\alpha\mathrm{n},\gamma)^9\mathrm{Be}$ reaction to keep light nuclei such as n, p, and ^{4}He, and nuclei beyond $A = 12$ in an NSE during declining temperatures, when the densities are small. This causes a large alpha abundance after freeze-out of reactions. This effect, most pronounced for core-collapse supernovae, is a function of entropy, being proportional to T^3/ρ in a radiation-dominated plasma (Figure 15.1).

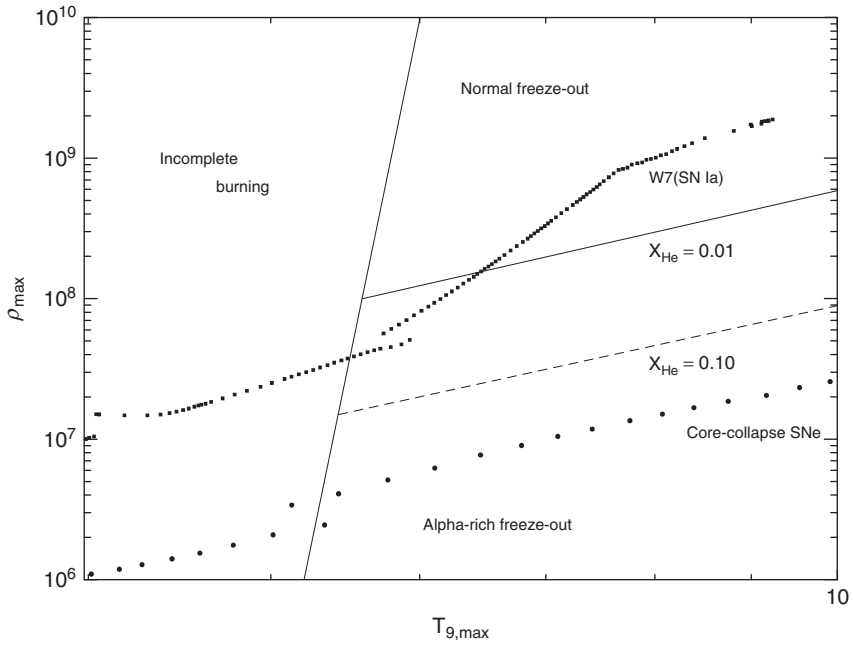


Figure 15.1 Final results of explosive Si-burning as a function of maximum temperatures and densities attained in explosions before adiabatic expansion. For temperatures in excess of 5×10^9 K, any fuel previously existing is photodisintegrated into nucleons and α-particles before reassembling in the expansion. For high densities, this is described by a full NSE with an Fe-group composition favoring nuclei with maximum binding energies and proton/nucleon ratios equal to Y_e. For lower densities, the NSE breaks into local equilibrium groups (QSE) with group boundaries due to reactions being slow in comparison to the dynamics. α-rich freeze-out (insufficient transfer from α-particles to carbon) is such a QSE-behavior. 1 and 10% lines indicate remaining α mass fractions.

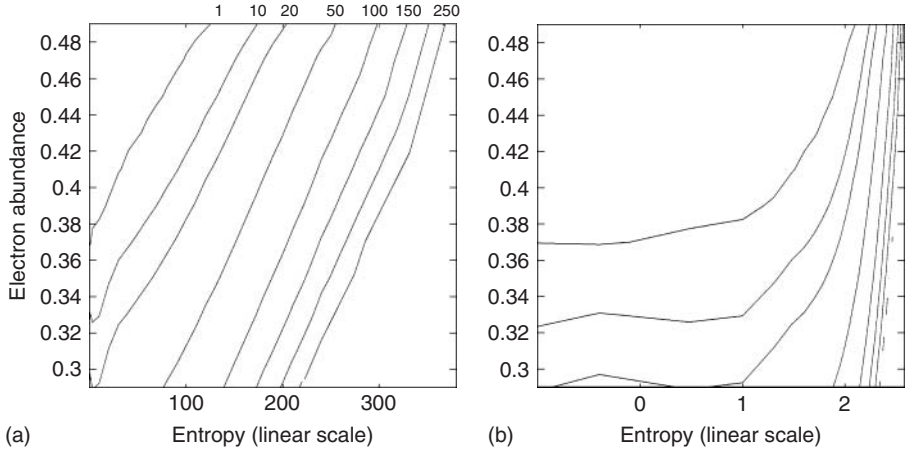


Figure 15.2 A full r-process requires A_{seed} + n/seed = $A_{actinide}$, that is, an n/seed ratio of at least 150. This can be achieved for high entropies, measured in k_B/nucleon, with a moderate Y_e (a) or for low entropies with a very low Y_e (b). The lines indicate n/seed ratios (same for both figures). For low entropies, they are only a function of Y_e (b), a value of 150 is attained at $Y_e \approx 0.15$.

The r-process nucleosynthesis (rapid neutron capture) relates to subsets of explosive Si-burning, either with low or high entropies, experiencing a normal or α-rich freeze-out. The requirement of a neutron/seed ratio of about 150 after freeze-out of charged-particle reactions (in order to produce the heaviest r-process nuclei via neutron capture from seed abundances), translates for a normal freeze-out into $Y_e = 0.12 - 0.3$. Such conditions are attained in neutron stars with a high n/p ratio of the order of 10 (i.e., $Y_e = 0.1$). For a moderate $Y_e > 0.40$, an extremely α-rich freeze-out is needed. The large mass fraction in ^{4}He (with $N = Z$) permits high ratios of remaining free neutrons to (small) abundances of heavier seed nuclei. In many cases, QSE groups of neutron captures and photodisintegrations are formed in isotopic chains of heavy elements during the r-process. The two options are displayed in Figure 15.2 for typical hydrodynamic expansion timescales. They could stand for a so-called neutrino wind in supernovae or expansions of highly compressed neutron-rich matter, occurring in neutron star mergers or jet ejecta of supernovae (Section 15.4).

15.3 End Stages of Stellar Evolution

A major impact of supernovae is related to the question: How do we understand cosmic "chemical" evolution from the first stars formed after the Big Bang (which produced only H, He, and Li) up to present? Here we present briefly the present understanding of stellar evolution as a function of initial mass (mass limits given here vary because of wind/mass loss, which depend on the initial heavy element content/metallicity and rotation). Extended discussions can be found in Heger *et al.* [11] and Maeder [12].

0. For less than 50–75 times the mass of Jupiter, that is, $5 - 7 \times 10^{-2} M_{\odot}$, objects might burn deuterium during the initial contraction, but never ignite by burning H to He and finally form so-called brown dwarfs.
1. Stars are objects that ignite hydrogen-burning. For initial masses between $5 - 7 \times 10^{-2}$ and 0.4–0.5 $M_{\odot}$, they produce a He-core with a degenerate electron gas. No He-ignition takes place as the degenerate electron gas provides sufficient pressure to prevent contraction and increasing temperatures. The end stage is a He white dwarf.
2. In the mass range 0.4/0.5–8 $M_{\odot}$ He-ignition sets in, producing after H- and He-burning, degenerate C/O-cores smaller than the Chandrasekhar mass M_{Ch} (= 1.4 $M_{\odot}$, the maximum mass of an object stabilized by the pressure of degenerate electrons). The outer H- as well as He-layers are driven off by He-burning flashes and induced pulsations, leading to C/O white dwarfs.
3. Stars with $8 < M/M_{\odot} < 10$ start C-ignition, but capture of degenerate electrons on Ne/Mg isotopes from C-burning causes pressure reduction and a fast contraction. Nuclear burning leads quasi-instantaneously to a full chemical equilibrium of nuclear reactions (NSE), producing a stellar core dominated by Fe-group nuclei. If stellar winds or mass loss in binary stellar systems prevent the growth of the core beyond M_{Ch}, the possibility of an Ne–O–Mg white dwarf, as a result of C-burning, exists.
4. Stars with $M > 10 M_{\odot}$ pass through all further burning stages (Ne-, O-, and Si-burning), with a final Fe-core of $M > M_{Ch}$ ($1.4 \times 2 < Z/A > ^2 M_{\odot}$) that contracts/collapses to nuclear densities. The degenerate Fermi gas of nucleons, plus short-range interactions, cause a bounce, not sufficient for an explosive ejection of all outer layers. Two-dimensional simulations of 10–15 $M_{\odot}$ stars can already explain explosions due to neutrino energy deposition, escape from the hot protoneutron star (PNS), and ongoing 2D and 3D simulations are at the verge of confirming the regular core-collapse supernova mechanism for 10–20 (or 25) $M_{\odot}$.
5. The next frontier includes the question what happens if a black hole rather than a neutron star is the remaining central object? This transition is expected somewhere in the range of 20–25 $M_{\odot}$ (for models without mass loss during stellar evolution) and it can come in two scenarios: (i) there exists a possible range where initially a successful explosion with an outgoing shock wave is launched, but reverse shocks (resulting from intershell density gradients) lead to the fallback of material, and thus an increase of the central neutron star mass. This could encounter the maximum mass limit of neutron stars and proceed toward the formation of a black hole. Such a mechanism would be visible like a supernova, but a large fraction of the innermost (initially ejected) material will not contribute, including the key nucleus ^{56}Ni, which is central to the observed light curve via its radioactive decay. (ii) More massive stellar progenitors will – without reaching the stage of an initially successful supernova explosion with neutron star formation – experience continuous accretion of infalling matter, and the

direct formation of a (stellar mass) black hole. Dependent on rotation and magnetic fields, an accretion disk forms with polar jet ejection (similar to jets from active galaxies with central supermassive black holes). This causes (highly energetic) explosions with large masses of ^{56}Ni, probably linked to collapsars/hypernovae/long-duration γ-ray bursts. The increasing (and extensive) amount of large-scale observational surveys of supernova explosions has led to many observations of (i) faint supernovae and (ii) type Ibc supernovae/hypernovae/GRBs, linked to massive stars that lost their hydrogen (and frequently) helium envelopes (classified as type I supernovae). Their explosion mechanism is strongly linked to understanding the nuclear EoS, the maximum neutron star mass, and magnetohydrodynamic (MHD) effects combined with fast rotation.

6. In principle, a further end stage of stellar evolution is expected for stars with initial masses between 140 and 260 $M_\odot$ (without mass loss), so called pair-creation or pair instability supernovae (PISNe). The internal pressure of massive stars is dominated by radiation pressure of thermal/blackbody photons. For such massive stars, the Planck spectrum of photons extends beyond 1 MeV during the evolution after C-burning, thus permitting to form electron–positron pairs. This reduction in photons (and their energy/momentum) leads to a sudden pressure decrease, accelerated contraction, and a violent ignition of O-burning. For such (not yet compact) stars, this energy release is sufficient to explode the whole star without a central remnant. Despite active research and predictions, it is uncertain whether these objects do form or stellar evolution might not lead to these end stages.
7. More than half of all stars form in binary or multiple systems. Dependent on their distance and sizes (as a function of evolution stage), mass transfer is possible, which also involves transfer (accretion) onto compact end stages (white dwarfs, neutron stars, and black holes). Accretion onto neutron stars leads to X-ray bursts, onto white dwarfs to novae (in both cases igniting explosively unburned, accreted hydrogen and/or helium). In case of higher accretion rates, the surface of white dwarfs can become hot enough to permit continuous H- and He-burning, increasing the C/O white dwarf beyond the maximum stable M_{Ch}, causing contraction and a thermonuclear explosion, known as type Ia supernovae (SNe Ia). (The Physics Nobel Prize in 2011 has made use of them as standard(izable) candles, permitting cosmological distance measurements.)

This chapter will mostly concentrate on topics 3–5 and 7, that is, discuss the supernova explosion mechanism with major focus on the nuclear physics involved and the composition of ejected matter as a result of explosive nuclear burning.

15.4 Core-Collapse Supernovae

For late burning stages after core He-burning, leading to C/O-cores, the partial or full degeneracy of the electron

gas is important. The critical Chandrasekhar limit $M_{Ch}(\rho Y_e, T)$ decides on further contraction and the central ignition of subsequent burning stages, that is, C-, Ne-, O-, and Si-burning. Dependent on the Fermi energy of the degenerate electron gas, electron capture on the C-burning products ^{20}Ne and ^{24}Mg can initiate a collapse, leading directly via NSE to a central Fe-core (case 4 of Section 15.3). More massive stars will proceed through all burning stages until Si-burning will finally produce an Fe-core. All late burning stages proceed on timescales that are orders of magnitude shorter than those for core H- and He-burning. The reason is that the energy carried away by freely escaping neutrinos dominates over radiation losses. These neutrinos are mostly not produced by reactions or decay of nuclei. They can be generated by $e^- + e^+$-pair annihilation (pair neutrinos), electron–photon scattering with neutrino–antineutrino pair creation (photo neutrinos), and neutrino–antineutrino pair creation from plasma oscillations (plasmon neutrinos). The timescales for the individual burning stages given in Table 15.1 in Section 15.2 refer to a solar metallicity $20\,M_\odot$ star without mass loss by Woosley *et al.* [1]. [Effects of mass loss, rotation, and metallicity can change these timescales to some extent (up to 20%).] Owing to the large difference in evolution timescales, the dominant mass loss by stellar winds occurs during H- and He-burning, while the final outcome of stellar evolution is determined by the CO-core size after He-burning. Therefore, given all dependencies of stellar evolution via initial metallicities and rotation, the initial main sequence mass of a star is less indicative of the final outcome than the size of the C/O-core.

15.4.1 The Onset of Core Collapse

In the late phases of O- and Si-burning, electron captures make matter more neutron rich, that is, it decreases Y_e, the electron- or proton-to-nucleon ratio. In high-density O-burning ($\rho > 2 \times 10^7\,g\,cm^{-3}$), two-electron capture reactions become important and lead to a decrease in Y_e, $^{33}S(e^-, \nu)^{33}P$, and $^{35}Cl(e^-, \nu)^{35}S$. Such effects become more extensive at even higher densities in Si-burning and a wide range of isotopes of Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, and Se have been identified to be of major importance. The amount of electron capture and the resulting Y_e has consequences on burning core sizes. The final size of the resulting inner Fe-core, representing the maximum mass that can be supported by the dominant pressure of the degenerate electron gas, is a function of Y_e but also reflects temperature effects if the electron gas is not completely degenerate [13]

$$M_{Ch}(Y_e, S_e) = 1.44(2Y_e)^2 \left[1 + \left(\frac{S_e}{\pi Y_e}\right)^2\right] M_\odot \tag{15.2}$$

Here, S_e denotes the entropy of the electron gas in k_B per baryon. At this point, the core collapses and, if not too massive, bounces and explodes in spectacular core-collapse events known as *type II* or *Ib/c supernovae*. These explosions mark the formation of a neutron star or black hole at the end of the life of the star and play a preeminent role in the nucleosynthesis and chemical evolution of the galaxy. The collapse is initiated by the capture of degenerate electrons on nuclei, which reduces the pressure of the dominant contribution by the degenerate electron gas. Alternatively, for lower densities and higher temperatures (in more

massive stars), endoergic photodisintegrations of nuclei reduce the thermal energy, leading to a pressure reduction as well. The evolution in the core is determined by gravity, which produces the collapse of the core, weak interactions, which determine the rate at which electrons are captured, and the rate at which neutrinos are trapped during the collapse. Stellar evolution, until core densities of about $10^{10}\,\mathrm{g\,cm^{-3}}$ and temperatures between 5 and $10\times10^{9}\,\mathrm{K}$ are attained, requires the consideration of extensive nuclear reaction networks, but is simplified by the fact that neutrinos need only be treated as a sink of energy and lepton number (owing to their immediate escape). Neutrino scattering increases with increasing density and in later phases the neutrino mean free path becomes shorter, so that neutrinos eventually proceed through all phases of free streaming, diffusion, and trapping. An adequate handling of the transitions between these transport regimes necessitates a detailed time- and space-dependent bookkeeping of the neutrino distributions in the core, that is, neutrino radiation transport. During initial collapse, only the electron neutrinos ν_e are important; during later stages, heavy neutrinos, ν_μ, ν_τ, and their antiparticles need to be included in numerical simulations of core-collapse and postbounce evolution.

The temperatures during collapse and explosion are high enough so that the matter composition is given by NSE, avoiding the need for solving nuclear reaction networks. The main weak interaction processes during the final evolution of a massive star are electron capture and β-decays. This requires the calculation of Fermi and Gamow–Teller (GT) transitions. β^--decay converts a neutron inside the nucleus into a proton and emits an electron. The presence of a degenerate electron gas, with fully populated levels up to the Fermi energy E_F, blocks decays that would produce electrons with smaller energies. Then, the decay rate of a given nuclear state is greatly reduced or even completely blocked at high densities. However, owing to the finite temperatures, excited states in the decaying nucleus can be thermally populated. Some of these states are connected by large GT transitions to low-lying states in the daughter nucleus, producing electrons above the Fermi energy and can therefore contribute significantly to the stellar β-decay rates. The importance of these states in the parent nucleus for the β-decay is discussed in detail in Bethe [13]. An accurate understanding of these effects is essential for a reliable evaluation of the stellar weak-interaction rates. The most advanced approach is based on large-scale shell-model calculations [6]. The resulting evolution of Y_e during the presupernova phase is essential as it determines the actual $M_{Ch}(Y_e, T)$, which influences the dynamical evolution of core collapse.

15.4.2 The Explosion Mechanism

Once the inner Fe-core grows beyond the maximum mass supported by the dominant pressure of the degenerate electron gas, it enters a dynamical phase of gravitational collapse during which it compactifies by ~ 5 orders of magnitude. The nuclear saturation density (i.e., the density of stable nuclei $\approx 2\times10^{14}\,\mathrm{g\,cm^{-3}}$) is exceeded at the center of the collapse and a PNS is formed. The dynamical timescale reduces from a few hundreds of milliseconds at the onset of collapse to a few milliseconds after the core has bounced back at nuclear densities (Figure 15.3).

The accretion phase onto the PNS with fluid instabilities and radiative transfer phenomena is not yet fully understood.

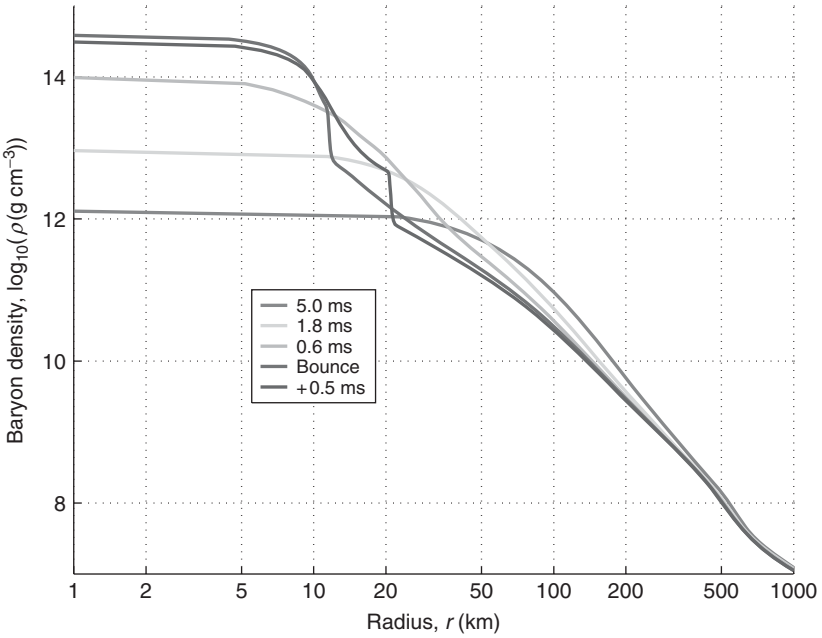


Figure 15.3 A sequence of density profiles of a 13 $M_\odot$ star before (5, 1.8, and 0.6 ms), at core bounce, and after core bounce (+0.5 ms). For such a relatively low-mass supernova with a small Fe-core, the bounce occurs at a maximum density of less than twice nuclear matter density. At bounce, one recognizes the size of the homologous core (with roughly constant density). After bounce, the emergence of an outward moving density (shock) wave can be witnessed.

The crucial weak processes during the collapse and postbounce evolution involve electrons/positrons, neutrinos, and nuclei $(A, Z) : \nu + (A, Z) \leftrightarrow \nu + (A, Z)$, $\nu + \mathrm{e}^{\pm} \leftrightarrow \nu + \mathrm{e}^{\pm}$, $\mathrm{p} + \mathrm{e}^{-} \leftrightarrow \mathrm{n} + \nu_\mathrm{e}$, $(A, Z) + \mathrm{e}^{-} \leftrightarrow (A, Z-1) + \nu_\mathrm{e}$, $\nu + N \leftrightarrow \nu + N$, $\mathrm{n} + \mathrm{e}^{+} \leftrightarrow \mathrm{p} + \overline{\nu_\mathrm{e}}$, $(A, Z) + \mathrm{e}^{+} \leftrightarrow (A, Z+1) + \overline{\nu_\mathrm{e}}$, $\nu + (A, Z) \leftrightarrow \nu + (A, Z)^*$, $(A, Z)^* \leftrightarrow (A, Z) + \nu + \overline{\nu}$, $N + N \leftrightarrow N + N + \nu + \overline{\nu}$, $\nu_\mathrm{e} + \overline{\nu_\mathrm{e}} \leftrightarrow \nu_{\mu,\tau} + \overline{\nu_{\mu,\tau}}$, $\mathrm{e}^{+} + \mathrm{e}^{-} \leftrightarrow \nu + \overline{\nu}$. This phase may last 0.5–10 s, much longer than the dynamical or transport timescale. Eventually, it will lead to the observed vigorous supernova explosion with an outward-propagating shock wave. The remaining PNS at the center enters another phase during which it cools by neutrino emission and contracts–or even collapses to a black hole in a last dynamical phase.

Improved electron capture rates on heavy nuclei overcame the idealized blocking of GT transitions in the traditionally applied single-particle model. In the single-particle picture of nuclei, the so-called pf-shell is filled for $Z = 40$ or $N = 40$ for protons or neutrons, respectively. Neutron numbers beyond $N = 40$ require a filling of the gd-orbits. If during core collapse, nuclei (or the Y_e) become(s) so neutron-rich that nuclei with $Z < 40$ and $N > 40$ dominate the NSE composition, electron capture would require the conversion of an *fp* proton to a *gd* neutron as all *pf* neutron orbits are filled. This Pauli-blocked transition would lead to the dominance of

electron capture on protons rather than nuclei under such conditions. The recent finding that configuration mixing and finite temperature effects result in unfilled *pf* neutron orbits removes this Pauli-blocking and results in the fact that under these conditions, electron capture rates on nuclei dominate those on free protons. Two effects are found due to new sets of electron capture rates [14]: (i) at low densities, for less neutron-rich nuclei, the total amount of electron capture is reduced with an improved description of GT transitions, affecting the early collapse phase. (ii) At high densities in the late collapse phase, the total amount of electron capture is enhanced, leading to smaller Y_e values than before. Such changes cause a reduction of maximum stable core sizes down to $M_{Ch} = 0.5M_{\odot}$. A faster deleptonization in the collapse phase in comparison to captures on free protons alone results in a 20% smaller inner core at bounce.

Taking this improved physics into account, leads, in the entire simulations (i.e., all mass zones involved) to conditions in densities ρ, electron abundance Y_e, and entropy s per baryon, which require the knowledge of the EoS [15] or other microscopic physics in current supernova simulations. The detailed treatment of the neutrino transport and interactions is of great importance for nucleosynthesis. This also opens an opportunity to investigate neutrino flavor oscillations among electron, muon, and tau neutrinos. On the one hand, the long-term explosion runs achieve (low) density structures that allow for the MSW (Mikheyev–Smirnov–Wolfenstein neutrino flavor oscillations in the outer layers. These may give additional hints on the expansion velocity and density distribution in the case that the neutrinos can be observed from a near-by supernova. On the other hand, collective flavor transitions have recently been postulated in regions where the neutrino density exceeds the electron density [16]. Such conditions will be attained in the evacuated zone that surrounds the PNS after the onset of an explosion. The impact of these collective neutrino flavor oscillations on accurate neutrino transport and spectra still needs to be determined.

Figure 15.4 shows a 2D slice through a 3D simulation of core-collapse and postbounce evolution. Figure 15.5 gives an optimistic outlook from present core-collapse simulations for a number of initial stellar masses.

In addition to the neutrino-driven explosions (see for reviews [17–19]), an EoS-caused explosion scenario is possible, based on the quark–hadron phase transition at supranuclear densities, but depending on choices of EoS properties in a narrow parameter range [20]. Fast rotation and strong magnetic fields have been discussed for more than 30 years. The major outcome was that high rotation rates and (possibly unrealistically high) magnetic fields can launch explosions. The question is whether such magnetic fields can be attained during collapse with rotation and on which timescale after collapse. This topic has recently been readdressed for the (probably rare) scenario of stellar core collapse with high magnetic fields and fast rotation rates. The results of a 3D MHD simulation are given in detail in Winteler *et al.* [21]; see also [22] for earlier investigations. They employed a precollapse 15 $M_{\odot}$ model with initial shellular rotation $\Omega(r) = \Omega_0 R_0^2/(r^2 + R_0^2)$, $\Omega_0 = \pi s^{-1}$, and $R_0 = 1000$ km. The magnetic field was approximated by a homogeneous, purely poloidal field of 5×10^{12}G. The conservation of angular momentum in combination with the collapse leads to a

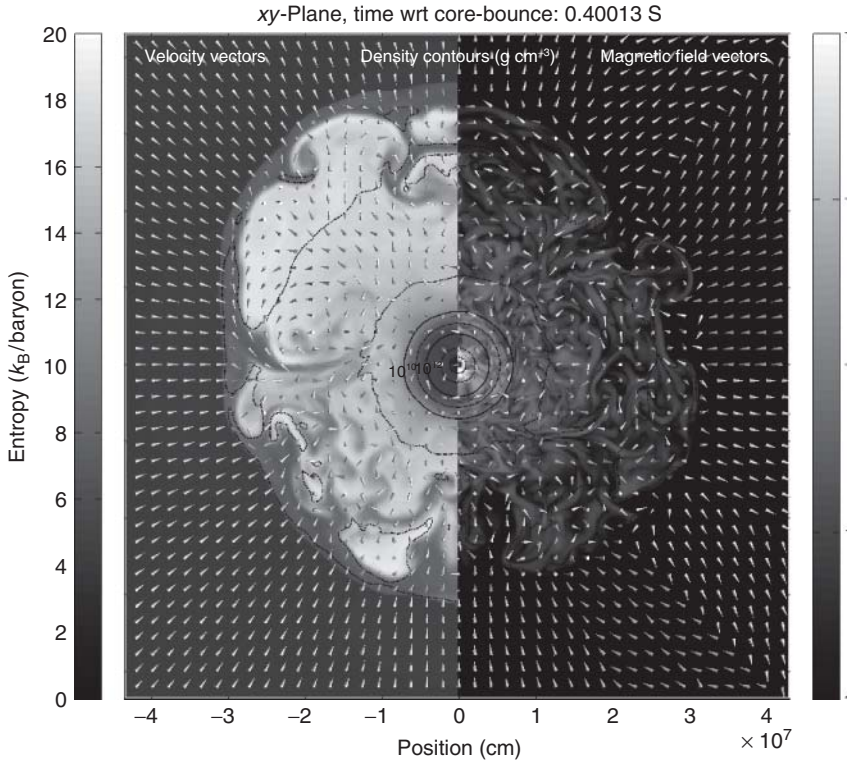


Figure 15.4 Illustration of a 3D simulation with a resolution of 600^3 zones. Shown are density contours as black lines for a 15 $M_\odot$ star 0.4 s after bounce. (a) The grayscale indicates the specific entropy and the cones the direction of the velocity. (b)The grayscale refers to the magnetic field strength and the cones to its direction. The cool high-density interior of the PNS and the hot low-density accreted matter behind the standing accretion front are clearly distinguishable.

massive spin-up of the core and significant deformations. The magnetic field is amplified by magnetic flux conservation up to $\approx 5 \times 10^{15}$G.

After bounce, differential rotation winds up the poloidal field into a very strong toroidal field, increasing the magnetic energy/pressure at the expense of rotational energy. Strongly magnetized regions appear near the rotational axis with an associated magnetic pressure quickly reaching and exceeding that of the local gas pressure. The Lorentz force becomes dynamically important and drives a bipolar outflow, that is, jets. Figure 15.6 displays a snapshot at 31 ms. This specific model is the one exception where 3D ejecta and their conditions really exist in model predictions. For 8–10 $M_\odot$ models (electron capture supernovae), similar 2D simulations exist. Otherwise, optimistic preliminary investigations are available, but no full exploding hydrodynamic calculations that would permit detailed nucleosynthesis studies. Therefore, in the following section, we first describe purely phenomenological calculations based on artificially induced explosions via a ''piston'' or energy deposition

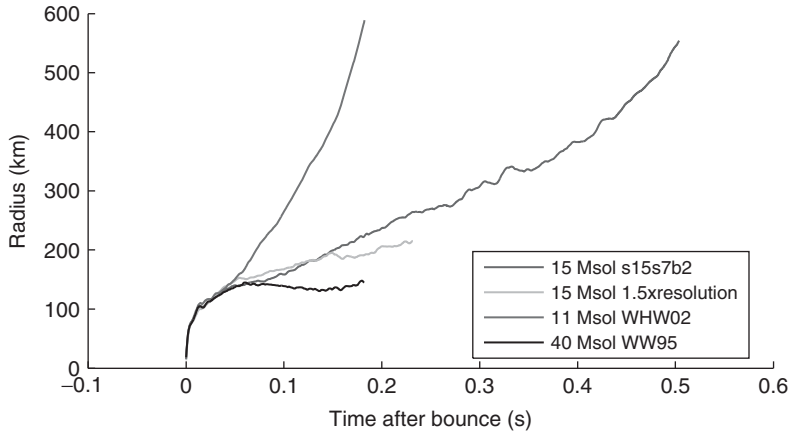


Figure 15.5 Three-dimensional results produced by the [3D] magnetohydrodynamics code FISH. Shown is a comparison of the (3D-averaged) shock propagation radius as a function of time for progenitor stars with initial masses of 11, 15 (with two resolutions) and 40 $M_{\odot}$, indicating an increase in the shock propagation time until explosion.

in terms of a ''thermal bomb,'' in order to discuss expected nucleosynthesis effects.

15.4.3 Nucleosynthesis in Explosions

15.4.3.1 Major Explosive Burning Processes

Supernova nucleosynthesis predictions have a long tradition, but all of them relied on artificially induced explosions of the progenitor star model, either via a piston or a thermal bomb [1, 5]. The mass cut between the ejecta and the remnant does not emerge from this kind of simulations but has to be determined from additional conditions. While the use of artificially introduced explosions is justifiable for the outer stellar layers, provided we know the correct explosion energy (on the order of 10^{51} erg seen in observations), it clearly is incorrect for the innermost ejected layers, which should be directly related to the physical processes causing the explosion. This affects especially the Fe-group composition and also the so-called neutrino wind, emitted seconds after the supernova explosion, and considered as a possible source of the r-process to produce the heaviest elements via neutron captures (the discussion that follows).

In the outer explosively burning layers, essentially only the energy in the shock front matters. The behavior of these zones can be easily understood from the maximum temperatures attained in the radiation bubble, which can be approximated by an artificially induced thermal bomb. For a given/known Y_e and density ρ, the most significant parameter in explosive nucleosynthesis is the temperature. A good prediction for the composition can be made by only knowing T_{max}. Matter behind the shock front is strongly radiation dominated. Assuming an almost homogeneous density and temperature distribution behind the shock (which is approximately correct), one can equate the supernova energy with the radiation energy

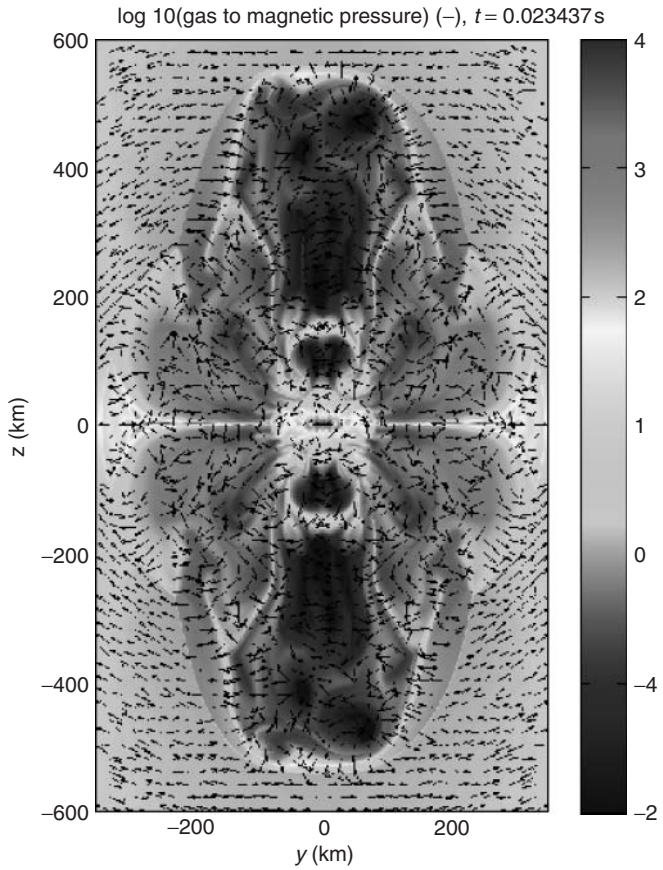


Figure 15.6 Two-dimensional cut through 3D simulation 23 ms after bounce, showing the ratio of gas pressure and magnetic pressure as contours on a color scale. Arrows indicate the movement of the fluid. Eventually polar jets will form to eject matter in both directions along the rotation axis, where the magnetic pressure is strongest (and the matter is very neutron rich).

inside the radius r of the shock front

$$E_{\mathrm{SN}} = \frac{4\pi}{3} r^3 a T(r)^4 \qquad (15.3)$$

This equation can be solved for r. With $T = 5 \times 10^9$ K, the lower bound for explosive Si-burning with complete Si-exhaustion, and an explosion energy of $E_{\mathrm{SN}} = 10^{51}$ erg, the result is $r \approx 3700$ km. For a 20 $M_\odot$ star, this radius corresponds to 1.7$M_\odot$, in excellent agreement with exact hydrodynamic calculations. Temperatures that characterize the edge of the other explosive burning zones correspond to the following radii: incomplete Si-burning ($T_9 = 4, r = 4980$ km), explosive O-burning (3.3,6430), and explosive Ne/C-burning (2.1,11 750). This relates to masses of 1.75, 1.81, and 2.05 $M_\odot$ in case of a 20 $M_\odot$ star. The explosive nucleosynthesis due to burning in the shock front is shown in Figure 15.7 for a few major nuclei.

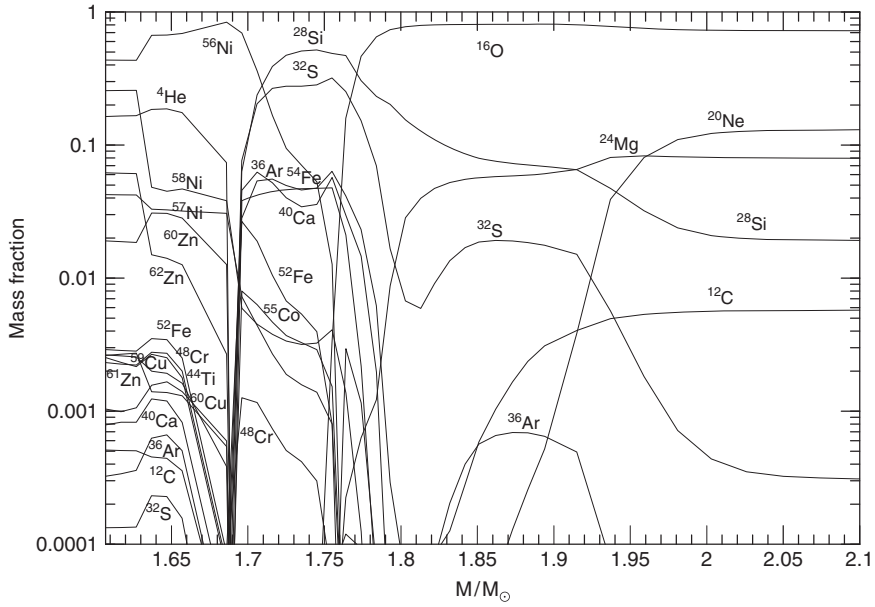


Figure 15.7 Mass fractions of a few major nuclei after passage of the supernova shock front through a star with an initial mass of 20 $M_\odot$. Matter outside 2 $M_\odot$ is essentially unaltered. Mass zones further in experience explosive Si-, O-, Ne-, and C-burning. For ejecting 0.07 $M_\odot$ of ^{56}Ni, the mass cut between neutron star and ejecta is required to be located at 1.6 $M_\odot$. The innermost ejecta are affected by the initial Y_e, here taken unaltered from the collapsing stellar model.

At less than 1.7 $M_\odot$, all Fe-group nuclei are produced in explosive Si-burning during the supernova event. At 1.63 $M_\odot$, Y_e changes from 0.494 to 0.499 and leads to a smaller ^{56}Ni abundance further inside, where more neutron-rich Ni-isotopes share the abundance with ^{56}Ni. This is an artifact of the Y_e gradient in the precollapse model, which can be changed in a consistent explosion treatment via neutrino interactions with this matter. Only α-rich freeze-out and incomplete Si-burning are encountered (see also Figure 15.1). The most abundant nucleus in the α-rich freeze-out is ^{56}Ni.

15.4.3.2 Explosive Burning Off Stability

The p-Process Up to now we discussed the production of heavy nuclei beyond the Fe-group only via the s-process in hydrostatic stellar evolution and rapid neutron captures (r-process) in explosive environments. A number of proton-rich (p-)isotopes of naturally occurring stable heavy nuclei cannot be produced by neutron captures along the line of stability. The currently most-favored production mechanism for those 35 p-isotopes between Se and Hg is photodisintegration of intermediate and heavy elements at high temperatures in late (explosive) evolution stages of massive stars. However, not all p-nuclides can be produced satisfactorily, yet. A well-known deficiency in the model is the underproduction of the Mo–Ru region, but the region $151 < A < 167$ is also underproduced, even in recent calculations [23].

In explosive Ne/O-burning zones of a supernova explosion, that is, at $\approx 2-$

3×10^9 K, where partial (but not complete) photodisintegration of preexisting nuclei occurs (from prior hydrostatic evolution or inherited metallicity), the process, acts like ''spallation'' and commences with a sequence of (γ, n)-reactions, moving the abundances to the proton-rich side, until (γ, p) and/or (γ, α)-reactions become faster than neutron emissions, and the flow branches and feeds other isotopic chains. Finally, the remaining unstable nuclei decay back to stability. Numerous experimental and theoretical efforts have been undertaken to improve the reaction input, especially with respect to open questions in optical potentials for α-particles and protons. However, nuclear uncertainties cannot change the underproduction of especially the light p-nuclei. Another process or different initial seed distributions seems to be required to supply these missing abundances.

The *ν*p-Process Neutron-deficient nuclei can also be produced in the recently discovered *ν*p-process in core-collapse supernovae. The neutrino and antineutrino flux determined via the reactions $\overline{\nu}_e + p \rightarrow n + e^+, \nu_e + n \rightarrow p + e^-$ influences the n/p-ratio. If the average antineutrino energy does not exceed the neutrino energy by more than $4 \times (m_n - m_p)$, matter turns proton rich ($Y_e > 0.5$) in the early neutrino wind. This includes the innermost ejecta of core-collapse supernovae. The discussion of these innermost ejected mass zones has been skipped above, when utilizing a thermal bomb and precollapse conditions (Y_e). For explosive nucleosynthesis in a 20 $M_\odot$ star, the *ν*p-process occurs the when this proton-rich matter is ejected, experiencing an α- and proton-rich freeze-out which does not stop at ^{56}Ni but continues up to ^{64}Ge (later decaying to ^{64}Zn). Thus core collapse yields include Fe-group nuclei up to essentially ^{64}Zn. Without neutrinos, synthesis of nuclei beyond the iron peak becomes very inefficient due to bottleneck (mainly even–even $N = Z$) nuclei with long β-decay half-lives and small proton-capture cross sections (e.g., ^{64}Ge).

When being subject to a large neutrino/antineutrino flux from the PNS, antineutrinos are readily captured on free protons on a timescale of a few seconds, leading to residual neutron densities of 10^{14}–10^{15}cm^{-3} for several seconds. These neutrons are easily captured by heavy neutron-deficient nuclei, for example, ^{64}Ge, inducing (n,p) reactions with timescales much shorter than the β-decay half-life. This permits further proton captures and allows the nucleosynthesis flow to continue to heavier nuclei. In Figure 15.8, we show *ν*p-process nucleosynthesis results from the innermost early neutrino wind ejecta produced in the explosion of a 15 $M_\odot$ star, based on two sets of nuclear mass predictions/measurements far from stability. These results show that the light p-nuclei of Mo and Ru, which are deficient in p-process calculations, can easily be produced. Further processing depends on the expansion (speed) of matter and the overlying mass of ejecta (for an overview see Ref. [5]).

The r-Process r-process conditions, discussed in Section 15.2.2, leading to the production of isotopes 10–15 units off stability, require high neutron densities and the knowledge of properties of such nuclei [9, 10]. Such conditions are attained in expanding neutron star matter with an n/p ratio of the order of 10 (i.e., a proton/nucleon ratio $Y_e = 0.1$). Another option is the so-called very α-rich freeze-out from explosive Si-burning with a moderate neutron excess,

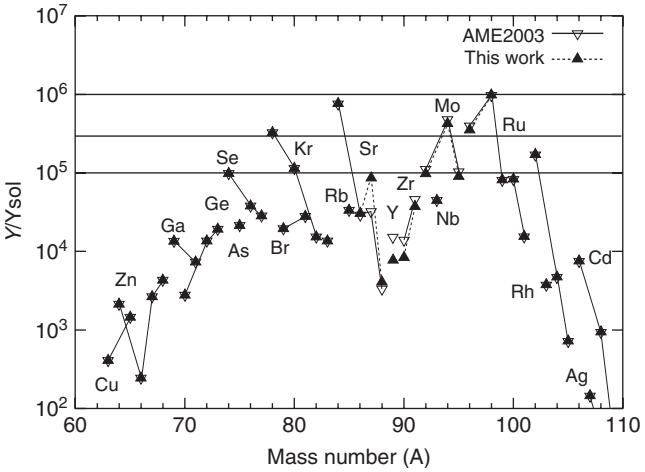


Figure 15.8 Final abundances in mass zones experiencing the νp-process, that is, the innermost ejecta of explosive, α-rich freeze-out Si-burning, normalized to solar after decay for two sets of thermonuclear reaction rates/masses. Matter up to $A = 100$ can be produced easily.

that is, $Y_e \approx 0.4–0.48$. Here the main question is whether such moderately neutron-rich conditions and very high entropies are attained in the late neutrino wind of core-collapse supernovae (Figure 15.2). Figure 15.9 shows that, in case of high entropies, a full r-process pattern can be produced. Figure 15.10 shows the results for the expansion of highly compressed neutron-rich matter, occurring in jet ejecta of rotating MHD supernovae [21].

With initial Y_es of the order 0.15–0.3, conditions are found close to those in neutron star merger ejecta (a scenario not discussed here). Nuclei around the second ($A = 130$) and third ($A = 195$) r-process peaks, and up to the actinides, can be synthesized in the jets with peak positions that agree with solar system abundances. The large trough in the mass range 140–160 is due to the strong $N = 82$ shell-closure of a specific mass model and a simplified prescription of fission fragments.

15.5 Type Ia Supernovae

SNe Ia are related to mass transfer in binary stellar systems involving a white dwarf. The prominent scenario is the accretion of matter from the binary companion at a high mass transfer rate, leading to steady H- and He-burning at the base of the accreted matter, and thus increasing the mass of the C/O white dwarf. When the maximum stable mass M_{Ch} is approached, contraction sets in, causing central ignition of C. As this ignition occurs in (electron-) degenerate matter under conditions where the pressure is only density and not energy dependent, a thermal runaway occurs, disrupting the whole star ([24], see Figures 15.11 and 15.12). The merger of two white dwarfs in a binary system can also (but not necessarily) lead to a combined object beyond M_{Ch}, also causing a white dwarf explosion but with probably widely varying conditions. In dense

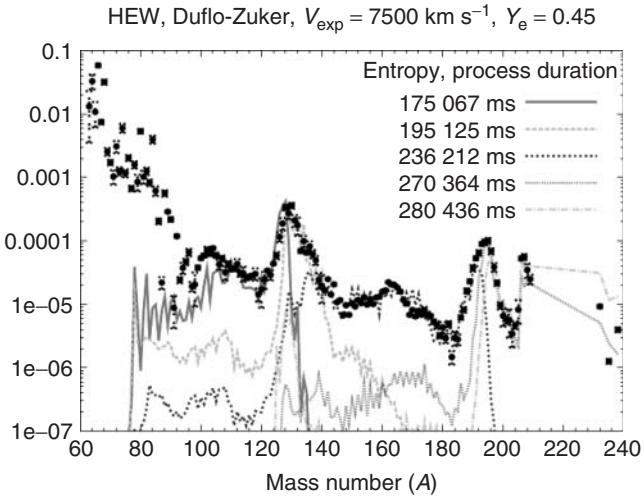


Figure 15.9 High-entropy neutrino wind results for expansion parameters and proton/nucleon ratio Y_e as given in the label, for a variation in extropies per baryon and k_B.

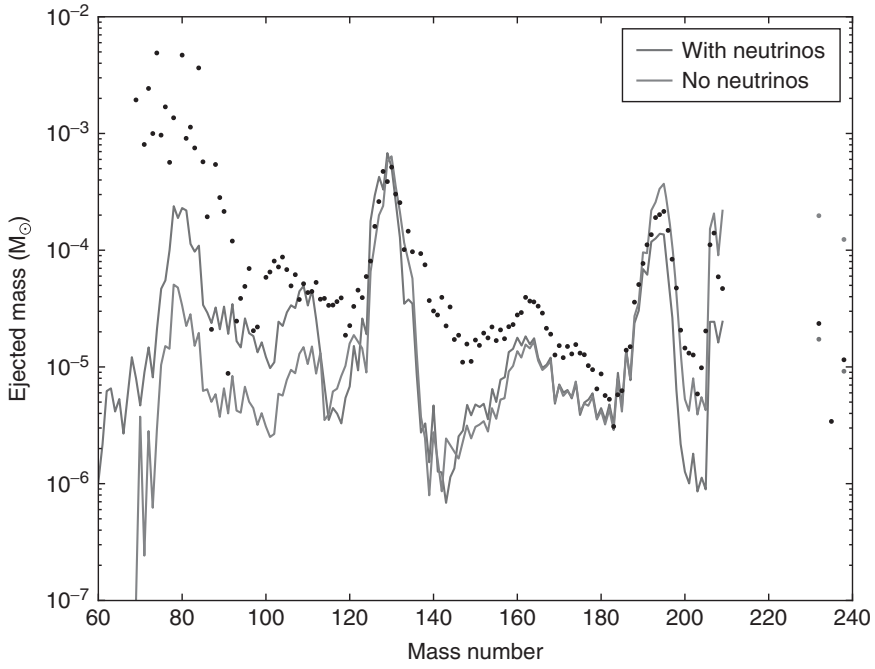


Figure 15.10 Integrated mass fractions of jet ejecta with (slightly higher abundances in the range $A = 50 - 110$) and without (higher abundances beyond $A = 190$) late neutrino heating. Black dots represent solar r-process abundances scaled to fit abundance predictions at $A = 130$.

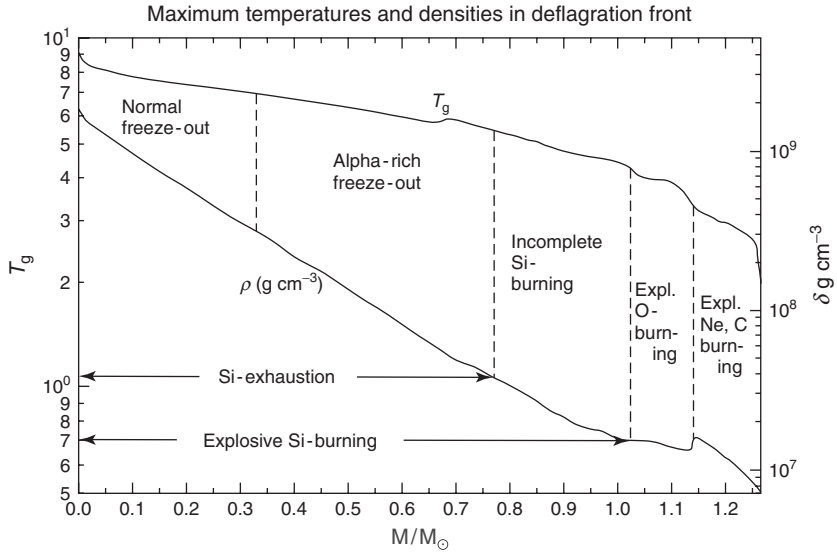


Figure 15.11 Maximum densities and temperatures during the passage of the burning front and expected nucleosynthesis processes (from the W7 model of Nomoto *et al.* [24]).

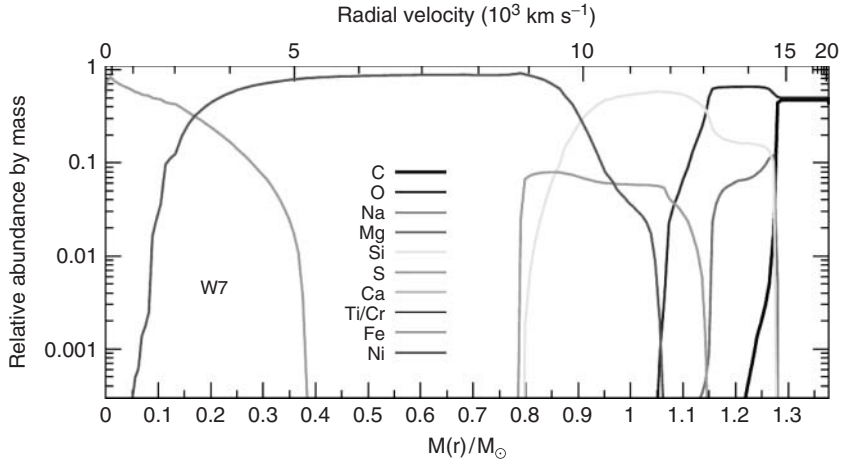


Figure 15.12 Major abundances (indicated in grey scale), resulting during the passage of the burning front for the W7 model of type Ia supernovae by Nomoto *et al.* [24]. The ejecta are dominated by Ni (^{56}Ni) from about 0.15 to 0.9 M$\odot$. The vanishing of $N = Z$ ^{56}Ni in the center and emergence instead of ^{56}Fe is due to electron captures and neutronization (reduction of Y_e). After the explosion, about 0.6 M$_\odot$ of ^{56}Ni will decay to ^{56}Co and ^{56}Fe, which dominates the observed light curve.

systems such as the center of galaxies or globular clusters, direct white dwarf collisions can also take place, again with a similar outcome. Finally, low accretion rates of He will assemble unburned He on the white dwarf surface. Without any instabilities, this He would ignite in the form of a detonation after a relatively high mass is accumulated, causing, in its explosion, fast ejection of Ni, which is not found in observations. However, if the detonation can be ignited already at an earlier stage, an outcome that is not in conflict with observations can result. For recent reviews and observations, see [25, 26].

15.5.1 Explosion Modeling

The major question is then related to the propagation of the burning front through the exploding white dwarf. A simmering phase precedes the degenerate ignition, eventually leading to the thermonuclear runaway. The initial propagation of the burning front is driven by heat conduction, requiring a modeling resolution of the size of the electron mean free path. Hydrodynamic instabilities and mixing of hot, burned material into adjacent zones leads over to a so-called deflagration phase, that is, propagation speeds less than the speed of sound. This can turn over into a detonation, that is, propagation faster than the sound speed of the upstream matter. The mechanism for this transition is still highly debated, but observations seem to support it. The explosion model imprints the conditions experienced on the nucleosynthetic yields. In all cases, this is dominantly ^{56}Ni and a smaller amount of intermediate mass nuclei from Si to Ca. Whether the ignition occurs in the center or off-center can have an effect on the Fe-group composition, as higher densities cause higher electron Fermi energies and more electron captures, neutronizating the compositions. In the following, we concentrate only on the most relevant case of single white dwarfs approaching M_{Ch} before ignition. This case is probably responsible for about 60% of all SNe Ia observed, showing a narrow range of observational characteristics and utilized as standard(izable) candles for cosmic distance measurements.

15.5.2 Nucleosynthesis

The SN explosion phase, investigated by ''classical'' 1D, spherically symmetric models (e.g., W7, see above), has successfully explained the basic contribution of type Ia to Galactic chemical evolution, as well as basic features of observed spectra and light curves of individual SNe Ia of a ''normal'' class. Some improvement in these observational aspects has been obtained by introducing a delayed detonation. These 1D models, however, treat the propagation speed of the deflagration flame as a parameter. The deflagration flame is hydrodynamically unstable and nonsphericity is essential. Recent explosion models have intensively addressed these issues with high-resolution, multi-D hydrodynamic simulations, coupled with an appropriate subgrid model to capture turbulence effects on unresolved scales. They provide essentially ''parameter-free'' simulations for the initial deflagration stage, where only the distribution of the initial deflagration ignition sparks are uncertain.

Earlier 1D simulations have tested variations in 1D burning front approximations, which tried to mimic averaged multi-D approaches. This permitted to test the impact on nucleosynthesis and realized

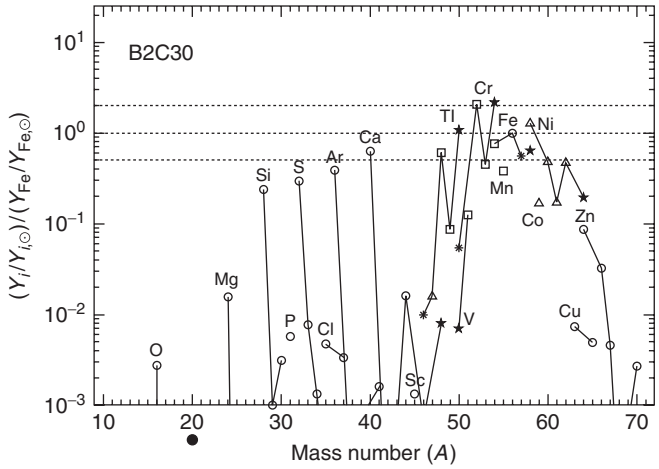


Figure 15.13 Ratio of isotopes for elements from O to Zn (after decay) to their solar values, normalized at ^{56}Fe (from Thielemann *et al.* [27]) for a model with a central ignition density of 3×10^9 g cm^{-3} and a typical burning front speed.

that slower burning fronts permit longer durations of electron captures at high densities and temperatures and thus smaller Y_es affecting the neutron richness of the central composition. Higher central densities (i.e., also electron Fermi energies) act in the same way. A similar (but opposite) effect was obtained when introducing a new set of electron capture rates from large-scale shell model calculations [6].

Dominant species in the central region with complete silicon burning affected by electron capture are stable ^{56}Fe,^{54}Fe,^{58}Ni, and radioactive ^{56}Ni (which decays to ^{56}Fe). In the case of even stronger electron capture effects, the main products are ^{50}Ti,^{54}Cr, and ^{58}Fe [27]. The resulting composition ratios to solar system abundances, normalized at ^{56}Fe, are shown in Figure 15.13 for a 1D simulation with a central ignition density of 3×10^9 g cm^{-3} and a typical burning front speed. Similar results can also be found in 2D self-consistent hydro simulation featuring slightly off-center ignitions [28]. It can be recognized that the Fe-group abundances are reproduced within a factor 2 of solar, Si, S, Ca, Ar, and Ca are underproduced by a factor 2–3, leaving these intermediate mass elements for a dominant production in core-collapse supernovae.

15.6 Summary and Outlook

Supernovae are not only stellar explosions that are as bright as whole galaxies (during maximum light), they also are responsible for the majority of elements found in nature. Core-collapse supernovae dominate in the production of O through Ti, contribute probably about 1/3 to the Fe-group nuclei, and are possibly the origin of the heaviest elements up to Th and U. The major fraction is probably responsible for elements up to $A \approx 120$, a rare fraction for the heaviest elements, including and beyond the $A = 130$ peak. SNe Ia are a ''mirror image,'' dominating the Fe-group, and contribute only about 1/3 to the

elements from Si to Ti. They provide possibly insignificant contributions to heavy elements. Both explosion mechanisms still involve open questions. In core-collapse supernovae, modern 3D MHD simulations with neutrino transport are by now successful to attain explosions via neutrino heating from the collapsed core, but with smaller explosion energies than observed. SNe Ia, based on a thermonuclear runaway explosion of white dwarfs in binary systems, exceeding the Chandrasekhar mass, might have a variety of progenitor systems leading to this end point, including white dwarf mergers and collisions. But the dominant fraction seems related to mass accretion from the binary companion until the critical white dwarf mass is attained. Here, the open questions are related to the exact ignition conditions (not possible to be obtained in so-called explicit hydrodynamic codes and thus utilized as a free parameter), the propagation speed of the burning front driven by multi-D hydrodynamic instabilities, plus the possible transition from a sub- to supersonic behavior (deflagration/detonation transition).

Glossary

Big Bang: Expansion of the observable Universe, commencing about 13.6 Giga years ago and resulting during the Big Bang nucleosynthesis in the production of $^{1-3}$H, ^{3,4}He, and ^{7}Li.

CNO-cycle: The version of H-burning, where – starting with preexisting ^{12}C – three proton captures, two β^+decays, and a final (p,α)-reaction turn 4 hydrogen nuclei into ^{4}He.

Core collapse: When contraction ends in almost free-fall infall due to pressure reduction via electron capture (of degenerate electrons) or (low-energy) particle production from high-energy photons.

Electron capture: Capture of an electron by an atomic nucleus, changing its charge, mass number (Z, A) to $(Z-1, A)$. This occurs in atoms usually by capturing an electron from inner orbits and in astrophysical fully ionized plasmas by capturing continuum electrons.

γ-ray burst, hypernova: An object with a total energy about 10-fold that of a supernova with an electromagnetic spectrum shifted toward high energies, probably related to the formation of a (stellar mass) black hole.

Neutrino wind: After core collapse to nuclear densities a (hot) protoneutron star is born with a high temperature due to the gravitational binding energy gain of the order 10^{53} erg. This energy is released and escapes in particles with the smallest interaction cross sections, neutrinos. The, although small, interaction drives matter ejection from the surface of the neutron star, the so-called neutrino wind.

Nuclear burning: The transformation of element/isotope abundances via nuclear reactions in hot astrophysical plasma, which permits to overcome Coulomb barriers of fusion reactions due to the temperature-related velocities of reaction partners. Neutron captures do not experience this Coulomb hindrance, but – due to their short half-life – are required to be produced in charged-particle-induced reactions either in stellar evolution or during explosions.

(nuclear) Equation of state: Relating thermodynamic quantities such as pressure, internal energy, entropy to density, and temperature (and composition) of an astrophysical plasma.

Nuclear statistical equilibrium (NSE): If forward and reverse reactions are in a chemical equilibrium, the abundance ratios between reaction partners are determined just by density and temperature. If all reactions across the nuclear chart are in chemical equilibrium (for sufficiently high temperatures to overcome Coulomb barriers in fusion reactions and which also provide photons with sufficient energies for photodisintegrations) a complete equilibrium materializes, which is in addition to T and ρ only dependent on the overall neutron/proton ratio of matter of Y_e.

pp-cycle: The version of H-burning, where four protons are turned into ^{4}He, starting with ^{1}H(p,e$^+\nu$)^{2}H, proceeding via ^{3}He and leading in the simplest version to the final reaction ^{3}He(^{3}He, 2p)^{4}He.

p-process: Probably a combination of two different processes (γ-process and νp-process) which produce the proton-rich stable nuclei above Fe, which are separated from their more neutron-rich isotopes by an unstable isotope, and can thus not be formed via neutron capture reactions. These two processes are related to sequences of photodisintegration reactions of stable heavy elements via temperature increase in supernova shock fronts or proton captures for the lighter elements up to $A = 100$.

Primary elements: Stars start their lives with essentially H and He. All elements that are (during stellar evolution or final explosions as for supernovae) produced from this initial H and He are produced in a primary manner, so to say, in one event from H and He. Examples are C, O, Ne, Mg, Si, S, Ar, Ca, Ti, Fe, Ni, Zn, Ge, and the r-process-dominated elements, for example, Eu or Ag, Au, Th, and U.

Quasi-equilibrium (QSE): Similar to NSE, but separated in local equilibrium clusters, linked by slow reactions that are not in chemical equilibrium.

r-process: A sequence of rapid neutron captures, producing isotopes far from stability, and – combined with short β-decay half-lives – permitting to produce the heaviest elements up to Th, U, and beyond.

Secondary elements: If prior stellar generations produced elements such as C, N, O, and Fe, the next generation will burn C, N, and O to ^{14}N in CNO H-burning. ^{14}N will be burned to ^{22}Ne in He-burning and can release neutrons in the ^{22}Ne(α, n)^{25}Mg reaction. The neutrons can, via neutron captures and β-decays, produce elements up to Pb and Bi in the s-process (slow neutron capture process). All the named elements are of so-called secondary nature and will continue to increase in their abundance with metallicity, but in a different manner from that of primary elements.

s-process: See secondary elements

Supernova: A bright object due to a stellar explosion with a total kinetic energy of the order 10^{51} erg. (observational: type I without H-lines in spectra, type II with H-lines; type Ia due to exploding white dwarfs in binary systems, type Ib,c,II due to core collapse of massive stars with extended, small, no H-envelope or even no He-envelope).

White dwarf: Stellar end stage prevented from further contraction/collapse by degenerate Fermi pressure of electrons, mostly with a C/O composition from He-burning, but He (after H-burning) and/or Ne/O/Mg (after C-burning) white dwarfs also exist.

References

1. Woosley, S.E., Heger, A., and Weaver, T.A. (2002) *Rev. Mod. Phys.*, **74**, 1015.

2. Haxton, W.C., Parker, P.D., and Rolfs, C.E. (2006) *Nucl. Phys. A*, **777**, 226.
3. Adelburger, E.G. *et al.* (2011) *Rev. Mod. Phys.*, **83**, 195.
4. Wiescher, M., Käppeler, F., and Langanke, K. (2012) *Ann. Rev. Astron. Astrophys.*, **50**, 165.
5. Thielemann, F.-K., Hirschi, R., Liebendörfer, M., and Diehl, R. (2011) *Massive Stars and their Supernovae*, LNP 812, Springer, p. 153–231.
6. Langanke, K. and Martínez-Pinedo, G. (2003) *Rev. Mod. Phys.*, **75**, 819.
7. Käppeler, F. and Mengoni, A. (2006) *Nucl. Phys. A*, **777**, 291.
8. Rauscher, T. and Thielemann, F.-K. (2000) *At. Data Nucl. Data Tables*, 75, 1; 79, 47; 88, 1.
9. Goriely, S., Chamel, N., and Pearson, J.M. (2009) *Phys. Rev. Lett.*, **102**, 152503.
10. Möller, P., Myers, W.D., Sagawa, H., and Yoshida, S. (2012) *Phys. Rev. Lett.*, **108**, 052501 66, 131.
11. Heger, A., Fryer, C.L., Woosley, S.E., Langer, N., and Hartmann, D.H. (2003) *Ap. J.*, **591**, 288.
12. Maeder, A. (2009) *Physics, Formation and Evolution of Rotating Stars*, Springer.
13. Bethe, H.A. (1990) *Rev. Mod. Phys.*, **62**, 801.
14. Martinez-Pinedo, G., Liebendörfer, M., and Frekers, D. (2006) *Nucl. Phys. A*, **777**, 395.
15. Lattimer, J. (2012) *Ann. Rev. Nucl. Part. Sci.*, **62**, 485.
16. Duan, H., Fuller, G.M., and Qian, Y.-Z. (2010) *Ann. Rev. Nucl. Part. Sci.*, **60**, 569.
17. Liebendörfer, M. (2011) *Massive Stars and their Supernovae*, LNP 812, Springer, p 439–460.
18. Janka, H.-T. (2012) *Ann. Rev. Nucl. Part. Sci.*, **62**, 407.
19. Burrows, A. (2013) *Rev. Mod. Phys.*, **85**, 245.
20. Sagert, I., Fischer, T., Hempel, M., Pagliara, G., Schaffner-Bielich, J., Mezzacappa, A., Thielemann, F.-K., and Liebendörfer, M. (2009) *Phys. Rev. Lett.*, **102**, 081101.
21. Winteler, C., Käppeli, R., Perego, A., Arcones, A., Vasset, N., Nishimura, N., Liebendörfer, M., and Thielemann, F.-K. (2012) *Ap. J.*, **750**, L22.
22. Burrows, A., Dessart, L., Livne, E., Ott, C.D., and Murphy, J. (2007) *Ap. J.*, **664**, 416.
23. Arnould, M. and Goriely, S. (2003) *Phys. Rep.*, **384**, 1.
24. Nomoto, K., Thielemann, F.-K., and Yokoi, K. (1984) *Ap. J.*, **286**, 644.
25. Mazzali, P., Röpke, F.K., Benetti, S., and Hillebrandt, W. (2007) *Science*, **315**, 825.
26. Gall, E.E.E., Taubenberger, S., Kromer, M. *et al.* (2012) *Mon. Not. Roy. Astron. Soc.*, **427**, 994.
27. Thielemann, F.-K., Brachwitz, F., Höflich, P., Martinez-Pinedo, G., and Nomoto, K. (2004) *New Astron. Rev.*, **48**, 605.
28. Maeda, K., Röpke, F.K., Fink, M., Hillebrandt, W., Travaglio, C., and Thielemann, F.-K. (2010) *Ap. J.*, **712**, 624.

16
Accelerator Mass Spectrometry and its Applications

Ragnar Hellborg, Göran Skog, and Kristina Stenström

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

16.1 Introduction

There is a need to detect very low concentrations of atoms in a sample in a number of research fields and technical applications. Several nuclear physics techniques have been developed, some of them with depth resolution and sensitivity that cannot be achieved with any other technique, either physical or chemical. In a particle accelerator, beams of fast, ionized atoms are produced. The type of ion, ion energy, intensity, and geometrical dimensions of the beam can often be chosen rather freely. This has made it possible to develop a number of analytical techniques with extremely high resolution and sensitivity, see, for example, Ref. [1].

One of the applications of nuclear physics techniques, which has been of great benefit to other fields of scientific endeavor, is accelerator mass spectrometry (AMS). The capability of AMS in extremely sensitive radioisotope measurements has been extensively demonstrated over the past nearly 35 years. For example, AMS has allowed refinements in the technique of ^{14}C dating in the fields of archaeology and quaternary geology. The most important improvement compared to traditional radiometric methods is the possibility to date small samples with AMS. Using AMS, the radiocarbon age of a 10 000 year old sample of 1 mg or less can be determined with a precision of 40 years in $<$30 min. To achieve the same precision with radiometric methods, the sample must contain more than 1 g of carbon and counting must proceed for over 24 h.

As the number of AMS facilities has grown, the number of applications has increased. ^{14}C is still the most important AMS isotope, but other applications, apart from traditional dating, are directed toward studies of atmospheric processes and ocean circulation to gain information about past climates. ^{14}C-AMS has also found applications in biomedical studies. Other isotopes such as ^{10}Be and ^{36}Cl have been used to gain hydrogeological information. These two isotopes are also trapped in the ice sheets of Greenland and Antarctica, where they can be used as tracers of the solar and geomagnetic modulation of the cosmic radiation reaching the earth. Cosmic-ray exposure also causes the buildup of ^{10}Be, ^{26}Al, and ^{36}Cl in surface rocks. The accumulated amounts of these isotopes can be measured by AMS and used to date the rock (so-called exposure dating). ^{26}Al-AMS has also been used to study metabolic processes in living systems. The production of ^{41}Ca from nuclear weapons testing has been measured by AMS, and ^{36}Cl and ^{129}I have been used to

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

trace the migration of nuclear waste from nuclear storage and reprocessing plants and nuclear power plants.

Although various types of accelerators can be used, almost all AMS systems employ the electrostatic or cascade tandem accelerator. Rare isotopes from a sample material placed in the ion source of the tandem accelerator are measured by counting individual atoms with nuclear detection techniques after acceleration up to energies in the range of 0.2–40 MeV (a few very large electrostatic tandem accelerators produce even higher energy beams). Compared to conventional mass spectrometry (MS), the dramatic improvement in background rejection in the AMS systems has, in some cases, led to a 10^8 times increase in the sensitivity of isotope ratio measurements.

AMS evolved at nuclear physics laboratories where tandem accelerators were originally installed during the 1960s and 1970s. Most of these tandems cannot produce beams with sufficient energy for today's nuclear physics experiments, and were therefore adapted for AMS owing to the need to date small samples of ^{14}C. During the late 1970s, a great deal of effort was devoted to the evaluation of tandem accelerators for AMS detection of ^{14}C. One important discovery was that ^{14}N does not form stable negative ions [2]. It was then soon shown that ^{14}C can be detected at the very low, natural isotope ratio of $^{14}C/^{12}C = 10^{-12}$ [3, 4]. The first dedicated AMS systems were installed in the early 1980s [5]. During the last 15 years (up to 2012), the best achievements have been obtained with relatively small accelerators designed for and dedicated to AMS measurements. Accelerators of types other than tandems have been used in only a limited number of investigations. This is more described in Ref. [6].

AMS is an expensive technique and it is also technically complicated. Therefore, the most significant technical development during the last 15 years is the trend toward smaller and simpler AMS systems. The reduction in floor space (down to 2.5×3.0 m^2!) and overall cost (down to 1 M€) are the most attractive features of these new systems, as well as the smaller technical staff required to run the systems. In fact, some of these new AMS systems can be operated by general laboratory staff as the system is designed to be robust and user friendly [7].

The development of AMS has been reviewed in many articles, for example, in Ref. [6]. The first 10 years of development of AMS has recently been reviewed by Gove *et al.* [8]. Every third year, an international conference on AMS is organized. The most recent one was held in Wellington, New Zealand in March 2011 [9]. The next one will take place in Aix-en-Provence (August 2014). Another series of conferences related to AMS are the International Radiocarbon Conferences. The most recent one was held in Paris in July 2012.

The number of AMS laboratories around the world, completely or partially dedicated to AMS, is about 100. Several of these AMS accelerators are listed in a table in Ref. [6]. Two commercial companies today offer complete AMS facilities: High Voltage Engineering Europe (HVEE) in Amersfort, the Netherlands, and National Electrostatic Corporation (NEC) in Middleton, Wisconsin, United States.

The instrumentation and general principles of AMS are described in Section 16.2. The trend toward smaller accelerators is outlined in Section 16.3. Some examples of the application of AMS are presented in Section 16.4. Finally, in Section 16.5, conclusions and future perspectives are discussed.

16.2
General Principles and Instrumentation

16.2.1
Principles of AMS

The two standard methods used for a long time to determine the isotopic composition of an element are MS and decay counting. While MS can be used for all isotopes, decay counting is restricted to radioisotopes. For long-lived radioisotopes, decay counting is inefficient because only a small fraction of the nuclides in a sample decays during a reasonable measurement time. Conventional MS has a high efficiency but it is limited to isotope ratios greater than about 10^{-7}. In AMS, the efficiency of MS is combined with extremely good discrimination against isobaric, isotopic, and molecular interference.

In Figure 16.1, the efficiencies of decay counting and AMS are shown as a function of half-life ($T_{1/2}$). The efficiency is defined as the number of atoms detected compared to the number in the sample. The counting time for AMS is assumed to be 1 h, while the efficiency for decay counting periods of 24 h and 14 days are shown in the figure. To obtain a statistical precision of 0.5% (which is often required in radiocarbon dating) using 1 g of carbon, it is necessary to count the decays for more than 48 h. For AMS, a sample of <1 mg carbon is sufficient.

The high discrimination of AMS is obtained by accelerating the ions to a high energy, usually employing an electrostatic or cascade tandem accelerator. For a detailed description of electrostatic and cascade accelerators, see, for example, [10]. In a few cases, other types of accelerators are used; this is further outlined in Ref. [6]. A typical AMS system includes the following:

- production of negative ions in a multisample, negative-ion source: the sample material containing the rare isotopes to be counted is placed in the ion source;
- first mass selection in the low-energy section of the accelerator system;

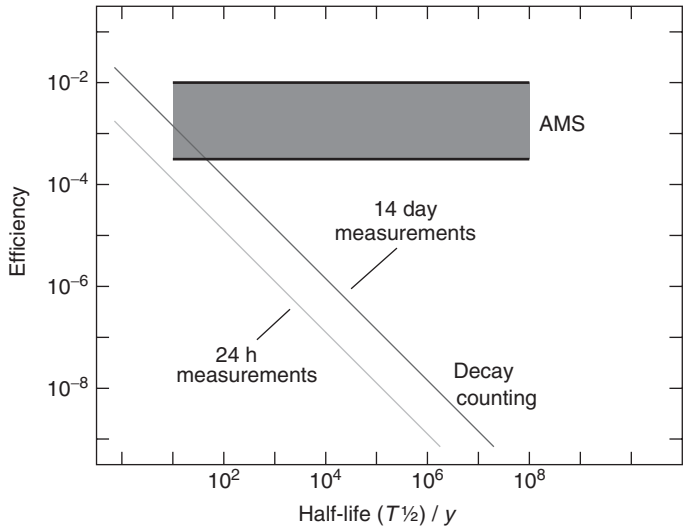


Figure 16.1 The efficiencies of AMS and decay counting as a function of half-life. For details, see the text.

- acceleration of negative ions from ground potential to a high positive potential;
- recharging of all ions to positive by stripping off electrons and, at the same time, dissociation of all molecular ions;
- acceleration of the now positive ions back to ground potential (exceptions, see Section 16.3);
- the removal of unwanted ions using electric and magnetic fields;
- identification and counting of the individual rare isotopes with nuclear detection techniques;
- computer control of the accelerator system to allow for unattended operation and to provide control of the AMS system parameters.

AMS is an extension of MS including an accelerator. In Figure 16.2, an MS system is compared with a simple AMS system. The introduction of the tandem accelerator, followed by several ion-filtering devices reduces the background by a factor of the order of 10^8. Three especially important characteristics of AMS allow the measurement of low isotope ratios (e.g., for $^{14}C/^{12}C$ down to 10^{-15}):

- Interference by some isobars is avoided by using a negative-ion source (e.g., $^{14}N^-$, $^{26}Mg^-$, $^{55}Mn^-$, and $^{129}Xe^-$, all with negative electron affinity, when detecting ^{14}C, ^{26}Al, ^{55}Fe, and ^{129}I).
- Interference by molecules is avoided by using a stripper system in the high-voltage terminal of the accelerator (e.g., $^{12}CH_2$ and ^{13}CH when detecting ^{14}C). In the stripping process, the negative ions become positive, and at the same time, nearly all the molecular ions will be dissociated. A few molecular ions with a maximum charge state of 2+

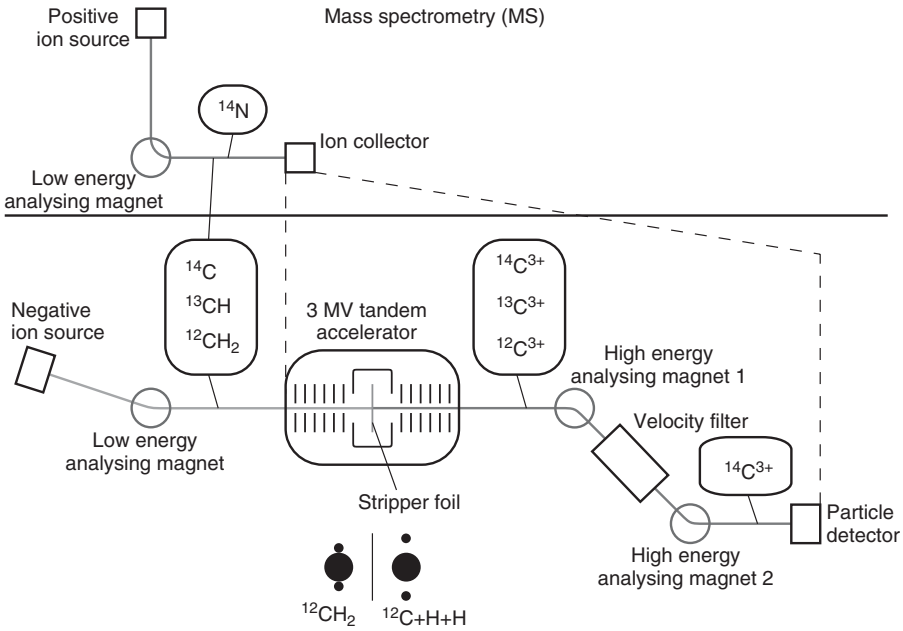


Figure 16.2 Comparison between MS and a typical AMS system.

will survive. Molecular ions can thus be avoided by choosing a charge state higher than 2+ with the high-energy analyzing system after the accelerator.
- The counting of individual ions is possible owing to their high final energy. The ions that reach the detector are easily separated by their energy difference.

The various components of an AMS system are described in more detail in [6].

16.2.2 Traditional AMS Instrumentation

16.2.2.1 The Ion Source

In a tandem accelerator, negatively charged ions gain energy by attraction to the high positive voltage at the geometric center of a pressure vessel. Thus, a negative-ion source is needed. The requirements on such a source designed for AMS are a high, stable ion beam current and a high efficiency for negative ions. The source should also be equipped with a multiple sample holder for fast switching between different samples. The cross-contamination or ''memory effect'' between successive samples should be negligible. A source working with solid samples, fulfilling these requirements, is the cesium negative-ion source, which was introduced around 1970 [11]. The principle of such a source is that cesium ions of a few kiloelectronvolts are focused on the surface of a solid sample and enough energy is transferred to the target material to produce free atoms and ions of the sample material. This process is called *sputtering*. For more technical details about the sputtering process and ion sources, see Refs. [6, 12].

Certain amounts of the Cs vapor will condense on the sample and produce a thin surface layer. The advantage of Cs (and some other group IA elements) over other heavy elements, which could also be used for sputtering, is that this surface layer will considerably reduce the work function (the energy difference between a free atom and an atom bound at the sample surface).

Some examples of the type of sample material used for different isotopes, the type of ions produced, the corresponding stable isotope, and typical beam currents for the stable isotope are given in Table 16.1.

A sputtering source fed with a gaseous CO_2 sample, as an alternative to solid graphite samples, was introduced at the AMS laboratory in Oxford, United Kingdom, a number of years ago [13]. Today, a number of AMS laboratories have reported the use of sputtering sources using CO_2

Table 16.1 Sample material, rare and stable isotopes, and current of stable isotopes for a few isotopes used in AMS.

Isotope	Sample material	Rare ion	Stable ion	Current of stable ion (μA)
^{10}Be	BeO	$^{10}Be^{16}O^-$	$^{9}Be^{16}O^-$	5–10
^{14}C	Graphite	$^{14}C^-$	$^{12}C^-$, $^{13}C^-$	30–100
^{26}Al	Al_2O_3	$^{26}AlO^-$	$^{27}AlO^-$	1–5
–	Al_2O_3	$^{26}Al^-$	$^{27}Al^-$	0.1–1
^{36}Cl	AgCl	$^{36}Cl^-$	$^{35}Cl^-$	10–20
^{59}Ni	Ni	$^{59}Ni^-$	$^{58}Ni^-$	1–5

(see, e.g., Fahrni *et al.* [14] and references therein). The advantage of using gaseous carbon samples is twofold. A much simpler sample preparation procedure (Section 16.2.2.3) can be used, as the second step of reducing the gas to solid carbon is unnecessary, and very small samples can be analyzed. The disadvantage of gaseous material is a lower beam current and a higher risk of cross-contamination between different samples. An overview of different types of ion sources for AMS can be found in Ref. [12].

16.2.2.2 Analyzing and Detecting Systems

The ion source is followed by an injector, a low-energy analyzing system, a tandem accelerator, a high-energy analyzing system, and, finally, a detector system. A detailed discussion of these different components can be found in an overview article presented a few years ago [6].

A number of different types of detectors, originally developed for nuclear physics research, are used to count the rare isotopes in AMS experiments. The most common types of detectors, such as gas ionization detectors, surface barrier detectors, time-of-flight detectors, gas-filled magnets, and X-ray detectors, are briefly described in a separate article [12].

16.2.2.3 ^{14}C Sample Preparation

^{14}C measurements using AMS require that the sample be transformed into elemental carbon in order to produce a stable ion beam with negligible memory effect in the ion source. In general, the samples follow three steps: (i) pretreatment, (ii) combustion to CO_2, and (iii) reduction to graphite.

Before combustion of the organic samples to be dated, they are first chemically prepared to remove any contaminants. The three-step acid–alkali–acid (AAA) method is the most common. The sample is washed in hot, diluted HCl, followed by hot, diluted NaOH solution. The NaOH may absorb CO_2 from the surrounding air. The final HCl wash ensures that any such contamination is removed. Between two successive steps, the sample is rinsed to neutral pH with deionized water. A more detailed overview of the pretreatment procedure can be found in Ref. [15].

After the pretreatment, the samples are combusted and converted to CO_2. Combustion using copper(II)oxide as oxidizing agent can be performed online with the rest of the graphitization system (see, e.g., Ref. [16]), or in sealed tubes (see, e.g., Ref. [17]). An alternative combustion is achieved using an elemental analyzer [18]. For carbonate samples (e.g., mollusks and foraminifers), the carbon dioxide is released by hydrolysis using phosphoric acid. In the form of CO_2, the sample is transferred to a small reaction volume for graphitization [19]. The transfer is usually achieved cryogenically. In 2010, Wacker, Němec, and Bourquin [20] introduced a new automated combustion and graphitization system where the CO_2 coming from an elemental analyzer is absorbed on a single column filled with zeolite. The CO_2 can then be easily released by heating the zeolite trap and transferred to the reactor by gas expansion.

The graphitization process is still being improved and methods to reduce the sample size and to minimize background levels are being developed [18]. Another important aspect is high throughput in sample preparation processes, especially when using AMS for quantitative isotope ratio analysis in the biosciences, where the number of measurements is often very high and the time available for each AMS analysis is short. A method for the rapid production of graphite from biochemical samples has been developed and is

described in Ref. [17], and a refinement of the technique was presented by Ognibene, Bench, and Vogel in 2003 [21].

16.3 The Trend Toward Smaller Accelerators

As has been mentioned above, in the case of ^{14}C, it is known that its isobar, ^{14}N, does not form stable negative ions. The remaining problem is therefore the destruction of hydrocarbon molecular ions such as $^{13}CH^-$ and $^{12}CH_2^-$.

Until the middle of the 1990s, the method used to eliminate molecular background in AMS was to allow the ions to undergo charge exchange at high energy (>2.5 MeV) in the high-voltage terminal of a tandem accelerator, to create charge states of 3+ or higher, where molecular bonds are no longer stable. An AMS facility accelerating ions to several megaelectronvolts has high investment and running costs due to the size of the installation. The possibility of using an accelerator of lower voltage and a lower charge state after charge exchange would thus make the AMS facility both smaller and cheaper. The pioneering work in this field was done by the IsoTrace group in Toronto, Canada [22], but it was not until the end of the 1990s, when it was realized through the work of the ETH group in Zurich, Switzerland, that it was possible to carry out radiocarbon dating using charge states lower than 3+.

In 1997, Suter *et al.* [23] discussed the use of a 0.5–1 MV AMS accelerator, which would operate with the 1+ or 2+ charge state. They laid the groundwork for this development with studies showing that molecular interference could be removed with a higher stripper gas pressure than previously used. Later, they reported the design of a prototype 0.5 MV accelerator that would use the 1+ charge state [24]. These studies resulted in an original design of a 0.5 MV accelerator, constructed at ETH in Zurich in cooperation with NEC [25].

A number of commercial instruments based on these developments are now on the market. NEC offers a compact AMS system based on a 0.5 MV Pelletron accelerator as reported in Ref. [26]. Klein *et al.* [27] report on a 1 MV, compact multielement AMS system, built by HVEE and comparable in size to the NEC 0.5 MV machine. The authors claim that this system, apart from high-precision, low-background measurements of ^{10}Be, ^{14}C, and ^{27}Al, is also capable of measuring iodine and plutonium. The first of these HVEE machines is in operation at Seville in Spain [28] and a second system is now in operation in Trondheim, Norway.

Introducing a second magnet in the high-energy analyzing beam line, Suter *et al.* [29] have shown that their compact AMS system (TANDY) based on the NEC 0.5 MV prototype accelerator can be used for heavy ions such as iodine as well as the actinides. Initial tests indicated that the background from neighboring isotopes is now significantly reduced and, in most cases, not a limiting factor any more. Surprisingly good performance has been obtained. Especially for actinides, transmissions are higher than those at any of the large AMS facilities. The best large facilities have significantly higher transmissions for ^{10}Be and also ^{26}Al. ^{36}Cl, another important cosmogenic radionuclide, cannot be measured at natural concentrations with low-energy AMS. Compact facilities therefore cannot fully replace larger instruments.

Striving for cheaper and smaller AMS systems, it has become evident that voltages even below 200 kV are feasible in some applications. In 2004, Synal *et al.* [30]

presented the results of some test experiments on a 200 kV tandem accelerator. They reported that a background $^{14}C/^{12}C$ ratio of $0.4 - 1.3 \times 10^{14}$ could be obtained with the instrument, suggesting that it could be useful for radiocarbon dating.

Acceleration using only a few hundred kilovolts does not require that the high voltage is enclosed in a pressure tank. The high voltage can instead be sustained in air, promoting the development of open-air AMS systems. Schroeder *et al.* [31] have presented the design of an open-air system, with only one standard accelerating tube section. This technique is called *single-stage* accelerator mass spectrometry (SSAMS), alluding to the fact that the two-step accelerator has been replaced by a single acceleration stage. It operates at a maximum voltage of 250 kV. The stripper is maintained at a high-voltage and high-energy MS is performed following the stripper without any further acceleration. The first accelerator of this type built by NEC was installed at Lund University in 2004 and to date nine other accelerators of the same type have been installed worldwide and two more are under construction (autumn 2012). This instrument gives backgrounds similar to the 0.5 MV machines discussed previously [32, 33] and has proved to be a useful tool for radiocarbon dating [34, 35]. A detailed description of the Lund SSAMS system can be found in Refs. [6, 35]. Data on the typical system performance for the SSAMS in Lund are given in Table 16.2.

Transmission is defined as the ratio between the $^{12}C^+$ current after the high-energy bending magnet and the $^{12}C^-$ current after the low-energy bending magnet. BP = before present, which means before AD 1950.

In 2007, a prototype radiocarbon dating system (MICADAS), which was based on a vacuum insulated accelerator with an acceleration voltage of 200 kV, was presented [36]. This small accelerator, with dimensions of only 2.5×3 m^2, was shown to fulfill the requirements for radiocarbon dating applications. Later Wacker *et al.* [37] have shown that this instrument is capable of high-precision radiocarbon measurements that goes beyond the precision of large accelerators.

The most recently proposed development toward smaller systems actually excludes the accelerator. Synal *et al.* [38] have demonstrated in a proof-of-principle experiment that radiocarbon detection is feasible using a pure mass spectrometer system at ion energies below 50 keV when using He as stripper gas.

Table 16.2 Performance data at the Lund SSAMS for standard measurements.

Parameter	Numerical values
Ion source output	20 μA of $^{12}C^-$
Transmission[a]	44% using He as stripper gas
Background (chemical blank)	0.25% of modern standard
Background (processed anthracite)	0.4% of modern standard
Dating precision (0–8000 BP)	0.5% or ∼ ±40 radiocarbon years BP

16.4 Applications

16.4.1 Calibration of the Radiocarbon Timescale

The radiocarbon method is based on the rate of decay of the radioactive carbon isotope ^{14}C, which is formed in the upper atmosphere through the reaction

between cosmic-ray neutrons and ^{14}N: $n + {}^{14}N \rightarrow {}^{14}C + p$, where n is a neutron and p is a proton. The ^{14}C atoms are rapidly oxidized to $^{14}CO_2$ molecules, which are taken up by plants. ^{14}C is further transported to animals through the food chain. Thus, plants and animals take up radioactive carbon while alive, but cease to do so when they die. However, the steady decay of ^{14}C in their tissues continues over the years. The ^{14}C atoms decay back to ^{14}N, with a half-life of circa 5700 years, according to the reaction: $^{14}C \rightarrow {}^{14}N + \beta$, where β is an electron. The detection limit of ^{14}C for an AMS facility is 1–2‰ of modern carbon (corresponding to a $^{14}C/^{12}C$-ratio of around 10^{-15}), which sets an upper limit for the radiocarbon method to slightly above 50 000 years.

The intensity of the cosmic radiation on earth varies with the sun's solar activity and fluctuations in the earth's magnetic field, and thus the production of ^{14}C in the atmosphere also varies. Moreover, the uptake of carbon dioxide in the oceans varies with changes in the ocean ventilation rate when the earth's climate changes rapidly. This means that the radiocarbon level in the atmosphere changes according to a complex pattern. As a result, a raw radiocarbon date may correspond to several possible calendar dates and may diverge from real calendar years by hundreds or even thousands of years. Thus, the radiocarbon clock must be calibrated to account for these fluctuations. For example, the start of the Holocene period, the period in time when the last ice age ended, is dated to 10 000 radiocarbon years ago. But the start of the Holocene, when agriculture began, cannot be determined with better precision than about 600 years using the radiocarbon method alone. Using the tree-ring radiocarbon calibration curve (see below), this transition has been shown to occur somewhere between 11 800 and 11 200 years ago.

Recent research to determine the pattern of fluctuation of the ^{14}C content of the earth's atmosphere has led to better and more detailed calibration model for the past 50 000 years. Using ^{14}C data from tree rings, corals, lake sediments, ice cores, and other sources, the radiocarbon community has now created a detailed record of ^{14}C variations over the millennia and extended the "official" radiocarbon calibration curve 50 000 calendar years back in time (IntCal09) [39]. Of these data records, the terrestrial tree-ring curve, is the most accurate and precise. The tree-ring curve is based on several millennia-long chronologies, providing an absolute time frame within the possible error of dendrochronology. These long chronologies come from wood from Europe and North America, see, for example [40, 41]. The oldest part of the tree-ring chronology has been constructed from German pine, and has been successfully linked to the younger German oak chronology. The tree-ring chronology from central Europe now extends back to circa 10 640 BC [42, 43].

Beyond the tree-ring curve, the calibration data sets rely mainly on marine samples [44]. Although there is a discrepancy between the different data sets in the time interval 26 000–50 000 radiocarbon years BP (before present), they all show that the calendar timescale sets the radiocarbon clock back several thousand years in this interval. The data sets come from radiocarbon dating of a sequence of deep-sea sediments in the Cariaco Basin near Venezuela, and from a similar sequence of dating from deep-sea sediments adjacent to the Iberian coast. These sequences were translated into a "calendar" scale by reference to patterns of oxygen isotopes ($^{18}O/^{16}O$) in the Greenland ice core records. The data

sets also include series of radiocarbon and uranium/thorium (U/Th) measurements on fossil coral formations from the tropical Atlantic and Pacific [45, 46, 47]. However, despite detailed continuous records, there remain uncertainties of how well marine radiocarbon records represent atmospheric ^{14}C content [48].

A new core was recently recovered from the varved sediments of Lake Suigetsu, Japan [49]. By dating terrestrial macrofossils (Section 16.4.3) in the varved sediment, it should be possible to generate a terrestrial radiocarbon model beyond the IntCal09 tree-ring limit and back to the radiocarbon detection limit.

The ultimate objective, which is to extend the tree-ring calibration curve beyond 50 000 years, may never be achieved and lies certainly several decades ahead. The most promising candidate, known at present, to realize this goal is the kauri trees buried in bogs in northern New Zealand [50].

16.4.2
Calibration with the ^{14}C Bomb Peak

Nuclear weapon testing in the atmosphere caused in the 1960s almost a doubling of the ^{14}C activity in the atmosphere, see Figure 16.3. After cessation of these nuclear tests in the atmosphere, the peak started to decrease through interaction of the atmosphere with the other carbon reservoirs. Today (2012) the remaining ''excess activity'' is around 4%. Samples originating from the time period after circa 1955 can be radiocarbon dated utilizing the ^{14}C bomb peak as a calibration curve. Clearly, the ^{14}C-bomb peak can be used to retrieve very precise dates for certain samples (within one year at the steepest part of the curve, that is, during the 1960s and 1970s) [51]). A detailed overview of how radiocarbon can be used to date events in the recent past is given in Ref. [52].

The bomb peak has also been used to determine the turnover time for cells in the human body, and the bomb tests offer a unique opportunity to study various processes involved in the global carbon cycle. It had already been realized at the beginning of the nuclear weapons test era that bomb-^{14}C could be employed for replacement rates for human tissue [55–57]. Since then, the technique has been used in a number of studies using decay-counting techniques as well as AMS. Mok *et al.* [58] used gas-proportional counting (GPC) to show that human gallstones begin to form about 10 years before they give rise to symptoms. This technique required 0.3 g of material per sample to generate a sufficient amount of CO_2 to analyze for ^{14}C. The entry of AMS with the possibility of using samples containing less than a few tens of micrograms of carbon, have stimulated further applications of the technique [59]. It has been shown that the formation of human cortical neurons [59] as well as the eye lens [60] almost entirely takes place around the time of birth. Other studies have shown that about 10% of all fat cells [61] and 0.5–1% of all of cardiomyocytes [62] are turned over annually in humans. Gonçalves *et al.* [63] used the ^{14}C bomb pulse to determine that the biological age of human atherosclerotic plaques, which are constrictions of the blood vessels that arise when fatty materials are accumulated in the artery walls. Atherosclerotic plaques are the main underlying cause of heart attacks and strokes. The study showed that the plaques develop slowly: the age of the plaque components varied between 5 and 15 years of age, with the cap being the youngest.

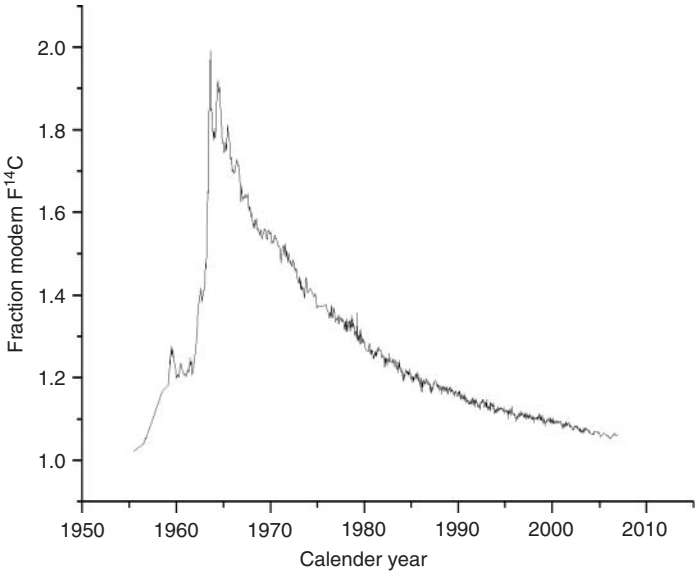


Figure 16.3 The bomb pulse, that is, the atmospheric concentration of ^{14}C in the northern hemisphere. ($F^{14}C$ concentration is approximately equal to 1.0 before the nuclear weapons tests.) ([53, 54].)

16.4.3 Dating of Lake Sediments

The common method of dating in quaternary geology and paleoclimatology studies is to use the archives of lake sediments. An important question is the chronology of the cores that are taken from the lakes. Earlier studies were based on large, bulk sediment samples, mainly consisting of gyttja (as much as 100 g of wet material or 5 g of dry material were common). Gyttja is made up of organisms that have lived in the lake and have taken their carbon from the lake water. The lake reservoirs contain carbon from many different sources: dissolved limestone, humic acid from soil and peat in the surroundings, or carbon dioxide from ''old'' groundwater reservoirs. Of course, a large part also comes from carbon dioxide exchanged with the atmosphere at the lake surface. However, the radiocarbon level in the lake does not reflect the contemporary atmosphere and, therefore, we have the problem of the so-called lake reservoir age, which gives older dates than expected from the atmospheric record.

Before AMS was established, the problem of ''hard-water'' lakes was well known and dating lakes with limestone in the surrounding bedrock was avoided. However, AMS dating showed that the reservoir problem was also severe for ''soft-water'' lakes. Instead of dating whole bulk samples, it was now possible to date small macrofossils from the vegetation around the lake, samples that reflect the atmospheric level of radiocarbon rather than the level in the lake. This has been a very important development for quaternary geologists and has led to completely new sampling techniques.

This is clearly demonstrated in an investigation that partly deals with the

tree limit in ancient times in the Scandes Mountains in northern Sweden and Norway [64]. The radiocarbon chronology in this investigation was very carefully determined, both with terrestrial macrofossils found in the cores and with bulk gyttja samples. The investigation was performed in a ''hard-water'' lake (Lake Tibetanus) and a ''soft-water'' lake (Lake Voulep Njakajaure). The results of this investigation and other studies, for example, [65], clearly demonstrated the importance of selecting the correct material for radiocarbon dating. The terrestrial macrofossils gave younger ages than the bulk samples in all cases. The reservoir age in Lake Tibetanus can be estimated to be 1000–2000 years, and furthermore it varies with time. In Lake Voulep Njakajaure, the reservoir effect is less, but there is still a reservoir age of about 500 years. This is quite typical for Swedish ''soft-water'' lakes [66].

Given the imprecision of individual calibrated radiocarbon determinations, it is often useful to use information from the deposition process to refine the chronology [67, 68, 69, 70].

Radiocarbon dating of macrofossils and pollen is not possible without the AMS technique, and it is therefore relevant to talk of a ''second radiocarbon revolution'' in quaternary geology after the introduction of ^{14}C dating with AMS. An excellent overview of the application of radiocarbon in quaternary studies is given in [15].

16.4.4 Atmospheric ^{10}Be

The AMS technique has had considerable influence on radiocarbon dating, but it has had an even greater impact in other fields of quaternary geology, which use the other cosmogenic isotopes. Before 1990, cosmogenic nuclides apart from ^{14}C were more or less unknown to quaternary geologists. This was simply due to the fact that there was no practical way to measure the extremely small concentrations of these cosmogenic nuclides produced in the geological archives. ^{10}Be, the most important cosmogenic nuclide apart from ^{14}C has a half-life that is long enough (1.6 million years) to cover the whole quaternary period. The best archives for ^{10}Be are the marine sediments and the inland ices of Greenland and Antarctica. One kilogram of ice contains about 50 million ^{10}Be atoms. Concentrated to a 1 mg substrate, this is an excellent sample for an AMS measurement. Decay counting of ^{10}Be, on the other hand, is virtually impossible due to its long half-life. All the ^{10}Be nuclides in 1 ton of ice would only give about 65 counts in 24 h. Thus, without AMS, the ^{10}Be signal would have remained hidden in the archives.

The ^{10}Be concentration in ice and marine sediments is an indicator of solar fluctuations, which in turn control the cosmic-ray flux that reaches the earth from outer space. The more active the sun, the less intensive the cosmic-rays on earth, and vice versa. During the period 1660–1740 AD, there is a distinct peak in the ^{10}Be record [71, 72]. It is known from visual observations that during this period there were hardly any sunspots on the solar surface and the sun remained quiet. Consequently, the cosmic-ray intensity increased, as did the ^{10}Be signal. The same is also true for other cosmogenic isotopes such as ^{14}C. The ^{14}C signal is, however, more difficult to interpret as it also depends on the ocean ventilation rate, which determines the amount of carbon dioxide that is dissolved in ocean waters.

By combining data from marine records of the varying production rates of ^{10}Be with data on the variations in the earth's

magnetic field intensity, it has proved to be possible to calculate the variations in the solar magnetic activity 200 000 years back in time [73]. This study indicated that the glacial and interglacial periods on earth during the past 200 000 years appear to be strongly linked to solar activity (see also [74, 75]). The amount of ^{10}Be was compared to the oxygen isotopic record, which is closely related to the global temperature. The two records follow each other very closely, which is a strong indication that small solar variations may have a great impact on the climate.

In annually laminated sediments from a German lake, Martin-Puertas *et al.* [76] have found a simultaneous sharp increase in windiness and cosmogenic ^{10}Be deposition around 2800 years ago. They conclude that changes in atmospheric circulation amplified the solar signal and caused abrupt climate change about 2800 years ago, coincident with a grand solar minimum. The range of variability of ^{10}Be and ^{14}C records also allows us to set constraints about long-term changes in solar activity. These records indicate that present solar activity levels were reached or exceeded regularly in the past [77].

Evidence of enhanced ^{10}Be deposition deep in the Antarctic ice has been reported [78]. The authors interpret this as a result of the low dipole field during the Matuyama–Brunhes geomagnetic reversal, which occurred about 780 000 years ago. If this is correct, it will be an important time marker connecting ice cores, marine cores, and radiometric timescales.

16.4.5 Exposure Dating

When a bedrock surface or a sediment is exposed to cosmic-rays, a buildup of cosmogenic nuclides will occur within minerals in the uppermost few meters below the surface of the rock. Measuring the concentrations of these nuclides allows determination of the time period for which rocks or sediment have been exposed. The most widely used of the cosmogenic nuclides are the radioactive nuclides ^{14}C, ^{10}Be, ^{26}Al, and ^{36}Cl and the stable noble gases ^{3}He and ^{21}Ne. The ability of AMS to measure low concentrations of rare cosmogenic nuclides has led to new methods of addressing long-standing geological questions, and has provided new insights into the rates and types of surface processes. Their different physical and chemical properties make it possible to apply surface exposure dating methods to rock surfaces of virtually any lithology at any latitude or altitude, for exposures ranging from the late Holocene to the Pliocene (>2.65 million years). The terrestrial in situ cosmogenic nuclide method is beginning to revolutionize the manner in which landscape evolution is studied. Single or multiple nuclides can be measured in a rock surface to obtain erosion rates on boulder and bedrock surfaces. A particularly interesting system is the ^{10}Be–^{26}Al pair that is produced in quartz, where ^{26}Al is mainly produced from ^{28}Si and ^{10}Be is mainly produced from ^{16}O. The in situ produced cosmogenic nuclides from continental and mountain erosion records have been used to reconstruct quaternary ice volume variations. The *in situ* methods have also been used in significant breakthroughs in establishing the rates and types of local- and large-scale erosion, soil development, and landscape evolution. A complete review of the subject is given in Ref. [79]. A recent review, which focuses on the applications of the method, is given in Ref. [80].

16.4.6 Recent Developments in Radiocarbon Dating and Archaeology

Since its introduction in the early 1950s, the radiocarbon dating method has been an essential tool for archaeologists. Collecting charcoal and bones for dating during excavations is standard practice, and systematic dating for research purposes is also becoming more and more common. With AMS, it is possible to date short-lived material such as a single grain of wheat or a blade of grass. A small piece of bone of subgram weight [81] and also charred bones [82] can now be dated, which was previously impossible. Analyzing more than one sample from a site or object can give indications of contamination, improving the reliability.

Until quite recently, cave paintings were dated according to stylistic criteria loosely associated with dates obtained from archaeological remains found in the vicinity of decorated surfaces. Advances in radiocarbon dating with AMS now make it possible to date prehistoric cave paintings by sampling the pigment itself. The ages obtained from paintings decorating two French caves at Cosquer and Chauvet have so far shown that the art of cave painting appeared early in the Upper Paleolithic period, which is much earlier than previously believed. The high artistic quality of the earliest paintings underlines the importance of absolute chronology in any attempt to study the evolution of prehistoric art. Prehistorians, who have traditionally interpreted the evolution of prehistoric art as a steady progression from simple to more complex representations, may have to reconsider existing theories of the origins of art. In the Chauvet caves, which consist of several chambers, radiocarbon dates of between 29 700 and 32 400 years BP have been obtained for submilligram charcoal samples [83, 84].

Radiocarbon dating has been applied to the study of modern human origins and dispersal in Eurasia [85]. Developments involving ultrafiltration of the prepared gelatin samples derived from bone collagen to separate out the smaller and lower-molecular-weight fractions [86] have led to radical improvements in the procedures for the effective purification of bone collagen to eliminate contamination by more recent carbon. Removing recent contaminants is of special importance in older bone samples, which have always provided the most widely available materials for dating from early human sites. Applications of this procedure have led to dates that are frequently between 2000 and 7000 years older than the original age estimates [87].

These new developments are of crucial importance when trying to determine the time of the extinction of the Neanderthals in Europe. Recently, an international team reported several radiocarbon dates from Gorham's cave in Gibraltar [88]. The dates cluster at about 28 000 years raw "radiocarbon years," indicating that the Neanderthals survived much longer than previously thought. It is believed that the Neanderthals took refuge in southern Europe where the environment was favorable and where modern humans were still fairly rare.

Art and human history are of general interest and therefore the dating of some objects by AMS has given rise to considerable publicity. The best-known cases are the dating of the Turin Shroud [89], the Iceman in Ötztal [90], and the Dead Sea Scrolls [91]. A ceremonial book from the earliest period of Christianity in Sweden was dated at the Lund AMS laboratory. The book, originating from

around AD 1100, is regarded as the oldest book in Sweden [92].

16.4.7 Environmental Science

Radiocarbon has proved to be an excellent tool in studies of the global carbon cycle, that is, the complex interactions between the atmosphere, oceans, land and living organisms. Studies of pathways and dynamics of carbon exchange between different reservoirs are made possible because the $^{14}C/C$ ratio, and also the $^{13}C/^{12}C$ ratio, differs among the specific compartments of the carbon cycle. Natural as well as anthropogenic processes, in the past and at present, can be studied. Understanding the natural climate processes and changes in the past is essential to assess more recent anthropogenic influences.

One example of immediate importance and interest is how mankind contributes to the global warming of earth, also known as the *greenhouse effect*. The large-scale combustion of fossil fuels since the late nineteenth century is believed to be one of the major contributions to the warming effect. As fossil fuels originate from very old dead organic matter, these are essentially ^{14}C-free, which is taken advantage of in tracer studies of the global carbon cycle. Furthermore, the $^{13}C/^{12}C$ ratio of fossil fuels differs from that of the preindustrial atmosphere, and thus, stable isotope measurements using the technique of isotope ratio mass spectrometry (IRMS) offer additional information in the studies. Bomb peak (Section 16.4.7) has proved to be an invaluable tool with a high time resolution in tracer studies of the global carbon cycle.

Cook *et al*. [93] present an excellent review of radiocarbon as a tracer in the global carbon cycle, including atmospheric, ocean, and soil studies. In addition, local effects due to releases from the nuclear fuel cycle are discussed in this review (see also Section 16.4.9). In all environmental studies using natural or bomb-^{14}C as a tracer, potential local effects from anthropogenic ^{14}C must be considered and avoided. Nuclear power plants, reprocessing facilities of nuclear waste, waste incinerators, hospitals, and laboratories using ^{14}C can all be local sources influencing the investigation in question [94, 95].

In the review by Cook *et al*. [93], the different methods for ^{14}C measurement – GPC, liquid scintillation counting (LSC), and AMS – and their relative advantages are discussed. In studies of the global carbon cycle, all three analytical techniques are represented. However, in studies where the amount of available sample material is very limited, AMS is the only possible technique. One such topic of immediate interest is ^{14}C-based source apportionment of organic atmospheric aerosols [96, 97].

Atmospheric aerosols – originating from natural processes or human activities – influence the earth's radiation balance and climate by scattering sunlight and by promoting cloud formation giving a net cooling effect on earth's climate [98]. Thus, atmospheric aerosols are believed mainly to counteract the global warming caused by greenhouse gases. However, certain aerosols, such as soot, can absorb sunlight and contribute to a warming effect. The magnitudes of the different effects are to date shrouded by very large uncertainties. Aerosols also have negative effects on human health, and are estimated to cause 800 000 deaths annually, mainly due to lung diseases [99].

A large fraction of the aerosol mass is carbon-containing (carbonaceous) substances. These roughly originate from three

sources; fossil fuel combustion, biomass burning, and biogenic sources. Information about the source (fossil or modern) can be obtained from the $^{14}C/^{12}C$ ratio in various aerosol fractions. Carbon of fossil origin (e.g., coal) is ^{14}C-free, while ^{14}C in contemporary carbon (e.g., biomass) can be estimated from atmospheric ^{14}C values. Analysis of certain organic tracers related to incomplete combustion of wood is a valuable additional tool in source apportionment [97].

Natural ^{14}C, which is produced only in the atmosphere, is very useful in understanding deep ocean circulation, because any natural ^{14}C atoms found at depth in the ocean must have arrived there by means of exchange between the atmosphere and the ocean surface water and subsequent transport within the body of water. Because ^{14}C decays with a half-life close to 5730 years, its abundance in the deep ocean is a direct measure of how much ^{14}C is supplied to the deep ocean from the surface by ventilation. Temperature measurements of deep ocean water reveal that much of the deep water is cold, and warm water is confined to a thin layer near the surface, indicating that the deep cold water in the subtropics must derive from the polar surface waters. This is known as the *thermohaline circulation* or the *conveyor belt* circulation [100]. A description of the conveyor belt has been made possible by AMS ^{14}C measurements.

As part of the World Ocean Circulation Experiment (WOCE), thousands of small-volume water samples were collected and the ^{14}C content of the dissolved CO_2 analyzed by AMS. During 1990–2004, more than 8000 articles were produced as a result of this program. A summary of the results concerning ^{14}C is given in [101].

The main reason for studying deep circulation is that the deep ocean is the major component of the global carbon cycle. The world's oceans contain approximately 93% of all carbon on earth. As most of the oceanic carbon resides in the deep ocean, even a small change in its carbon budget can significantly affect the atmospheric budget and hence the global climate. The chances of such an event may seem remote, because changes in the deep ocean are slow compared to those in the atmosphere, upper ocean, and terrestrial biosphere. However, measurements from polar ice cores provide abundant evidence of abrupt climate changes during the most recent glacial cycle [102], some of which may have involved changes in the deep ocean [103]. Abrupt climate change may be a real possibility today, as human activities that modify the physical environment are increasing globally [104].

Another important reason for accurately characterizing the deep ocean is the need to validate ocean carbon cycle models. These models are frequently used to predict the response of the ocean to increasing atmospheric CO_2. Projections of future carbon uptake by the ocean [105] inevitably involve the deep ocean. Global mapping of the ^{14}C abundance at 3500 m depth using the WOCE database shows the highest concentrations of ^{14}C in the North Atlantic. The Northeast Pacific has the lowest ^{14}C concentration, indicating the end of the conveyor belt circulation, while the Southern Ocean (the ocean south of 60° S latitude) has an intermediate ^{14}C concentration. Translating the difference in ^{14}C concentration between the North Atlantic and Northeast Pacific deep waters into years reveals an age difference of about 1000 years, which thus characterizes the turnover of the global circulation [106].

16.4.8 Biomedicine

Isotopic labeling has been used for many years for tracing chemicals in living systems. Three kinds of labeling can be used. Short-lived radioactive isotopes have a high signal-to-background ratio, but they have the disadvantage of exposing the organism to radiation. Stable isotopes emit no radiation, but owing to an often high natural background, they are not easy to detect. Long-lived radioactive isotopes have a high signal-to-background ratio and can be used in biomedical AMS in very small amounts. Because of their low radioactivity, the radiation to the individual will be kept to a minimum. The radiation dose deposited in a human as a function of the biological mean life has been calculated by Vogel [107]. The result for a 70 kg person who has been given a compound labeled with ^{14}C ($T_{1/2} = 5730$ year) with an activity of 3700 Bq is shown in Figure 16.4. The dose obtained during a 1 h flight and the dose obtained from natural isotopes inside one's own body (^{14}C, ^{40}K, etc.) during a week are also indicated in the figure for comparison. Another advantage of biomedical AMS is the high sample throughput; the measuring time often being <10 min/sample.

Numerous AMS studies on the human biochemistry have been conducted using ^{14}C (see, e.g., Ref. [108]). AMS is sensitive and precise to a few attomoles of ^{14}C per gram carbon, which renders a number of advantages to biomedical tracing. The use of such small amounts of material enables the use of subtoxic amounts of a chemical substance, the analysis of small-tissue biopsies or a few microliters of blood, as well as the analysis of highly specific biochemical substances and subcellular fractions, including purified DNA. Examples of studies are in drug development and absorption and in distribution, metabolism, and excretion (ADME) studies [108]. Microdosing is another concept in drug development that uses the ultrasensitivity of the AMS technique to open opportunities for early studies in man [109].

AMS has also been used to carry out detailed long-term biokinetic studies of ^{14}C from ^{14}C-labeled pharmaceuticals in humans. The studies conducted were mainly related to so-called breath tests, where the ^{14}C-labeled compound is ingested and metabolized, resulting in the end product, $^{14}CO_2$, which is exhaled and easily collected for measurement [110]. In conclusion, the use of ultralow activities in combination with AMS (down to 1/1000 of those used for LSC) has demonstrated the possibility of metabolic investigations on children, as well as on other sensitive patient groups such as newborns and pregnant or breast-feeding women.

The main AMS isotope in the bioanalytical field is ^{14}C. ^{14}C-AMS is used mainly when its high sensitivity is required, for example, for keeping the effective dose to human test subjects as low as possible. Most laboratory studies using ^{14}C is however performed using LSC. LSC-labeling studies usually demand amounts of ^{14}C that are several orders higher than what is required and suitable for AMS. Extreme care has to be taken not to contaminate biochemical AMS samples from LSC studies or from stock solutions of ^{14}C-labeled materials [111]. These kinds of high-activity samples and solutions should never even enter the AMS sample preparation laboratory.

Apart from ^{14}C, a number of other long-lived radioisotopes have been applied to research on human biochemistry using AMS. Hydrogen is used together with carbon for organic tracing *in vivo* with molecules labeled with ^{3}H

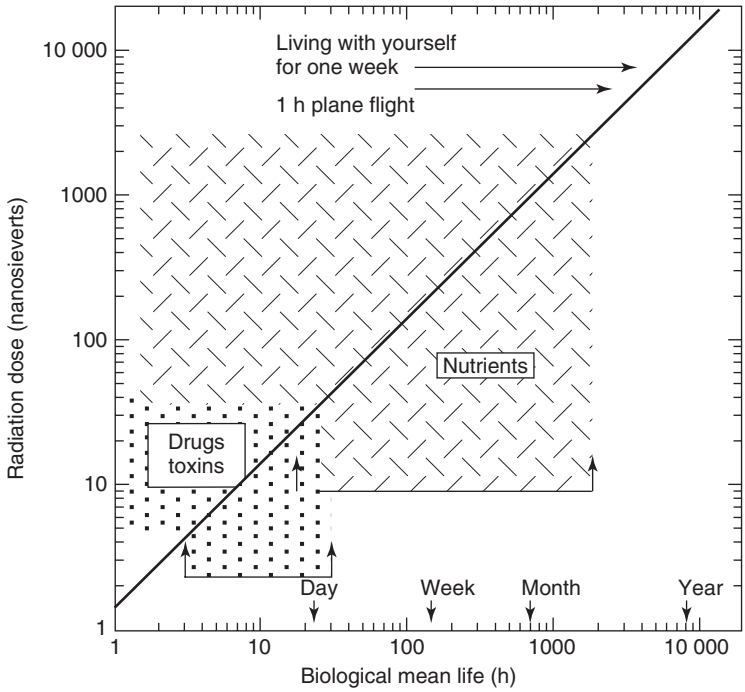


Figure 16.4 The radiation dose deposited in a 70 kg person as a function of the biological mean life of 3700 Bq in a ^{14}C-labeled compound. A 1 h plane flight produces the exposure indicated by "1 h plane flight." Natural radioisotopes within an individual produce the exposure indicated by "living with yourself." (Reprinted from [107] copyright 2000, with permission from Elsevier.)

($T_{1/2} = 12.33$ year) and ^{14}C. The ^{3}H/^{1}H ratio of a milligram-sized water sample can be measured at a level of a few times 10^{-16}, which means that AMS provides a factor of 10^3 improvement in sensitivity with milligram-sized ^{3}H-samples compared to decay counting [112]. Aluminum, the most common metal in the earth's crust, is a nonessential element in biological systems. However, it is known to have a deleterious effect on neurological systems. Aluminum is highly neurotoxic and inhibits prenatal and postnatal development of the brain in humans and animals. The introduction of ^{26}Al ($T_{1/2} = 7.16 \times 10^5$year) as an AMS isotope has enabled the study of aluminum metabolism under physiological conditions. Calcium is an important element in the human body. Many diseases are related to calcium in organs and cells. ^{41}Ca ($T_{1/2} = 1.04 \times 10^5$ year) is therefore an ideal tracer. Several studies of calcium have been reported, such as long-term bone resorption, calcium uptake and deposition in heart tissue, and the metabolism of calcium in the skeleton [113].

The cost and size of conventional AMS accelerators have restricted their penetration into the bioanalytical instruments market. The recently introduced AMS accelerators with voltages well below 1 MV (Section

16.3), having a much reduced size and complexity, have changed the situation. Several commercial companies involved in biomedical research now have their own AMS facility [108, 114–116].

The more extensive use of AMS in biomedical research will require continued development of cost-effective, laboratory-sized AMS systems that can be used in conjunction with gas- and liquid-phase separation techniques. Considerable progress has been made in coupling gas and liquid chromatography directly to AMS to allow online, compound-specific ^{14}C analysis, but no commercial system is yet available [117].

16.4.9 Studies Related to Nuclear Technology

A number of radioactive isotopes are distributed throughout the environment as a result of nuclear weapons testing, nuclear fuel reprocessing, nuclear reactor operation, and, to a small extent, also accidental releases from nuclear facilities. ^{3}H, ^{14}C, ^{99}Tc, ^{129}I, $^{135,134,137}Cs$, ^{237}Np, ^{236}U, $^{239,240,241}Pu$, and other actinides are all examples of isotopes that can be found in the environment surrounding nuclear facilities. These isotopes can be found in air, water, sediments, aerosol particles, plants, animals, and humans. The most commonly applied analytical tool is β-detection and (for the actinides) α-spectrometry. Applications relevant to human health effects often require significantly higher sensitivity than these two standard methods can provide. AMS has demonstrated improved detection limits for all these isotopes. For example, for plutonium isotopes, the reported sensitivity of AMS is $\sim 10^6$ atoms per sample during routine $^{239,240,241,242,244}Pu$ measurements [118].

Fallout of ^{36}Cl from nuclear weapons testing in the 1950s and 1960s has been preserved in glaciers around the world. AMS measurements of this isotope preserved in ice cores have improved estimates of historical, worldwide atmospheric deposition and have allowed the sources of ^{36}Cl in groundwater to be better identified [119]. Another important radionuclide used as a tracer of marine current movements is ^{129}I. The pioneering work on ^{129}I in the ocean was done by Kilius *et al.* [120]. ^{14}C and ^{129}I have been measured in seawater around radioactive waste dump sites. Half a liter of water is enough to identify traces of these isotopes by AMS [121]. The contamination of ^{129}I and ^{127}I have recently been quantified around the coastal waters of Sweden [122]. A review of the ^{129}I topic is given in [123]. The isotope ^{36}Cl has been detected in groundwater samples taken not far from a disposal site for processed nuclear waste in the United States [124].

Today's standards for neutron exposure [in the dosimetry system of 1986 (DS86) and the recently published DS02] are largely dependent on studies of survivors from the Hiroshima and Nagasaki nuclear bombs in August 1945. The radiation effects observed on survivors have to be related to the neutron dose obtained. These doses have until now mostly been obtained by calculations from data regarding the bombs, distances, and so on. AMS investigations of the isotopes ^{36}Cl, ^{41}Ca, ^{63}Ni, and others in samples such as concrete, granite, copper, and so on, irradiated by the nuclear explosion, as a function of the distance from the hypocenter, have shown the need for revision of the adopted dose–response relation for neutrons. The $^{36}Cl/Cl$ ratio in samples of concrete [125] and granite [126] from Hiroshima has been investigated. These three investigations all indicate a much "harder" neutron spectrum

than has previously been estimated. The radioisotope ^{41}Ca is produced by thermal neutron capture by stable ^{40}Ca. In a recent investigation, tooth samples were collected from exposed survivors as well as from large, distant survivors for comparison. The ^{41}Ca/Ca ratios for the exposed survivors show a significant inverse correlation with distance from the hypocenter [127].

^{59}Ni is an important radioisotope in nuclear waste management. The isotope is produced by neutron activation, mainly through the nuclear reaction ^{58}Ni(n, γ)^{59}Ni, in the stainless steel close to the core of a nuclear reactor. The radionuclide ^{59}Ni decays only via electron capture, and the radiation emitted consists primarily of characteristic X-rays. This, in combination with its long half-life of 7.6×10^4 years, makes it difficult to measure by decay counting. PIXEAMS (briefly described in Ref. [6]) provides an efficient way of measuring ^{59}Ni.

A number of steel samples obtained from the Swedish nuclear industry have been analyzed using the Lund AMS system. The samples were taken from different positions close to the core, such as the moderator tank, steam separator, guiding rods for the moderator head, and various flanges. The activity found ranged from a few megabecquerels per gram nickel down to a few kilobecquerels. Samples of recirculating water from a PWR reactor have also been analyzed. Activities of the water samples were found to be 10–30 kBq per liter water [128].

Another example related to the nuclear power industry concerns the release of ^{14}C from power plants, which leads to an increase in the ^{14}C-specific activity of the atmosphere and, hence, to an increased radiation exposure of the population. ^{14}C is one of the radionuclides produced to different degrees by neutron-induced reactions in all types of nuclear reactors. It is believed that, of all nuclides released in routine operation by the nuclear power industry, ^{14}C is likely to produce the largest collective dose to the human population.

Part of the ^{14}C created in reactors is continuously released as airborne effluents in various chemical forms (CO_2, CO, and hydrocarbons) through the ventilation system of the power plant during normal operation. Only a few liters of air are required using ^{14}C-AMS, compared to 100–1000 l for decay counting. In a 1 year study, the total airborne ^{14}C effluents from the stack of two light water reactors were measured continuously over 2 week periods [129, 130].

The incorporation of ^{14}C into living material, mainly leaves and grass, in the environment of power plants has also been studied. The ^{14}C content in annual tree rings of pine (Pinus), located at different distances from power plants, has also been measured [131, 132]. A detailed investigation concerning ^{14}C levels in terrestrial and freshwater samples from the vicinity of the Ignalina nuclear power plant in Lithuania has been presented in [133, 134]. The investigation showed ^{14}C levels in moss and soil samples taken close to the reactor that were up to 20 times higher than the contemporary background level. The excess ^{14}C could be associated with airborne ^{14}C particulates released from the plant.

The nuclear fuel reprocessing facility at Sellafield in Northwest England is known to release substantial amounts of ^{14}C. In Figure 16.5, the ^{14}C content in grass samples collected in September 1996 at various distances from the Sellafield plant is presented [132]. Excess ^{14}C has also been found in the vicinity of research reactors, laboratories for labeling of radioisotopes,

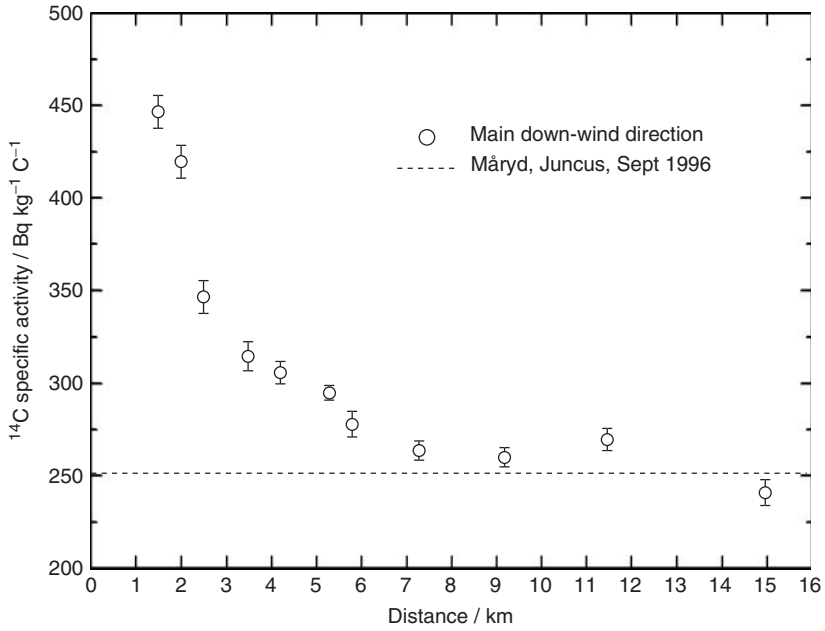


Figure 16.5 The ^{14}C-specific activity in grass collected at various distances in the NNE direction from the Thorp (Thermal Oxide Reprocessing Plant) at the Sellafield nuclear fuel-reprocessing facility. Måryd is a "clean air" site 10 km east of Lund.

and ships used for temporary storage of spent nuclear fuel [135].

16.5 Conclusions and Future Perspectives

AMS is a vital field with an increasing number of applications. Its vitality is demonstrated in the number of new dedicated facilities installed during recent years. Some of these newly installed accelerators are used by private companies for biomedical applications. During the 30 years since the introduction of AMS, it has evolved from the academic world of nuclear physics to the commercial world.

The sophistication of the applications of AMS and the range of applications has been increasing for several years and it appears that this will continue, at least, in the near future. The developments are the result of improvements in the technique itself, and the use of new AMS isotopes.

AMS is playing an increasing role in archaeology as well as the geosciences. The considerable reduction in sample size, compared to the decay-counting technique, has led to new applications. The introduction of gas-ion sources will further reduce the size of samples needed for radiocarbon dating, and the dating of samples as small as 10 μg carbon now seems possible. Surface exposure dating and *in situ* cosmogenic dating are already an important application of AMS and their rapid expansion is foreseen.

As a result of the investigations performed during the past 10 years, new types of instruments have become available.

Compact accelerators running at voltages as low as 200 kV for radiocarbon dating and biomedical applications are already commercially available. The use of lower voltages and simpler technology has already led to an increase in the number of laboratories that perform AMS and this expansion will certainly continue. With some modifications and extensions, the sub-MV accelerators should be able to measure many of the relevant isotopes, such as ^{10}Be, ^{26}Al, ^{41}Ca, ^{129}I, ^{236}U, and the Pu isotopes with acceptable sensitivity. Another important improvement during recent years is the interfacing of standard analytical instruments such as gas chromatographs and high-performance liquid chromatographs with a gas-fed, negative-ion source. This allows AMS to be used for real-time analysis. It is important that the interface provides efficient conversion of a wide variety of biological molecules for the ion source.

As smaller accelerators are installed in new laboratories, the existing larger facilities will be used to analyze isotopes with severe interference from isobars, or for the development of as yet unexplored isotopes. Isotopes that appear to be impossible to measure at low isotope ratios with sub-MV accelerators include ^{32}Si, ^{53}Mn, ^{59}Ni, ^{60}Fe, and ^{99}Tc.

Acknowledgment

Max Strandberg, from the Faculty of Engineering, Lund University, kindly prepared Figures 16.1 and 16.2.

Glossary

Accelerator Mass Spectrometry (AMS): Differs from ordinary mass spectrometry in that it (in most cases) uses a tandem accelerator to accelerate negative ions to high kinetic energies (typical a few megaelectronvolts) before mass analysis. The special strength of AMS among the mass spectrometric methods is its power to separate a rare isotope from an abundant neighboring mass. The method suppresses molecular isobars completely and can, in many cases, also separate atomic isobars (e.g., ^{14}N from ^{14}C). This makes it possible to detect naturally occurring, long-lived radioisotopes such as ^{14}C, ^{10}Be, ^{26}Al, and ^{36}Cl. Their typical isotopic abundance ranges from 10^{12} to 10^{16}. AMS is superior to the competing technique of decay counting for all isotopes where the half-life is long enough (more than a few years).

Calibration of Radiocarbon Years: A raw radiocarbon date cannot be used directly as a calendar date because the level of atmospheric ^{14}C has not been constant during the span of time that can be radiocarbon dated. The level is affected by variations in the cosmic-ray intensity, which is in turn affected by variations in the helio- and geomagnetic fields.

Cosmogenic Nuclides (Or Cosmogenic Isotopes): Rare isotopes created when a high-energy cosmic-ray interacts with the nucleus of an in situ solar system atom, causing cosmic-ray spallation. These isotopes are produced in the atmosphere within rocks or soil and in extraterrestrial items such as meteorites. By measuring cosmogenic isotopes, scientists are able to gain insight into a range of geological and astronomical processes. There are both radioactive and stable cosmogenic isotopes. Some of these radioisotopes are tritium, ^{14}C, and ^{10}Be.

Mass Spectrometry (MS): An analytical technique that measures the mass-to-charge ratio of charged particles. A beam of charged particles enters a magnetic

field and ions with different mass-to-charge ratio will get different trajectories. It is used for determining masses of particles and for determining the elemental composition of a sample or molecule. Typical energies of the ions are in the range of 10–100 keV.

Radiocarbon Dating: A radiometric dating method that uses the naturally occurring cosmogenic nuclide ^{14}C to estimate the age of organic materials up to about 50 000 years or slightly beyond. A radiocarbon date is reported in radiocarbon years ''before present'' (BP), with ''present'' defined as AD 1950. A raw radiocarbon age is then calibrated to give calendar ages. As a first approximation, we can consider that all living organisms that are in direct contact with the atmosphere contain the same concentration of ^{14}C as the atmosphere. ^{14}C decays via *beta*-emission to ^{14}N, with a half-life of circa 5730 years. Owing to the radioactive decay, when a living organism dies – either a plant or an animal – its ^{14}C concentration starts to decrease.

Tandem Accelerator: A type of electrostatic or cascade accelerator, in which negatively charged ions are accelerated through one potential difference before being stripped of two or more electrons inside a high-voltage terminal and (the now positive ions) are accelerated again.

References

1. Hellborg, R., Whitlow, H.J., Hamm, R.W., and Hamm, M.E. (2012) Industrial aspects of ion beam analysis, in *Industrial Accelerators and their Applications*, World Scientific Publ. Co, Singapore, ISBN 978-981-4307-04-8.
2. Purser, K.H., Liebert, R.B., Litherland, A.E., Benkens, R.P., Gove, H.E., Bennett, C.L., Clover, M.R., and Sondheim, W.E. (1977) An attempt to detect stable N^- ions from a sputter ion source and some implications of the results for the design of tandems for ultra-sensitive carbon analysis. *Rev. Phys. Appl.*, **12**, 1487–1492.
3. Nelson, D.E., Korteling, R.G., and Stott, W.R. (1977) Carbon-14: direct detection at natural concentrations. *Science*, **198**, 507–508.
4. Bennett, C.L., Beukens, R.P., Clover, M.R., Gove, H.E., Liebert, R.B., Litherland, A.E., Purser, K.H., and Sondheim, W.E. (1977) Radiocarbon dating using electrostatic accelerators. *Science*, **198**, 508–510.
5. Purser, K.H., Liebert, R.B., and Russo, C.J. (1980) MACS: an accelerator-based radioisotope measuring system. *Radiocarbon*, **22**, 794–806.
6. Hellborg, R. and Skog, G. (2008) Accelerator mass spectrometry. *Mass Spectrom. Rev.*, **27**, 5, 398.
7. Schulze-König, T., Dueker, S.R., Giacomo, J., Suter, M., Vogel, J.S., and Synal, H.A. (2010) BioMICADAS: compact next generation AMS system for pharmaceptical science. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **268**, 891–894.
8. Gove, H.E., Purser, K.H., and Litherland, A.E. (2010) Accelerator Mass Spectrometry (AMS) 1977–1987. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **268**, xvii–xxii.
9. The 12th International Conference on Accelerator Mass Spectrometry. (2013) *Nucl. Instrum. Methods Phys. Res., Sect B*,**294**.
10. Hellborg, R. (ed.) (2005) *Electrostatic Accelerators – Fundamentals and Applications*, Springer-Verlag, Heidelberg.
11. Middleton, R. and Adams, C.T. (1974) A close to universal negative ion source. *Nucl. Instrum. Methods*, **118**, 329–336.
12. Hellborg, Ragnar and Skog, Goran (2010) Sources and detectors in accelerator mass spectrometry, in *The Encyclopedia Mass Spectrometry*, (eds D. Beauchemin and D. E. Matthews), Vol. 5, Elsevier.
13. Bronk Ramsey, C. and Hedges, R.E.M. (1990) A gaseous ion source for routine AMS radiocarbon dating. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **52**, 322–326.
14. Fahrni, S.M., Wacker, L., Synal, H.-A., and Szidat, S. (2013) Improving a gas ion source for ^{14}C AMS. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **294**, 320–327

15. Hajdas, I. (2008) Radiocarbon dating and its applications in Quaternary studies. *Quaternary Sci. J.*, **57** (1-2), 2–24.
16. Genberg, J., Stenström, K., Elfman, M., and Olsson, M. (2010) Development of graphitization of μg-sized samples at Lund University. (eds. E. Bard *et al.*), *Radiocarbon*, **52**, 1270–1276.
17. Vogel, J.S. (1992) Rapid Production of Graphite without contamination for biomedical AMS. *Radiocarbon*, **34** (3), 344–350.
18. Ruff, M., Fahrni, S., Gäggeler, H.W., Hajdas, I., Suter, M., Synal, H.-A., Szidat, S., and Wacker, L. (2010) On-line radiocarbon measurements of small samples using elemental analyzer and Micadas gas ion source. *Radiocarbon*, **52** (4), 1645–1656.
19. Vogel, J.S., Southon, J.R., Nelson, D.E., and Brown, T.A. (1984) Performance of catalytically condensed carbon for use in accelerator mass spectrometry. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **5** (2), 289–293.
20. Wacker, L., Němec, M., and Bourquin, J. (2010) A revolutionary graphitisation system: fully automated, compact and simple. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **268**, 931–934.
21. Ognibene, T.J., Bench, G., and Vogel, J.S. (2003) A high-throughput method for the conversion of CO_2 obtained from biochemical samples to graphite in septa-sealed vials for quantification of ^{14}C via accelerator mass spectrometry. *Anal. Chem.*, **75** (9), 2192–2196.
22. Lee, H.W., Galindo-Uribarri, A., Chang, K.H., Kilius, L.R., and Litherland, A.E. (1984) The $^{12}CH_2^{2+}$ molecule and radiocarbon dating by accelerator mass spectrometry. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **5**, 208–210.
23. Suter, M., Jacob, S., and Synal, H.A. (1997) AMS of ^{14}C at low energies. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **123**, 148–152.
24. Suter, M., Jacob, S.W.A., and Synal, H.A. (2000) Tandem AMS at sub-MeV energies–Status and prospects. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **172**, 144–151.
25. Synal, H.A., Jacob, S., and Suter, M. (2000) The PSI/ETH small radiocarbon dating system. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **172**, 1–7.
26. Goslar, T., Czernik, J., and Goslar, E. (2004) Low-energy ^{14}C AMS in Poznań Radiocarbon Laboratory, Poland. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **223-224**, 5–11.
27. Klein, M., Mous, D., and Gottdang, A. (2006) A compact 1MV multi-element AMS system. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **249**, 764–767.
28. Chamizo, E., Lopez-Gutierrez, J.M., Ruiz-Gomez, A., Santos, F.J., Garcia-Leon, M., Maden, C., and Alfimov, V. (2008) Status of the compact 1 MV AMS facility at the Centro Nacional de Aceleradores (Spain). *Nucl. Instrum. Methods Phys. Res., Sect. B*, **266**, 2217–2220.
29. Suter, M., Müller, A.M., Alfimov, V., Christl, M., Schulze-König, T., Kubik, P.W., Synal, H.A., Vockenhuber, C., and Wacker, L. (2010) Are compact AMS facilities a competitive alternative to larger tandem accelerators? *Radiocarbon*, **52**, 319–330.
30. Synal, H.A., Döbeli, M., Jacob, S., Stocker, M., and Suter, M. (2004) Radiocarbon AMS towards its low energy limits. *Nucl. Instrum. Methods*, **223/224**, 339–345.
31. Schroeder, J.B., Hauser, T.M., Klody, G.M., and Norton, G.A. (2004) Initial results with low energy single stage AMS. *Radiocarbon*, **46** (1), 1–4.
32. Klody, G.M., Schroeder, J.B., Norton, G.A., Loger, R.L., Kitchen, R.L., and Sundqvist, M.L. (2004) New results for single stage low energy carbon AMS. Presented at the 8th European Conference on Accelerators in Applied Research and Technology (ECAART8).
33. Skog, G., Rundgren, M., and Sköld, P. (2010) Status of the single stage AMS machine at Lund University after 4 years of operation. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **268**, 895–897.
34. Freeman, S.P.H.T., Cook, G.T., Dougans, A.B., Naysmith, P., Wilcken, K.M., and Xu, S. (2010) Improved SSAMS performance. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **268**, 715–717.
35. Skog, G. (2007) The single stage AMS machine at Lund University: status report. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **259**, 1–6.

36. Synal, H.A., Stocker, M., and Suter, M. (2007) MICADAS: a new compact radiocarbon AMS system. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **259**, 7–13.
37. Wacker, L., Bonani, G., Friedrich, M., Hajdas, I., Kromer, B., Němec, M., Ruff, M., Suter, M., Synal, H.A., and Vockenhuber, C. (2010) MICADAS: Routine and high-precision radiocarbon dating. *Radiocarbon*, **52** (2–3), 252–262.
38. Synal, H.-A., Schultze-König, T., Seiler, M., Suter, M., and Wacker, L. Mass spectrometric detection of radiocarbon for dating applications. (2012) *Nucl. Instrum. Methods Phys. Res., Sect. B. http://dx.doi.org/10.1016/j.nimb.2012.01.026.*
39. Reimer, P.J., Baillie, M.G.L., Bard, E., Bayliss, A., Beck, J.W., Bertrand, C.J.H., Blackwell, P.G., Buck, C.E., Burr, G.S., Cutler, K.B., Damon, P.E., Edwards, R.L., Fairbanks, R.G., Friedrich, M., Guilderson, T.P., Hogg, A.G., Hughen, K.A., Kromer, B.Mc.C., Gerry, M., Sturt Ramsey, C.B., Reimer, R.W., Remmele, S., Southon, J.R., Stuiver, M., Talamo, S., Taylor, F.W., van der Plicht, J., and Weyhenmeyer, C.E. (2004) IntCal04 terrestrial radiocarbon age calibration, 0–26 Cal Kyr BP. *Radiocarbon*, **46** (3), 1029–1058.
40. Friedrich, M., Remmele, S., Kromer, B., Hofmann, J., Spurk, M., Klaus, F.K., Orcel, C., and Küppers, M. (2004) The 12,460-Year Hohenheim Oak and pine tree-ring chronology from central Europe–a unique Annual record for radiocarbon calibration and paleoenvironment reconstructions. *Radiocarbon*, **46**, 1111–1122.
41. Stuiver, M. (1982) A high-precision calibration of the AD radiocarbon time scale. *Radiocarbon*, **24** (1), 1–26.
42. Schaub, M., Büntgen, U., Kaiser, K.F., Kromer, B., Talamo, S., Andersen, K.K., and Rasmussen, S.O. (2008) Lateglacial environmental variability from Swiss tree rings. *Quaternary Sci. Rev.*, **27**, 29–41.
43. Schaub, M., Kaiser, K.F., Frank, D.C., Büntgen, U., Kromer, B., and Talamo, S. (2008) Environmental change during the Alleröd and Younger Dryas reconstructed from Swiss tree-ring data. *Boreas*, **37**, 74–86.
44. van der Plicht, J., Beck, J.W., Bard, E., Baillie, M.G.L., Blackwell, P.G., Buck, C.E., Friedrich, M., Guilderson, T.P., Hughen, K.A., Kromer, B., McCormac, F.G., Bronk Ramsey, C., Reimer, P.J., Reimer, R.W., Remmele, S., Richards, D.A., Southon, J.R., Stuiver, M., and Weyhenmeyer, C.E. (2004) NotCal04–Comparison/Calibration ^{14}C Records 26–50 Cal Kyr BP. *Radiocarbon*, **46** (3), 1225–1238.
45. Bard, E., Rostek, F., and Ménot-Combes, G. (2004) A better radiocarbon clock. *Science*, **303**, 178–179.
46. Hughen, K., Lehman, S., Southon, J., Overpeck, J., Marchal, O., Herring, C., and Turnbull, J. (2004) ^{14}C activity and global carbon cycle changes over the past 50,000 years. *Science*, **303**, 202–207.
47. Hughen, K., Southon, J., Lehman, S., Bertrand, C., and Turbull, J. (2006) Marine-derived C-14 calibration and activity record for the past 50,000 years updated from the Cariaco Basin. *Quaternary Sci. Rev.*, **25** (23-24), 3216–3227.
48. Muscheler, R., Kromer, B., Björck, S., Svensson, A., Friedrich, M., Kaiser, K.F., and Southon, J. (2008) Tree rings and ice cores reveal ^{14}C calibration uncertainties during the Younger Dryas. *Nat. Geosci.*, **1**, 263–267.
49. Nakagawa, T., Gotanda, K., Haraguchi, T., Danhara, T., Yonenobu, H., Brauer, A., Yokoyama, Y., Tada, R., Takemura, K., Staff, R.A., Payne, R., Bronk Ramsey, C., Bryant, C., Brock, F, Schlolaut, G., Marshall, M., Tarasov, P., and Lamb, H. (2006) Suigetsu Project Members (2012), SG06, a fully continuous and varved sediment core from Lake Suigetsu, Japan: stratigraphy and potential for improving the radiocarbon calibration model and understanding of late Quaternary climate changes. *Quaternary Sci. Rev.*, **36**, 164–176.
50. Turney, C.S.M., FiField, L.K., Hogg, A.G., Palmer, J.G., Hughen, K., Baillie, M.G.L., Galbraith, R., Ogden, J., Lorrey, A., Tims, S.G., and Jones, R.T. (2010) The potential of New Zealand kauri (Agathis australis) for testing the synchronicity of abrupt climate change during the Last Glacial Interval (60,000-11,700 years ago). *Quaternary Sci. Rev.*, **29**, 3677–3682.
51. Goodsite, M.E., Rom, W., Heinemeier, J., Lange, T., Ooi, S., Appleby, P.G., Shotyk, W., van der Knaap, W.O., Lohse, C., and

Hansen, T.S. (2001) High resolution AMS ^{14}C dating of post-bomb peat archives of atmospheric pollutants. *Radiocarbon*, **43** (2B), 495–515.

52. Hua, Q. (2009) Radiocarbon: a chronological tool for the recent past. *Quat. Geochronol.*, **4**, 378–390.

53. Levin, I. and Kromer, B. (2004) The Tropospheric $^{14}CO_2$ Level in Mid-Latitudes of the Northern Hemisphere (1959-2003). *Radiocarbon*, **46** (3), 1261–1272.

54. Levin, I., Hammer, S., Kromer, B., and Meinhardt, F. (2008) Radiocarbon observations in atmospheric CO_2: determining fossil fuel CO_2 over Europe using Jungfraujoch observations as background. *Sci. Total Environ.*, **391**, 211–216.

55. Broecker, W.S., Schulert, A., and Olson, E.A. (1959) Bomb carbon-14 in human beings. *Science*, **130**, 331–332.

56. Libby, W.F., Berger, R., Mead, J.F., Alexander, G.V., and Ross, J.F. (1964) Replacement rates for human tissue from atmospheric radiocarbon. *Science*, **146**, 1170–1172.

57. Harkness, D.D. (1969) Walton A. Carbon-14 in the biosphere and humans. *Nature*, **223**, 1216–1218.

58. Mok, H.Y., Druffel, E.R., and Rampone, W.M. (1986) Chronology of cholelithiasis. Dating gallstones from atmospheric radiocarbon produced by nuclear bomb explosions. *N. Engl. J. Med.*, **314**, 1075–1077.

59. Spalding, K.L., Bhardwaj, R.D., Buchholz, B.A., Druid, H., and Frisen, J. (2005) Retrospective birth dating of cells in humans. *Cell*, **122**, 133–143.

60. Lynnerup, N., Kjeldsen, H., Heegaard, S., Jacobsen, C., and Heinemeier, J. (2008) Radiocarbon dating of the human eye lens crystallines reveal proteins without carbon turnover throughout life. *PLoS ONE*, **3**, e1529.

61. Spalding, K.L., Arner, E., Westermark, P.O., Bernard, S., Buchholz, B.A., Bergmann, O., Blomqvist, L., Hoffstedt, J., Naslund, E., Britton, T., Concha, H., Hassan, M., Ryden, M., Frisen, J., and Arner, P. (2008) Dynamics of fat cell turnover in humans. *Nature*, **453**, 783–787.

62. Bergmann, O., Bhardwaj, R.D., Bernard, S., Zdunek, S., Barnabe-Heider, F., Walsh, S., Zupicich, J., Alkass, K., Buchholz, B.A., Druid, H., Jovinge, S., and Frisen, J. (2009) Evidence for cardiomyocyte renewal in humans. *Science*, **324**, 98–102.

63. Gonçalves, I., Stenström, K., Skog, G., Mattsson, S., Nitulescu, M., and Nilsson, J. (2009) Dating components of human atherosclerotic plaques. *Circ. Res.*, **106**, 1174–1177. doi: 10.1161/CIRCRESAHA.109.211201.

64. Barnekow, L., Possnert, G., and Sandgren, P. (1998) AMS ^{14}C chronologies of Holocene lake sediments in the Abisko area, northern Sweden –a comparison between dated bulk sediment and macrofossil samples. *GFF*, **120**, 59–67.

65. Andrée, M., Oeschger, H., Siegenthaler, U., Riesen, T., Moell, M., Amman, B., and Tobolski, K. (1986) ^{14}C dating of plant macrofossils in lake sediment. *Radiocarbon*, **28** (2A), 411–416.

66. Olsson, I. (1986) Radiometric dating, in *Handbook of Palaeoecology and Palaeohydrology* (ed. B. Berglund), John Wiley & Sons, Ltd, p. 273–312.

67. Bronk Ramsey, C. (2008) Deposition models for chronological records. *Quat. Sci. Rev.*, **27**, 42–60.

68. Bronk Ramsey, C. (2010) Developments in the calibration and modeling of radiocarbon dates. *Radiocarbon*, **52** (2-3), 953–961.

69. Unkel, I., Fernandez, M., Björcka, S., Ljung, K., and Wohlfarth, B. (2010) Records of environmental changes during the Holocene from Isla de los Estados (54.4° S), southeastern Tierra del Fuego. *Global and Planet. Change*, **74** (3–4), 99–113.

70. Lindvall, H., Björck, S., Holmgren, S., Ljung, K., Van der Putten, N., and Porter, C. (2011) A Holocene peat record in the central South Atlantic: an archive of precipitation changes. *GFF*, **133** (3-4), 195–206.

71. Bard, E., Raisbeck, G.M., Yiou, F., and Jouzel, J. (1997) Solar modulation of cosmogenic nuclide production over the last millennium: comparison between ^{14}C and ^{10}Be records. *Earth Planet. Sci. Lett.*, **150** (3-4), 453–446.

72. Beer, J., Baumgartner, S., Hannen-Dittrich, B., Hauenstein, J., Kubik, P., Lukasczyk, C., Mende, W., Stellmacher, R., and Suter, M. (1994) Solar variability traced by

cosmogenic isotopes, in *The Sun as a Variable Star: Solar and Stellar Irradiance Variations*. (eds. J.M. Pap, C. Fröhlich, H.S. Hudson, W.K. Tobiska), Cambridge University Press, Cambridge, pp. 291–300.

73. Sharma, M. (2002) Variations in solar magnetic activity during the last 200,000 years: is there a sun-climate connection? *Earth Planet. Sci. Lett.*, **199**, 459–472.

74. Bond, G., Kromer, B., Beer, J., Muscheler, R., Evans, M.N., Showers, W., Hoffmann, S., Lotti-Bond, R., Hajdas, I., and Bonani, G. (2001) Persistent solar influence on North Atlantic climate during the holocene. *Science*, **294**, 2130–2136.

75. Beer, J., Mende, W., and Stellmacher, R. (2000) The role of the sun in climate forcing. *Quat. Sci. Rev.*, **19** (1), 403–415.

76. Martin-Puertas, C., Matthes, K., Brauer, A., Muscheler, R., Hansen, F., Petrick, C., Aldahan, A., Possnert, G., and van Geel, B. (2012) Regional atmospheric circulation shifts induced by a grand solar minimum. *Nat. Geosci.*, doi: 10.1038/ngeo1460.

77. Muscheler, R. and Heikkil, U. (2011) Constraints on long-term changes in solar activity from the range of variability of cosmogenic radionuclide records. *Astrophys. Space Sci. Trans.*, **7**, 355–364.

78. Raisbeck, G.M., Yiou, F., Cattani, O., and Jouzel, J. (2006) ^{10}Be evidence for the Matuyama-Brunhes geomagnetic reversal in the EPICA dome C ice core. *Nature*, **444** (7115), 82–84.

79. Gosse, J.C. and Phillips, F.M. (2001) Terrestrial in situ cosmogenic nuclides: theory and application. *Quat. Sci. Rev.*, **20**, 1475–1560.

80. Ivy-Ochs, S. and Kober, F. (2008) Surface exposure dating with cosmogenic nuclides. *Quat. Sci. J.*, **57** (1–2), 179–209.

81. Hedges, R.E.M. and van Klinken, G.J. (1992) A review of current approaches in the pretreatment of bone for radiocarbon dating by AMS. *Radiocarbon*, **34** (3), 279–291.

82. Lanting, J.N., Aerts-Bijma, A.T., and van der Plicht, J. (2001) Dating of cremated bones. *Radiocarbon*, **43** (2A), 249–254.

83. Valladas, H., Clottes, J., Geneste, J.M., Garcia, M.A., Arnold, M., Cachier, H., and Tisnerát-Laborde, N. (2001) Evolution of prehistoric cave art. *Nature*, **413**, 479.

84. Valladas, H. (2003) Direct radiocarbon dating of prehistoric cave paintings by accelerator mass spectrometry. *Measur. Sci. Technol.*, **14**, 1487–1492.

85. Mellars, P. (2006) A new radiocarbon revolution and the dispersal of modern humans in Eurasia. *Nature*, **439**, 931–935.

86. Bronk Ramsey, C., Higham, T., Bowles, A., and Hedges, R. (2004) Improvements to the pre-treatment of bone at Oxford. *Radiocarbon*, **46**(1), 155–163.

87. Jacobi, R.M., Higham, T.F.G., and Bronk Ramsey, C. (2006) AMS radiocarbon dating of Middle and Upper Palaeolithic bone in the British Isles: improved reliability using ultrafiltration. *J. Quat. Sci.*, **21** (5), 557–573.

88. Balter, M. (2006) Mild climate, lack of moderns let last neandertals linger in gibraltar. *Science*, **313**, 1557.

89. Damon, P.E., Donahue, D.J., Gore, B.H., Hatheway, A.L., Jull, A.J.T., Linick, T.W., Sercel, P.J., Toolin, L.J., Bronk, C.R., Hall, E.T., Hedges, R.E.M., Housley, R., Law, I.A., Perry, C., Bonani, G., Trumbore, S., Woelfli, W., Ambers, J.C., Bowman, S.G.E., Leese, M.N., and Tite, M.S. (1989) Radiocarbon dating of the shroud of turin. *Nature*, **337** (6208), 611–615.

90. Bonani, G., Ivy, S., Hajdas, I., Niklaus, T.R., and Suter, M. (1994) AMS ^{14}C age determinations of tissue, bone and grass samples from the Ötztal Ice Man. *Radiocarbon*, **36** (2), 247–250.

91. Bonani, G., Ivy, S., Wölfli, W., Broshi, M., Carmi, I., and Strugnell, J. (1992) Radiocarbon dating of fourteen Dead Sea scrolls. *Radiocarbon*, **34** (3), 843–849.

92. Skog, G. (2002) Accelerator mass spectrometry in quaternary geology and archaeology, in *Proc Int Conf Applications of High Precision Atomic and Nuclear Methods* (eds A. Olariu, K. Stenström, and R. Hellborg), Editura Academiei Romane, pp. 232–242.

93. Cook, G.T., Scott, E.M., and Harkness, D.D. (2010) Radiocarbon as a tracer in the global carbon cycle. *Radioact. Environ.*, **16**, 89–137.

94. Stenström, K., Skog, G., Nilsson, C.M., Hellborg, R., Leide Svegborn, S., Georgiadou, E., and Mattsson, S. (2010) Local variations in ^{14}C –how is bomb-pulse dating of human cells and tissues affected?

Nucl. Instrum. Methods Phys. Res., Sect. B, **268**, 1299–1302. doi: 101016/jnimb200910157.

95. Stenström, K., Unkel, I., Nilsson, C.M., Rääf, C., and Mattsson, S. (2010) The use of hair as an indicator of occupational ^{14}C contamination. *J. Radiat. Environ. Biophys.*, **49**, 97–107. doi: 101007/s00411-009-0245-9.

96. Genberg, J., Hyder, M., Stenström, K., Bergström, R., Fors, E., Jönsson, J.A.A., and Swietlicki, E. (2011) Source apportionment of carbonaceous aerosol in Southern Sweden. *Atmos. Chem. Phys.*, **11**, 11387–11400. doi: 10.5194/acp-11-11387-2011.

97. Yttri, K.E., D Simpson, D., Stenström, K., Puxbaum, H., and Svendby, T. (2011) Source apportionment of the carbonaceous aerosol in Norway – Quantitative estimates based on ^{14}C, thermal optical and organic tracer analysis. *Atmos. Chem. Phys.*, **11**, 9375–9394. doi: 10.5194/acp-11-9375-2011.

98. Solomon, S., Qin, D., Manning, M., Alley, R.B., Berntsen, T., Bindoff, N.L., Chen, Z., Chidthaisong, A., Gregory, J.M., Hegerl, G.C., Heimann, M., Hewitson, B., Hoskins, B.J., Joos, F., Jouzel, J., Kattsov, V., Lohmann, U., Matsuno, T., Molina, M., Nicholls, N., Overpeck, J., Raga, G., Ramaswamy, V., Ren, J., Rusticucci, M., Somerville, R., Stocker, T.F., Whetton, P., Wood, R.A., and Wratt, D. (2007) Technical Summary Rep, IPCC.

99. WHO (2002) *The World Health Report 2002. Reducing Risks, Promoting Healthy Life*, World Health Organization, Geneva.

100. Broecker, W.S. (1991) The great ocean conveyor. *Oceanography*, **4** (2), 79–89.

101. Key, R.M., Quay, P.D., Schlosser, P., McNichol, A., Reden, K.F., Schneider, B.J., Elder, K.L., Stuiver, M., and Ostlund, H.G. (2002) WOCE Radiocarbon IV: Pacific Ocean results; P10, P13N, P14C, P18, P19, & S4P. *Radiocarbon*, **44** (1), 239–392.

102. Dansgaard, W., Johnsen, S.J., Clausen, H.B., Dahl-Jensen, D., Gundestrup, N.S., Hammer, C.U., Hvidberg, C.S., Steffensen, J.P., Sveinbjornsdottir, A.E., Jouzel, J., and Bond, G. (1993) Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, **364** (6434), 218–220.

103. Broecker, W.S. (2003) Does the trigger for abrupt climate change reside in the ocean or in the atmosphere. *Science*, **300** (5625), 1519–1522.

104. Alley, R.B., Marotzke, J., Nordhaus, W.D., Overpeck, J.T., Peteet, D.M., Pielke, R.A., Pierrehumbert, R.T., Rhines, P.B., Stocker, T.F., Talley, L.D., and Wallace, J.M. (2003) Abrupt climate change. *Science*, **299** (5615), 2005–2010.

105. Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., and Xiaosu, D. (2001) *Climate Change 2001–The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, p. 944.

106. Matsumoto, K. and Key, R.M. (2004) Natural radiocarbon distribution in the deep ocean, in *Global Environmental Change in the Ocean and on Land* (eds M. Shiyomi *et al.*), TERRAPUB, pp. 45–58

107. Vogel, J.S. (2000) Accelerator mass spectrometry for human biochemistry: the practice and the potential. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **172**, 884–891.

108. Arjomand, A. (2010) Accelerator mass spectrometry-enabled studies: current status and prospects. *Bioanalysis*, **2** (3), 519–541.

109. Lappin, G. (2011) Microdosing: pharmacokinetic and metabolism data early in the drug development process, in Pharmacokinetics in Drug Development. (eds. P.L. Bonate and D.R. Howard), American Association of Pharmaceutical Scientists , pp. 115–130. DOI 10.1007/978-1-4419-7937-7_5.

110. Gunnarsson, M., Leide Svegborn, S., Stenström, K., Skog, G., Nilsson, L.-E., Hellborg, R., and Mattson, S. (2007) Long-term biokinetics and radiation exposure of patients undergoing ^{14}C-glycocholic acid and ^{14}C-xylose breath tests for diagnosis of small intestinal bacterial overgrowth. *Cancer Biother. Radiopharm.*, **22**, 762–771.

111. Buchholz, A., Freeman, S., Haack, K., and Vogel, J.S. (2000) Tips and traps in the ^{14}Cbio-AMS preparation laboratory. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **172**, 404–408.

112. Ognibene, T.J., Bench, G., Brown, T.A., and Vogel, J.S. (2004) The LLNL accelerator mass spectrometry system for biochemical ^{14}C-measurements. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **223/224**, 12–15.

113. Jiang, S., He, M., Dong, K., Yue, D., Wu, S., Xu, G., Qin, J., You, Q., Zheng, Y., Guan, Y., Liang, Q., Zhang, G., Zhao, X., Wang, Q., and Liu, S. (2004) The measurement of ^{41}Ca and its application for the cellular Ca^{2+} concentration fluctuation caused by carcinogenic substances. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **223/224**, 750–753.

114. Seymour, M. (2011) Accelerator MS: its role as frontline bioanalytical technique. *Bioanalysis*, **3** (24), 2817–2823.

115. Young, G.C. and Ellis, W.J. (2007) AMS in drug development at GSK. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **259**, 752–757.

116. Zoppi, U., Crye, J., Song, Q., and Arjomand, A. (2007) Performance evaluation of the new AMS system at Accium Biosciences. *Radiocarbon*, **59** (1), 171–180.

117. McIntyre, C.P., Galutschek, E., Roberts, M.L., von Reden, K.F., McNichol, A.P., and Jenkins, W.J. (2010) A continuous-flow gas chromatography ^{14}c accelerator mass spectrometry system. *Radiocarbon*, **52** (2), 295–300.

118. Brown, T.A., Marchetti, A.A., Martinelli, R.E., Cox, C.C., Knezovich, J.P., and Hamilton, T.F. (2004) Actinide measurements by accelerator mass spectrometry at Lawrence Livermore National Laboratory. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **223/224**, 788–795.

119. Green, J.R., Cecil, L.D., Synal, H.A., Santos, J., Kreutz, K.J., and Wake, C.P. (2004) A high resolution record of chlorine-36 nuclear-weapons-tests fallout from Central Asia. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **223/224**, 854–857.

120. Kilius, L.R., Rucklidge, J.C., and Litherland, A.E. (1987) Accelerator mass spectrometry of ^{129}I at Isotrace. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **223/224**, 72–76.

121. Povinec, P.P., Oregioni, B., Jull, A.J.T., Kieser, W.E., and Zhao, X.L. (2000) AMS measurements of ^{14}C and ^{129}I in seawater around radioactive waste dump sites. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **172**, 672–678.

122. Hansen, V., Yi, P., Hou, X.L., Aldahan, A., Roos, P., and Possnert, G. (2011) Iodide and iodate (^{129}I and ^{127}I) in surface water of the Baltic Sea, Kattegat and Skagerrak. *Sci. Total Environ.*, **412–413**, 296–303.

123. Hou, X.L., Hansen, V., Aldahan, A., Possnert, G., Lind, O.C., and Lujaniene, G. (2009) A review on speciation of iodine-129 in the environmental and biological samples. *Anal. Chim. Acta*, **632**, 181–196.

124. Cecil, L.D., Welhan, J.A., Green, J.R., Grape, S.H., and Sudicky, E.R. (2000) Use of chlorine-36 to determine regional-scale aquifer dispersivity, eastern Snake River Plain aquifer, Idaho/USA. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **172**, 679–687.

125. Straume, T., Finkel, R.C., Eddy, D., Kubik, P.W., Gove, H.E., Sharma, P., Fujita, S., and Hoshi, M. (1990) Use of accelerator mass spectrometry in the dosimetry of Hiroshima neutrons. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **52**, 552–556.

126. Nagashima, Y., Seki, R., Matsuhiro, T., Takahashi, T., Sasa, K., Sueki, K., Hoshi, M., Fujita, S., Shizuma, K., and Hasai, H. (2004) Chlorine-36 in granite samples from the Hiroshima A-bomb site. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **223/224**, 782–787.

127. Wallner, A., Rühm, W., Rugel, G., Nakamura, N., Arazi, A., Faestermann, T., Knie, K., Maier, H.J., and Korschinek, G. (2010) ^{41}Ca in tooth enamel. Part I: A biological signature of neutron exposure in atomic bomb survivors. *Radiat. Res.*, **174**, 137–145.

128. Persson, P. (2002) Improved detection limit for ^{59}Ni using the technique of accelerator mass spectrometry, PhD Thesis, Lund University, Lund, Sweden

129. Stenström, K., Erlandsson, B., Hellborg, R., Wiebert, A., Skog, G., Vesanen, R., Alpsten, M., and Bjurman, B. (1995) A one-year study of the total air-borne ^{14}C effluents from two Swedish light-water reactors, one boiling water- and one pressurized water reactor. *J. Radioanal. Nucl. Chem.*, **198**, 203–213.

130. Stenström, K., Erlandsson, B., Hellborg, R., Wiebert, A., and Skog, G. (1996) Determination of the $^{14}CO_2$ and the total airborne ^{14}C releases from two Swedish

light- water reactors using accelerator mass spectrometry. *Radioact. Radiochem.*, **7**, 32–36.

131. Stenström, K., Erlandsson, B., Hellborg, R., Wiebert, A., and Skog, G. (1996) Environmental levels of carbon-14 around a Swedish nuclear power plant measured with accelerator mass spectrometry. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **113**, 474–476.

132. Stenström, K., Erlandsson, B., Faarinen, M., Hellborg, R., Kiisk, M., Persson, P., Skog, G., Thornberg, C., Mattsson, S., Olariu, A., Olariu, S., Baciu, F., and Cutoiu, D. (2002) Radioecological applications of ^{14}C measurements at the Lund accelerator mass spectrometry (AMS) facility, in *Proceedings of Applications of high precision atomic & nuclear methods* (eds A. Olariu, K. Stenström, and R. Hellborg), Academiei Romane, Bucarest, pp. 254–265.

133. Magnusson, A.A., Stenström, K., Skog, G., Adliene, D., Adlys, G., Hellborg, R., Olariu, A., Zakaria, M., Rääf, C., and Mattsson, S. (2004) Levels of ^{14}C in the terrestrial environment in the vicinity of two European nuclear power plants. *Radiocarbon*, **46**, 863–868.

134. Magnusson, A.A., Stenström, K., Adliene, D., Adlys, G., Dias, C., Rääf, C., Skog, G., Zakaria, M., and Mattsson, S. (2007) Carbon-14 levels in the vicinity of the Lithuanian nuclear power plant Ignalina. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **259**, 530–535.

135. Stenström, K., Skog, G., Nilsson, C.M., Hellborg, R., Leide Svegborn, S., Georgiadou, E., and Mattsson, S. (2010) Local variations in ^{14}C –how is bomb-pulse dating of human cells and tissues affected? *Nucl. Instrum. Methods Phys. Res., Sect. B*, **268**, 1299–1302.

Further Readings

Hamm, R.W. and Hamm, M.E. (eds) (2012) *Industrial Accelerators and their Applications*, World Scientific Publ. Co, Singapore. ISBN 978-981-4307-04-8.

Hellborg, R. (ed.) (2005) *Electrostatic Accelerators – Fundamentals and Applications*, Springer Verlag, Heidelberg.

Hellborg, R. and Skog, G. (2008) Accelerator mass spectrometry, *Mass Spectrom. Rev.*, **27:5**, 398.

The 12th International Conference on Accelerator Mass Spectrometry. (2013) *Nucl. Instr. Meth. B*, **294**.

Proceedings of the 20th International Radiocarbon Conference (31 May–5 June, 2009), Kona Hawaii, 2010. Radiocarbon Vol 52, no 2–3.

17
Nuclear Medicine

Ulli Köster, Ferid Haddad, Nicolas Chouin, François Davodeau, Jean-François Chatal, Jacques Barbet, and Françoise Kraeber-Bodéré

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

17.1 Introduction

Nuclear medicine is a medical specialty that consists in administering to patients some drugs coupled to radionuclides (radiopharmaceutical agents) for diagnostic imaging (mainly in cardiology, neurology, and oncology) or for therapy (mainly in oncology).

Diagnostic and therapeutic administration of radionuclides to patients through oral or intravenous route has been introduced in the 1940s, with the use of ^{131}I (iodine-131) for diagnosis and treatment of thyroid pathology. This was made easy because iodine is taken up by thyroid cells and can be directly used without the need to be coupled to a carrier molecule. During the following decades, some carrier molecules have been evaluated after labeling with ^{123}I (iodine-123) for imaging in patients with neurological, cardiological, and oncological diseases, and ^{131}I for therapy in patients with neuroendocrine tumors (NETs), liver cancer, and malignant lymphoma and, in the 1970s with ^{99m}Tc (technetium-99m) for imaging and with ^{90}Y (yttrium-90) for therapy. Some of them have been approved and routinely used.

At the end of the 1990s, a real technological revolution deeply changed the sphere of activity of nuclear medicine in imaging with the introduction and the routine use of ^{18}F (fluorine-18) and the radiopharmaceutical fluorodeoxyglucose (FDG) for imaging in oncology using positron emission tomography (PET) as an innovative imaging modality. With the availability of PET/CT (computed tomography) hybrid equipments (PET system combined to CT) and more recently PET/MRI (PET system combined to nuclear magnetic resonance imaging), there is no doubt that the number of radiopharmaceuticals will grow in the coming years for applications in oncology, neurology, and cardiology. Owing to the near ideal physical characteristics of ^{18}F and its availability, there is an obvious trend to use it for labeling of a lot of molecules of interest, but some of them are not easily and stably labeled with ^{18}F. That is why innovative short-lived positron-emitting radionuclides such as ^{68}Ga (gallium-68) or ^{64}Cu (copper-64) are needed. Moreover, the half-life of ^{18}F is too short for labeling of large molecules such as antibodies, which distribute slowly, in a few days, in the body. Long-lived radionuclides such as ^{89}Zr (zirconium-89) or ^{124}I are then required.

For therapy, a great promise can be expected from radioimmunotherapy (RIT) and peptide receptor radionuclide therapy (PRRT). For these applications, various radionuclides can be considered. ^{131}I is no more appropriate because it emits

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

also hard γ-rays; thus, at least in Europe, it requires to keep patients in shielded rooms, the number of which is quite limited. New β-emitting radionuclides are therefore requested for an easy and safe use. Finally, for the past decade, a growing interest has been demonstrated for the use of α-emitting radionuclides, which could be quite efficient for the treatment of microscopic tumors disseminated in the body.

17.2 Current Medical Use of Radionuclides and Radiopharmaceuticals

17.2.1 For Imaging

The main interest of radionuclides for imaging is due to their very high sensitivity, which makes them detectable in pico- to femtomolar concentration range. They allow us to visualize high-affinity receptor systems that have low concentrations. These low concentrations are not associated with some pharmacological effects and allow radionuclides to trace these receptor systems. Moreover, with regard to morphological imaging techniques such as CT, ultrasonography, and magnetic resonance imaging, which show small morphological abnormalities with high sensitivity without specifying their nature, nuclear medicine imaging techniques (single photon emission computed tomography, SPECT and PET) add functional information – for example, indicating the benign or malignant nature of a morphological abnormality, and thus are quite complementary with morphological imaging techniques.

17.2.1.1 Common SPECT Radionuclides

Imaging with radionuclides requires that emitted photons have energy high enough to cross the body before interacting with the detector. This energy should not be too low (under 60 keV) so as to avoid being attenuated too much by traversed tissues. Too high energies ($\gg$300 keV) are unfavorable because they require thicker collimators with lower transmission. Moreover, half-lives should not be too long to avoid excessive dose for the injected patients. ^{99m}Tc is by far the most common radionuclide used for scintigraphic imaging. It is daily available, for one week, after elution of a ^{99}Mo/^{99m}Tc generator. The energy of emitted photons (140 keV) is best matched to single photon imaging detectors, and the half-life (6 h) is short enough to minimize radiation burden and long enough to allow sufficient time for radiopharmaceutical preparation, injection, and distribution in the body, and for sought lesions. Moreover, its cost is acceptable. ^{123}I also has favorable characteristics and tends to be more and more used for labeling of molecules for SPECT applications. ^{111}In emits photons with favorable energy, and its half-life is well suited for labeling of molecules with slow pharmacokinetics such as antibodies. ^{201}Tl has been the first radionuclide used for myocardial perfusion imaging in the 1970s. It has favorable characteristics for scintigraphic imaging, but it causes a high effective dose equivalent, which limits the injected dose and consequently the image quality.

17.2.1.2 Common PET Radionuclides

^{18}F, a positron emitter with a physical half-life of 110 min, has been introduced in routine clinical practice of nuclear oncology in the second part of the 1990s. Its use, when coupled to deoxyglucose (^{18}F-FDG), has revolutionized the place of nuclear medicine with regard to other imaging techniques. FDG is a non-metabolized glucose analog that is a substrate for

glucose transporters on the cell membrane and cytoplasmic hexokinase enzyme [1]. After phosphorylation, FDG is trapped in the cell with a residence time compatible with PET recording. Many cancers have a large expression of glucose transporters and hexokinase, which is at the origin of a high FDG uptake. Clinical indications for FDG-PET cover the majority of solid cancers. This imaging technique provides crucial information to oncologists, and is considered as the most informative among other investigations. In a recent study conducted in 30 000 patients in the USA, FDG-PET had a major impact on patient care in 30% of cases. One very interesting and promising indication of FDG-PET is the early evaluation of response to treatment with regard to conventional response evaluation criteria in solid tumors (RECIST) based on the change in tumor volume, which can be delayed by comparison to FDG tumor uptake. Beyond FDG, a large number of molecules labeled with ^{18}F have been evaluated in clinical studies, but only a few of them has been considered as useful in clinical practice. It is highly likely that within the next decade 5–10 ^{18}F-labeled tracers will be commonly used for PET imaging, in addition to FDG.

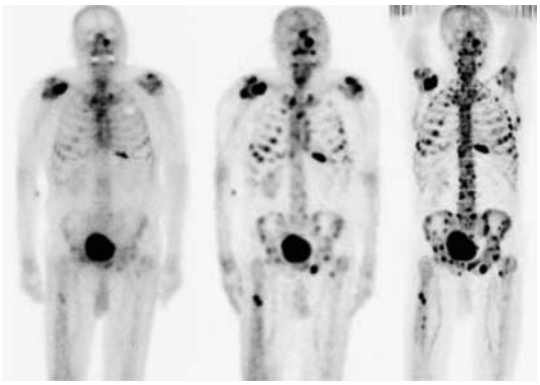
Owing to its short half-life, ^{18}F must be produced in a small cyclotron located not far from user hospitals (about 300 km). Usually three production runs are performed in early morning, and delivered to the user sites in the morning and early afternoon.

17.2.1.3 Radiopharmaceuticals

Some radioisotopes (e.g., radioiodine or ^{82}Rb) can be used directly as soluble ionic salts, thanks to their chemical and biological properties. In most cases, it is necessary to associate the radionuclide to a carrier molecule that will target the cells of interest. This radiolabeled molecule is called a *radiopharmaceutical*. An ideal radiopharmaceutical should localize rapidly to allow lesion visualization as soon as possible after injection. The count rate should be high to collect the data in the shortest interval time, and the target-to-background ratio should also be high to allow a clear visualization of the sought lesion. Finally, the radiation burden delivered to the patient should be as low as possible to allow repeated and safe injections.

A given radionuclide can be used for various pathologies provided it is coupled to a suitable vector. For example, there are numerous ^{99m}Tc radiopharmaceuticals that all use the stannous reduction method. The main clinically used ^{99m}Tc radiopharmaceuticals are:

- ^{99m}Tc-MDP (methylene diphosphonate). This radiopharmaceutical is taken up by *osteoblast* cells that build bone. It is used in nuclear oncology for bone scintigraphy (Figure 17.1), which is mainly used in search of metastases from prostate and breast cancer.
- ^{99m}Tc-MIBI (methoxyisobutylisonitrile) and ^{99m}Tc-tetrofosmin (Myoview®). ^{99m}Tc-MIBI (Cardiolite®). When injected intravenously into a patient during a stress test and at rest, it distributes in the myocardium proportionally to the myocardial blood flow. ^{99m}Tc-MIBI test is performed to diagnose coronary artery disease.
- ^{99m}Tc-DTPA (diethylene triamino penta acetic acid). This radiopharmaceutical is used to determine overall kidney function and also relative left and right kidney function. The scan can also be very useful in assessing the function of renal transplant grafts.
- ^{99m}Tc-MAA (macro aggregated albumin). Radioactive aggregated



(a) (b) (c)

Figure 17.1 Bone metastases detected with ^{99m}Tc-MDP on a planar scan (a) and a SPECT image (b) compared to a ^{18}F-fluoride PET image (c). Reprinted by permission of SNMMI from Ref. [2], figure 1.

particles are trapped in the pulmonary alveolar capillary bed following intravenous injection, and their distribution is uniform throughout the vascular bed, thus producing a uniform scintigraphic image. This radiopharmaceutical is mainly used for detection of pulmonary embolism.

17.2.2 Radiopharmaceuticals for Therapy

Radiopharmaceuticals can also be used for treatment of various cancer types through different routes of administration (intravenous, oral, or intra-cavitary). This internal radiotherapy is most often used as a unique treatment or can be associated with other treatment modalities such as chemotherapy.

^{131}I is routinely used for treatment of differentiated thyroid. After thyroidectomy, the objectives of radioiodine ablation are to eliminate remnants aiming at decreasing the risk for loco-regional disease. At the time of metastatic progression, the success of ^{131}I therapy depends on the size of metastases. In the 1980s, ^{131}I-labeled metaiodobenzylguanidine (^{131}I-MIBG) was the first radioiodinated molecule introduced into clinical practice to treat relatively rare phaeochromocytomas, paragangliomas, neuroblastomas, and carcinoids.

For the past 20 years, ^{90}Y has been used for several therapeutic applications. The first one, at the beginning of the 1990s, consisted in the intra-arterial delivery of micron-sized ^{90}Y-tagged microspheres that preferentially and permanently embed in tumor as opposed to normal tissue. This therapeutic modality has been used to treat primary or metastatic liver cancer via direct infusion into the hepatic arterial circulation.

17.2.3 Nuclear Medicine Tools for Tomorrow's Personalized Medicine

Usefulness of aforementioned radionuclides has been well documented. However, innovative radionuclides appear to be necessary for the development of new molecular imaging and therapy, which should be more specific and more efficient, within the scope of a personalized medicine approach, which has become a reality owing to an improved molecular understanding of pathological processes. Personalized medicine is today applied to clinical neurology and oncology [3],

and consists in giving the most adapted treatment for one individual patient (most appropriate doses and sequences for a given therapeutic agent) on the basis of reliable biological data with the aim to spare those non responding patients from secondary adverse effects.

Molecular diagnosis will play an important role in personalized medicine for which diagnosis and therapy will be integrated within the scope of a so-called theranostic approach. Nuclear medicine imaging allows us to characterize pathological processes such as tumor phenotype by injecting radiolabeled vectors that specifically target biological processes of interest [4]. The same vectors can be radiolabeled with β-emitting- or α-particle-emitting radionuclides as very specific molecular radiotherapy agents, with the hope to efficiently kill disseminated tumor cells in the body that are refractory to conventional treatments.

For both imaging and therapy applications, the use of new biological vectors requires innovative radionuclides with complementary physical characteristics. This choice of radionuclide depends on these characteristics, but also on the type of tumor, the molecular target, and the nature of the vector. Table 17.5 and Table 17.7 show the list of innovative radionuclides used in PET imaging and therapy.

Physical half-life of the radionuclide should match the biological half-life of the vector. As an example, the radionuclide suitable for labeling of a large molecule of intact immunoglobulin, which distributes in the body over a few days, should be different from that appropriate for labeling of a peptide that reaches its target in a few hours. The use of ^{68}Ga with a half-life of 68 min is limited to the labeling of small size molecules that distribute rapidly in the body, whereas ^{89}Zr and ^{124}I, with half-lives of respectively 3.3 and 4.2 days, are well suited to the labeling of large molecules such as antibodies.

The particle path length is another important parameter depending on nature and energy of the particle. In PET imaging, spatial resolution and image quality depend on the positron path length. ^{64}Cu allows us to obtain images with excellent resolution close to that of ^{18}F [5]. With ^{68}Ga and ^{124}I, image resolution is theoretically less good but clinical practice shows in fact very high image quality as well [6, 7].

In therapy, the range of α-particles is very short (<0.1 mm), which limits the use of α-emitters to isolated tumor cells or micro-clusters of malignant cells. The range of electrons for β-emitters is longer depending on energy of the considered radionuclide. Consequently, they can efficiently irradiate macro clusters of tumor cells with a few millimeters in diameter. As antigen or receptor expression is often quite variable among all tumor cells, β-emitters that are bound to antigen or receptor positive tumor cells can irradiate neighboring tumor cells that do not express antigen or receptor according to a so-called cross fire effect.

With regard to β-particles, α-particles have a much higher linear energy transfer (LET) and a much shorter range (<0.1 mm). They have a much higher killing effect for isolated tumor cells than β-particles [8]. ^{211}At is considered as one of the most promising α-particle-emitting radionuclide for medical applications owing to its physical features.

For diagnostic and therapeutic applications, emission spectrum of the radionuclide and especially associated γ-emissions have to be considered with regard to radiation safety and dosimetry. Moreover, branching ratios of β^+ emissions should be taken into consideration, and

for radionuclides with branching ratios smaller than that of ^{18}F, higher activity should be injected and image acquisition time should be longer. For therapy, energy spectra of ^{177}Lu and ^{67}Cu are more favorable than that of ^{131}I with γ-emissions much less energetic and abundant.

Within the scope of a Theranostic approach, matched pairs of β^+/β^- emitting radionuclides ($^{124}I/^{131}I$, $^{86}Y/^{90}Y$, $^{64}Cu/^{67}Cu$, $^{44}Sc/^{47}Sc$, $^{152}Tb/^{161}Tb$) are very promising because the same biodistribution is expected both for imaging dosimetry and for therapy using the same elements. A reliable dosimetry can thus be possible before a molecular radiotherapy and could help adapt injected activity and predict toxicity and efficacy.

Moreover, labeling stability, availability, and cost of radionuclides have to be taken into consideration. As an example, for PET imaging, ^{68}Ga has the great advantage of being easily and daily available from a generator for a few months.

17.3 Modes of Radioisotope Production

The medical use of a radionuclide is connected to its decay mode. Radionuclides emitting penetrating radiation will be used for imaging purposes. This is the case of radionuclides that emit γ-rays or for β^+ emitter that produce, after slowing down in matter and annihilation with an electron, two 511 keV γ-rays emitted back-to-back that can be detected in coincidence (PET imaging). Radionuclides emitting strongly ionizing radiation like charged particles (β^- emitter, α emitter, or Auger electron emitter) will be used for therapy.

To produce these radionuclides, one has to use nuclear reactions. Since the target material is usually made from stable isotopes, the production mode is strongly correlated to the decay mode of the nucleus of interest. Hence, neutron-rich isotopes are preferably produced using neutron-induced nuclear reactions in reactors. This is the case for most of the isotopes decaying by β^- emission. In contrast, proton-rich isotopes are preferably produced using nuclear reactions induced by charged particles provided by accelerators.

The two main tools to produce medical isotopes (research reactors and accelerators) are complementary, offering a large set of radioisotopes for nuclear medicine applications.

17.3.1 Radioisotope Production in Nuclear Reactors

17.3.1.1 Research Reactors

The generic term *research reactor* covers various types of reactors that are not primarily aimed at electricity production. There are prototypes of new reactor concepts (e.g., reactors with fast neutron spectrum and molten salt reactor), reactors for studying certain aspects of reactor physics or accident scenarios, neutron sources for intense extracted neutron beams that serve for neutron scattering and fundamental neutron physics and irradiation reactors that have in-pile irradiation positions with high flux for test irradiations of materials (fuel elements, components of existing, and future reactor types), for neutron transmutation doping of semiconductors, and finally for radioisotope production. Some reactors provide in-pile irradiation positions and extracted neutron beams, while the first listed types of research reactors are usually distinct.

Power reactors are aimed at an efficient conversion of heat to electricity. Therefore, the coolant is extracted at high

pressures and high temperatures, which require appropriate construction materials. The neutron flux in the reactor is deliberately limited in the design to avoid a combination of thermo-mechanical stress with high radiation damage. The massive pressure vessels of most power reactors are built for the entire lifetime of the reactor and cannot be exchanged.

Research reactors that aim at highest possible neutron fluxes are optimized antipodals to power reactors. The aim is to produce the maximum useful neutron flux at limited thermal power generation. Therefore the fuel elements and the entire core are far more compact. Only few kilograms up to tens of kilograms of uranium are used, but this lower mass is compensated by higher enrichment of the ''useful'' fissile ^{235}U. Today, generally so-called *LEU* (low enriched uranium with <20% enrichment of ^{235}U) is used for most research reactors, while those providing the highest neutron flux require higher enrichment (up to 93%). To minimize neutron losses by capture, all materials in and around the core are selected accordingly; thus often heavy water (D_2O) or Be is used as moderator and reflector, and Al or Zr for structural materials. As there is no need to drive a thermoelectric cycle, the average cooling water temperatures are kept at a moderate level, that is, between room temperature and few tens of degree C. Moreover, the water pressure is limited, ranging from the hydrostatic pressure of few meter water cover to few tens of bar for the primary cooling circuit of the most powerful research reactors. Research reactors up to few megawatt power are simply cooled by natural convection; more powerful types require forced convection during operation while natural convection is sufficient already shortly after shutdown. This, together with the tiny mass of fuel, contributes to the exceptional inherent safety of research reactors.

There is no straightforward ranking of the suitability of different research reactors for radioisotope production. Owing to very different core designs, the maximum neutron flux does not scale directly with the nuclear power, and even for comparable maximum flux, the effectively usable flux varies strongly with the position of the available irradiation positions. Reactors with a single compact fuel element (e.g., HFIR, RHF, FRM2) have a very high neutron flux just outside the core (at 10–20 cm distance) and inside the hollow fuel element (e.g., flux trap in HFIR). Reactors with multiple fuel elements (e.g., BR2, HFR, NRU) do not reach the same peak flux but provide far more irradiation positions distributed over a larger volume (Table 17.1).

17.3.1.2 Installations for Radioisotope Production

For radioisotope production, it is essential that the targets can be easily introduced into the irradiation position and recovered after irradiation.

The easiest access is in ''pool-type reactors,'' where the reactor core and control rods are situated at the bottom of an open reactor pool. Six meters of water layer or more acts as efficient radiation shield. Thus, operators can access the top of the pool during reactor operation, and load and unload the irradiation positions from the top with long manipulators or fishing lines.

At higher power, an efficient cooling of the core requires a faster water flow and/or higher pressure. Therefore, the core is situated in a closed cooling circuit operated at higher water pressures of few bar up to few tens of bar. In a ''tank reactor,'' the moderator is placed in the same pressure vessel as the core. Alternatively only the

Table 17.1 Prominent research reactors that are used for radioisotope production, see also the IAEA Research Reactor Database [*http://nucleus.iaea.org/RRDB/*].

Research reactor	Laboratory	Location	Medical isotopes produced	Thermal flux[a] (10^{14} cm^{-2} s^{-1})	Power (MW)
HFIR	ORNL	Oak Ridge, USA	^{188}W, ^{166}Dy, ^{186}Re, ^{188}Re, ^{117m}Sn, ^{195m}Pt, and so on	25	85
BOR-60	SSC RIAR	Dimitrovgrad, Russia	^{89}Sr, ^{153}Gd	20 (Fast flux)	60
SM-3	SSC RIAR	Dimitrovgrad, Russia	^{192}Ir, ^{188}W, ^{153}Gd, ^{60}Co, ^{32}P, ^{33}P, and so on	19	100
RHF	ILL	Grenoble, France	^{188}W, ^{177}Lu, ^{161}Tb, ^{71}Ge, and so on	15	58
BR2	SCK-CEN	Mol, Belgium	^{99}Mo, ^{192}Ir, ^{153}Sm, ^{186}Re, ^{177}Lu, ^{188}W, ^{125}I, and so on	10 (in tank), 3.6 (in pool)	100
HFR	NRG	Petten, Netherlands	^{99}Mo, ^{131}I, ^{133}Xe, ^{192}Ir, ^{177}Lu, ^{89}Sr, ^{186}Re, and so on	4.5	45
MURR	University of Missouri	Columbia, USA	^{177}Lu, ^{166}Ho, ^{153}Sm, ^{33}P, ^{186}Re, ^{188}Re, and so on	4.5	10
HANARO	KAERI	Daejeon, South Korea	^{131}I, ^{125}I, ^{166}Ho, ^{192}Ir, ^{99}Mo, ^{32}P, and so on	4	30
SAFARI	NECSA	Pelindaba, South Africa	^{99}Mo, ^{131}I, ^{192}Ir, ^{177}Lu	4	20
NRU	AECL	Chalk River, Canada	^{99}Mo, ^{131}I, ^{125}I, ^{60}Co	4	135
FRM2	TUM	Garching, Bavaria	^{177}Lu, ^{161}Tb, and so on	4 (Fast flux), 1.3 (thermal flux)	20
OSIRIS	CEA Saclay	Saclay, France	^{99}Mo, ^{192}Ir, ^{153}Sm	2.7	70
OPAL	ANSTO	Lucas Heights, Australia	^{99}Mo, ^{131}I	2.5	20
Maria	Polatom	Świerk-Otwock, Poland	^{99}Mo, ^{131}I, ^{32}P, ^{35}S, ^{192}Ir	2.5	30
RA-3	CNEA	Ezeiza, Argentina	^{99}Mo, ^{131}I, ^{133}Xe, ^{153}Sm, ^{192}Ir	2.4	10
BRR	KFKI	Budapest, Hungary	^{131}I, ^{125}I, ^{153}Sm, ^{32}P, ^{166}Ho, ^{90}Y, and so on	1.7	10
LVR-15	NRI	Řež, Czech Republic	^{99}Mo, ^{153}Sm, ^{166}Ho, ^{177}Lu, ^{192}Ir	1.4	10

[a]The maximum unperturbed thermal neutron flux that is usually accessible for production of medical radioisotopes is given.

core is placed in a pressure vessel, but the surrounding elements stay in an open pool: ''tank in pool reactor.'' Also, heavy-water-moderated reactors use this concept to enclose the precious heavy water in a tank and separate it thus from the light water in the surrounding pool (Figure 17.2).

When producing radioisotopes in tank reactors, the samples have to be loaded into and removed from the tank while the tank is not under pressure, that is, when the reactor is not operating. Thus, such irradiation positions are suitable for production of longer-lived radioisotopes with half-lives commensurate with the operation cycles of the reactor. Alternatively, the tank has to be equipped with dedicated inserts to introduce and remove the irradiation samples during reactor operation. This can be vertical thimbles that are loaded and unloaded as same as irradiation positions in pool reactors or closed circuits with mechanical, pneumatic, or hydraulic transport of the samples.

The difficulty to access positions close to the core during operation for very long cycles and the lower neutron flux are the main reasons why power reactors are not used for the production of medical isotopes.

The technical challenges for radioisotope production targets in nuclear reactors are quite different from cyclotron targets. Unlike charged particle beams, the simple passage of thermal neutrons through a target does not heat the latter significantly. However, the heating by γ-rays from the reactor core (several watts per gram close to the core of high flux reactors) and the self-heating of the sample (mainly for fissile targets that reach kilowatts per gram power) provide significant heat load. Thus, fissile samples or irradiation positions close to the reactor core require usually a forced convection for sample cooling.

For materials with high neutron capture cross section, the self-shielding effect has to be considered. It will reduce the effective neutron flux in the sample. Large strongly neutron-absorbing samples cause also a local flux depression around the irradiation position (''perturbed neutron flux''). Such samples and massive fissile samples may even affect the overall neutron balance of a reactor. This has to be accounted for if loading or unloading of such samples is envisaged during the reactor operation, that is, samples have to be moved slowly to stay in the range of automatic power regulation.

Nuclear reactors are inherently equipped with advanced systems for nuclear safety: multiple safety barriers, including the building, massive biological shielding, tight supervision of gaseous and liquid effluents, assure a safe management of the radioisotope inventory of operating and used reactor cores. The activities, dose rates, and radiotoxicities of samples irradiated for radioisotope production represent usually only a minor fraction of the respective values of the reactor core. Therefore, nuclear reactors are naturally predestined for a safe production of large activities of radioisotopes.

17.3.1.3 Importance of Specific Activity

A straightforward way to produce radioisotopes is via (n,γ) reactions on certain stable isotopes. However, only in few favorable cases, the neutron capture cross-section is high enough that a major fraction of the stable target can be converted in a high neutron flux to the wanted radioisotope. Usually, only a small fraction is converted to radioisotopes and the latter will be ''diluted'' by remaining stable isotopes. One of the most important quality criteria for radioisotopes for nuclear medicine applications is the *specific activity*

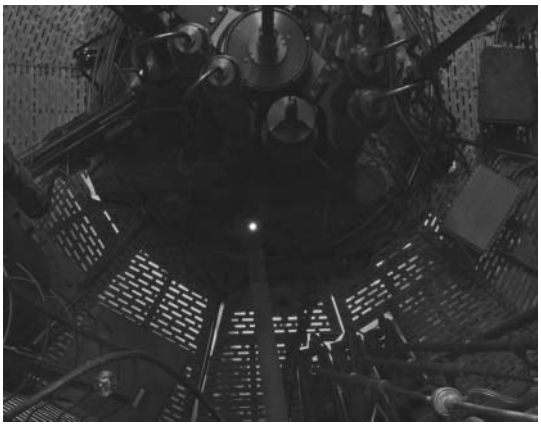


Figure 17.2 Top view of the central part of the high flux reactor of Institut Laue-Langevin seen through 7.5 m of light water. The heavy water tank in the middle is surrounded by a light water pool. The bottom of the latter is "illuminated" by Cherenkov light. The bright point below the center of the image is the V4 thimble tube, providing at its bottom a thermal neutron flux of $1.5 \times 10^{15}\,\mathrm{cm^{-2}\,s^{-1}}$. It can be manually loaded and unloaded with irradiation shuttles during reactor operation. (Photo U. Køster, ILL Grenoble).

S (usually expressed in gigabecquerel per milligram, curies per milligram, or similar units), which is the ratio of activity to mass. The theoretical specific activity of a pure radioisotope is given by

$$S\,\mathrm{max} = \frac{\ln(2) \times N_\mathrm{A}}{T_{1/2} \times M} \tag{17.1}$$

where $T_{1/2}$ is the half-life of the radioactive isotope, M is the molar mass, and N_A denotes Avogadro's constant. For example, an isotope with $M = 100\,\mathrm{g\,mol^{-1}}$ and $T_{1/2} = 7$ days would have a theoretical specific activity of nearly $7\,\mathrm{TBq\,mg^{-1}}$.

If other reactions (such as destruction of the target or product by nuclear reactions) do not interfere significantly, then the specific activity achievable by a one-step transmutation reaction is given by

$$S = \frac{N_\mathrm{A}}{M} \times \sigma \times \Phi \times \left(1 - \exp\left(-\frac{\ln(2) \times t_\mathrm{irr}}{T_{1/2}}\right)\right) \tag{17.2}$$

where σ is the cross section for transmutation of the target into the product, Φ the neutron flux density, and t_irr the irradiation time.

The finally reached specific activity is also determined by the undesired further transmutation (burnup) of the wanted reaction product. This product burnup becomes significant when the neutron capture cross-section of the product is high. For ^{153}Gd, ^{159}Dy, ^{169}Yb, or $^{195\mathrm{m}}$Pt, this seriously limits the achievable specific activity. Taking into account the product and target burnup, the specific activity is given by

$$S = \frac{N_\mathrm{A} \cdot \ln(2) \times \sigma \times \Phi}{M \times T_{1/2} \times \left(\frac{\ln(2)}{T_{1/2}} + (\sigma_\mathrm{p} - \sigma_\mathrm{t})\Phi\right)} \times \left(\exp\left(-\sigma_\mathrm{t}\Phi t_\mathrm{irr}\right) - \exp\left(-\left(\frac{\ln(2)}{T_{1/2}} + \sigma_\mathrm{p}\Phi\right) t_\mathrm{irr}\right)\right) \tag{17.3}$$

Here σ_t denotes the total burnup cross-section of the target isotope (i.e., also including side reactions that do not lead to the product) and σ_p the burnup cross-section of the product isotope.

The specific activity of a radioisotope product is a measure of quality familiar to the users in nuclear medicine. However, the theoretical specific activity varies with the half-life and mass. Hence a comparison of different isotopes and their achievable specific activities is less comprehensible. Therefore in addition, we define the ratio

$R = S/S\,\mathrm{max}$, which indicates how close the specific activity comes to the theoretical optimum. For $R = 0.5$, one radioactive atom will be accompanied by one stable atom (of the same element or target, respectively); for $R = 0.1$, only 1 of 10 atoms is the radioisotope of interest, and so on.

A pure radioisotope ($R = 1$) is called *carrier-free* or no-carrier-added (n.c.a.), and a radioisotope accompanied by a lot of stable isotopes ($R \ll 1$) carrier-added (c.a.).

To decrease the amount of stable isotopes in the final product, one often uses target material that has been enriched in the useful target isotope and depleted in other stable isotopes that do not contribute to the production of the wanted radioisotope. Enrichment is performed via ultracentrifugation of gaseous compounds, electromagnetic mass separation, or selective laser dissociation or resonant laser ionization, respectively.[1)] Most of these techniques are costly and complex; therefore the cost of highly enriched target material can contribute substantially to the overall cost of the produced radioisotopes. Table 17.2 and Table 17.4 show the natural isotopic abundance of the target isotopes as an indication of the importance to use enriched material and an indication of the cost of highly enriched material.

In principle, the specific activity could also be increased after irradiation by a physical mass separation of the radioisotope from remaining stable isotopes. However, all the mentioned enrichment techniques would be even more costly to implement if radioactive material has to be handled; they need time (decay losses!) and have an overall efficiency well below unity (loss of valuable radioisotopes). Therefore, post-irradiation enrichment with physical methods is presently not in commercial use, but is of great interest for R&D applications such as systematic studies of dose–effect relationships as function of specific activity.

Chemical separation methods are often more efficient than physical ones, but for (n,γ) reactions, the produced isotope belongs to the same element, and is therefore chemically identical to the stable target isotope.

As shown in Table 17.2, only in exceptional cases a high specific activity ($R > 0.1$) can be reached by the direct production route, for example, ^{177}Lu. Another way toward higher specific activity is via *indirect production* routes. Here an intermediate radioisotope is produced by (n,γ) reactions, and then decays by β-decay to the wanted daughter that differs in its chemical properties from the target are achieved and can therefore be chemically separated.

If the intermediate isotope is short-lived, a single separation is performed after decay to the longer-lived wanted radioisotope. If the intermediate isotope is longer-lived, it may serve as *generator* for repeated separations of the daughter. After each separation, new daughter activity will grow in and can be separated again. Such generators are particularly interesting because short-lived isotopes can be extracted again and again, even many half-lives after the end of irradiation. Hence, generators are essential to make shorter-lived isotopes easily available at remote places (Table 17.3).

Radioisotopes produced via the indirect route have usually a very high specific activity but they are not necessarily completely monoisotopic. Depending on the chemical purity of the target material and of reagents used for the separation, and owing to possible side reactions producing

1) Enrichment by fractional distillation or chemical methods works well for light elements (H/D, ^{3}He/^{4}He, ^{12}C/^{13}C, ^{16}O/^{18}O, etc.) but these play no important role as target material for medical radioisotope production in reactors.

Table 17.2 Radioisotopes produced directly by neutron capture reactions.

Product isotope	Half-life	Applications	Target isotope	Natural abundance (%)	Theoretical specific activity (TBq mg^{-1})	Specific activity			
						$\Phi = 10^{14}$ cm^{-2} s^{-1}		$\Phi = 10^{15}$ cm^{-2} s^{-1}	
						(GBq mg^{-1})	(*R*)	(GBq mg^{-1})	(*R*)
^{32}P	14.3 d	LDR	^{31}P	100	10.6	0.3	3.1E−5	3	3.1E−4
^{60}Co	5.27 y	EBRT, HDR	^{59}Co	100	0.04	14	3.4E−1	30	7.1E−1
^{64}Cu	12.7 h	PET	^{63}Cu	69	143	4.3	3.0E−5	40	2.9E−4
^{89}Sr	50 d	BM	^{88}Sr	83	1.1	0.004	3.7E−6	0.04	3.7E−5
^{90}Y	64 h	RS, PRRT, RIT	^{89}Y	100	201	0.8	4.2E−5	8	4.2E−4
^{99}Mo	66 h	Generator	^{98}Mo	24	177	0.08	4.5E−6	0.8	4.5E−5
^{103}Pd	17 d	LDR	^{102}Pd	1.0	2.8	1.9	6.8E−4	18	6.6E−3
^{117m}Sn	13.6 d	*BM*	^{116}Sn	15	3.0	0.003	1.0E−6	0.03	1.0E−5
^{153}Sm	46.3 h	BM	^{152}Sm	27	16	80	4.8E−3	640	3.9E−2
^{166}Ho	26.8 h	LDR, RS, BM, RIT	^{165}Ho	100	26	21	8.0E−4	200	7.8E−3
^{169}Er	9.4 d	RS	^{168}Er	27	3.0	0.8	2.7E−4	8	2.7E−3
^{169}Yb	32 d	HDR	^{168}Yb	0.13	0.89	190	2.2E−1	260	2.9E−1
^{177}Lu	6.65 d	PRRT, RIT	^{176}Lu	2.6	4.1	470	1.1E−1	1500	3.7E−1
^{186}Re	3.72 d	BM	^{185}Re	37	7.0	35	5.0E−3	300	4.4E−2
^{188}Re	17 h	HDR, BM, RIT	^{187}Re	63	36	24	6.5E−4	230	6.3E−3
^{192}Ir	73.8 d	EBRT, HDR	^{191}Ir	37	0.34	70	2.0E−1	90	2.6E−1
^{193m}Pt	4.33 d	*Auger*	^{192}Pt	0.78	5.8	0.6	1.1E−4	6	1.1E−3
^{195m}Pt	4.02 d	*Auger*	^{194}Pt	33	6.2	0.009	1.5E−6	0.02	3.3E−6

See Section 17.3.1.4 for the applications (experimental use in italics). The maximum achievable specific activity (saturation activity) was calculated for 100% enriched target isotopes and thermal neutron fluxes of 10^{14} and 10^{15} cm^{-2} s^{-1}, respectively. Experimental values may differ owing to additional resonance capture of epithermal neutrons. Often the irradiation time would be impractically long to reach saturation activity; thus in reality shorter irradiation times are used for higher throughput.

Table 17.3 Reactor-produced radioisotope generators.

Mother	$T_{1/2}$	Daughter	$T_{1/2}$	Applications
^{90}Sr	29 y	^{90}Y	64 h	RS, PRRT, RIT
^{99}Mo	66 h	^{99}Tc	6 h	SPECT
^{166}Dy	82 h	^{166}Ho	27 h	RS, BM, RIT
^{188}W	69 d	^{188}Re	17 h	HDR, RIT
^{191}Os	15 d	^{191m}Ir	5 s	*SPECT*
^{227}Ac	22 y	^{223}Ra	11 d	BM
^{229}Th	7880 y	^{225}Ac	10 d	RIT
^{225}Ac	10 d	^{213}Bi	0.76 h	PRRT, RIT

See Section 17.3.1.4 for the applications.

stable isotopes of the product, there might be traces of ''carrier'' present.

In the indirect route, the final quality (n.c.a.) does, to first order, not depend on the neutron flux. With higher flux, the yield (i.e., the activity produced per mass of target material) will rise, but even at lower flux the quality of product will remain the same. However, in reality it becomes more challenging to achieve a higher separation factor (more stable target material has to be removed) and the requirements to the chemical purity of target material and chemicals used in the separation process rise because traces of stable impurities become relatively more important for a batch with lower activity.

Certain radioisotopes can only be reached from a stable target by two consecutive neutron captures (Table 17.4). To bypass efficiently the intermediate shorter-lived isotope, a particularly high neutron flux is required. The specific activity of the product produced by double neutron capture increases roughly proportional to the square of the neutron flux; therefore irradiation positions with particularly high neutron flux are required to produce high specific activity isotopes via double neutron capture.

Even more successive neutron captures and β-decays are required to breed ^{229}Th from ^{226}Ra targets: ^{226}Ra(n,γ)^{227}Ra (β^-) ^{227}Ac (n,γ) ^{228}Ac (β^-) ^{228}Th(n,γ)^{229}Th. With long irradiation times (years) in a very high neutron flux, ^{229}Th could thus be produced and subsequently serve as generator to provide ^{225}Ac. The latter is an important isotope for targeted α-therapy, either for direct use as ''in vivo generator'' or as generator isotope for the α-emitter ^{213}Bi.

With actinide targets, another reaction becomes important for radioisotope production, namely neutron-induced fission. Here the fission products are chemically different from the target element and can be chemically separated. Consequently, radioisotopes with high specific activity ($R \geq 0.1$) can be produced. Industrial production of ^{99}Mo is today mainly performed by thermal-neutron-induced fission of highly enriched ^{235}U targets, and ^{133}Xe and ^{131}I can be extracted from the same targets.

As for other indirect production routes, the neutron flux is not very relevant for the quality of produced radioisotopes. Large-scale industrial production is performed in high flux reactors at neutron fluxes ranging from 5×10^{13} to 2×10^{14} cm^{-2} s^{-1}. Irradiation in still lower fluxes is possible (e.g., for local or regional supply of ^{99}Mo with medium-flux reactors) but leads to a less efficient use of the ^{235}U targets. Still higher fluxes are not favorable as the cooling of the fissile targets gets increasingly difficult.

Not every research reactor is equipped for irradiation of massive fission targets. This requires additional investments such as forced cooling, local monitoring of fission gas release, additional shielding, and correspondingly heavy equipment for handling the heavily shielded transport containers.

Table 17.4 Radioisotopes produced by double neutron capture or indirectly by single neutron capture reactions.

Product isotope	Half-life	Applications	Target isotope	Natural abundance (%)	Intermediate isotope	Half-life	$\Phi = 10^{14}$ cm^{-2} s^{-1}			$\Phi = 10^{15}$ cm^{-2} s^{-1}		
							t_{irr} (d)	t_d (d)	Yield (GBq mg^{-1})	t_{irr} (d)	t_d (d)	Yield (GBq mg^{-1})
Production via (2n,γ) reactions												
^{166}Dy	3.4 d	Generator	^{164}Dy	28	^{165}Dy	2.35 h	10	1	2	5	1	100
^{188}W	69.8 d	Generator	^{186}W	28	^{187}W	0.99 d	100	10	0.002	50	10	0.1
Indirect production via (n,γ)β$^-$												
^{47}Sc	3.3 d	*PRRT, RIT*	^{46}Ca	0.004	^{47}Ca	4.5 d	10	1	0.5	10	1	5
^{125}I	59.4 d	LDR	^{124}Xe	0.10	^{125}Xe	17 h	7	7	6	4	7	20
^{131}I	8.0 d	Thyroid, MIGB	^{130}Te	34	^{131}Te	25 min	28	2	0.1	28	2	1
^{131}Cs	9.7 d	LDR	^{130}Ba	0.11	^{131}Ba	12 d	7	7	0.7	7	7	7
^{161}Tb	6.9 d	*PRRT, RIT*	^{160}Gd	22	^{161}Gd	4 min	14	0.5	0.4	14	0.5	4
^{177}Lu	6.7 d	PRRT, RIT	^{176}Yb	12.8	^{177}Yb	1.9 h	14	1	0.6	14	1	4
^{199}Au	3.1 d	*PRRT, RIT*	^{198}Pt	7.2	^{199}Pt	31 min	7	0.5	0.7	7	0.5	7
^{227}Ac	21.7 y	Generator	^{226}Ra	0	^{227}Ra	42 min	100	30	0.02	28	30	0.03
Indirect production via (n,f)												
^{99}Mo	2.8 d	Generator	^{235}U	0.72	–	–	7	1	5.7	–	–	–
^{131}I	8.0 d	Thyroid, MIGB	^{235}U	0.72	–	–	7	12	0.7	–	–	–
^{133}Xe	5.3 d	SPECT	^{235}U	0.72	–	–	7	7	3	–	–	–

For the applications see Section 17.3.1.4 (experimental use in italics). The achievable yields (gigabecquerel of product isotope per milligram of 100% enriched target isotope, not to be confounded with specific activity of the product!) were calculated for thermal neutron fluxes of 10^{14} and 10^{15} cm^{-2} s^{-1}, respectively. Experimental values may differ owing to additional resonance capture of epithermal neutrons. Typical irradiation and decay times were chosen to keep coproduced radioisotope impurities at an acceptable level.

Nearly 100 kCi (3700 TBq) of ^{99}Mo have to be produced every week to cover the worldwide needs of ^{99}Mo/^{99m}Tc. For the past decades, most of this was produced as ''fission-moly'' by fission of highly enriched ^{235}U targets. After irradiation for about six days and decay of short-lived activities for one day, about 10% of the molybdenum fraction is ^{99}Mo, that is, fission-moly has high specific activity, which facilitates the use of a so-called chromatographic generators where the ^{99}Mo is adsorbed on acid alumina surfaces and ^{99m}Tc is eluted. Until recently, the main producers of fission-moly were the reactors NRU, HFR, BR2, SAFARI, and OSIRIS. Unplanned outages of NRU and HFR caused world-wide shortage of ^{99}Mo/^{99m}Tc. Since then additional reactors started medium-scale fission-moly production (Maria, LVR-15) and other reactors are being backfitted for fission-moly production (FRM2) or under construction (RJH). Production at SAFARI, OPAL, RA-3, and other smaller reactors is now based on LEU targets. Production for local needs is also performed by neutron capture on ^{98}Mo but the produced ^{99}Mo with low specific activity requires different, less practical, generator technologies.

Various alternative production methods are under investigation, but the enormous quantities required and the low market prices for ^{99}Mo did not yet allow the emergence of a real alternative to fission-moly production in research reactors [9].

Higher energy neutrons with hundreds of kiloelectronvolts to megaelectronvolts are particularly useful for reactions other than (n,γ). Thus, (n,p) and (n,α) reactions can populate β^--decaying isotopes that are chemically different from the target material. Hence high specific activities may be achieved by chemical separation. ^{32}P, ^{33}P, ^{35}S, ^{47}Sc, ^{67}Cu, and ^{89}Sr are produced with (n,p) reactions on the corresponding targets, and ^{33}P is also produced via ^{36}Cl(n,α).

Significant fluxes of fast neutrons are available in few dedicated fast reactors (BOR60) or in irradiation positions in the core or very close to the core of multi-purpose reactors (SM3, HFIR, BR2, FRM2).

17.3.1.4 Medical Applications of Reactor-Produced Radioisotopes

Reactor-produced radioisotopes cover a variety of medical applications with very different requirements with respect to the specific activity:

1. *Targeted therapy with β- and α-emitters.* For tumor cell types that have only few selective sites (peptide receptors or antigens respectively), it is crucial to use radioisotopes with very high specific activity. Otherwise the selective sites would be blocked with tumor-seeking agents carrying ''cold'' (stable) isotopes that have no therapeutic effect. Owing to their higher radiotoxicity, therapeutic activities of α-emitters are orders of magnitude lower compared to β-emitters; therefore correspondingly lower specific activity is acceptable. Two prominent α-emitters can be produced in research reactors: ^{227}Ac that serves as generator for ^{223}Ra, and ^{229}Th that serves as generator for ^{225}Ac.
2. *Auger electron therapy*. The logic continuation of the increase in selectivity from β- to α-emitters would be Auger electron emitters. Low-energy Auger electrons have a very short-range of nanometers to micrometers. Hence, they show high radiotoxicity only when specific internalizing agents carry them to the cell's nucleus, while non-targeted activity remains rather harmless. Thus, Auger emitters such as ^{117m}Sn, ^{125}I,

^{161}Tb, ^{193m}Pt, or ^{195m}Pt are very promising, but it is challenging to find suitable internalizing agents [10].

3. *Bone metastases* (*BM*). Many cancers (prostate cancer, breast cancer, etc.) are developing in a later stage frequently painful BM. Compared to normal, slowly growing bone, these metastases show often an enhanced metabolism of calcium and phosphates. This can be exploited to target them metabolically with bone-seeking radioisotopes. Chemical homologues of calcium, such as ^{89}Sr or ^{223}Ra, as well as phosphates or bisphosphonates carrying ^{32}P or stable complexes of radioactive metals (^{117m}Sn, ^{153}Sm, ^{186}Re, ^{188}Re) show an enhanced uptake in the BM and are used for palliative therapy [11]. A moderate specific activity is sufficient for such applications.
4. *Radiosynovectomy* (*RS*). Radiation synovectomy (or synoviorthesis) uses the intra-articular injection of colloids containing radioisotopes to treat severe inflammations (e.g., arthritis) of the flexible joints. Pure β-emitters without γ-ray emission (or with weak γ-ray emission) are employed where the β energy is adapted to the size of the joint: ^{90}Y for large joints (hip, knee), ^{186}Re for mid-sized joints (elbow), and ^{169}Er for small joints (finger). A low-to-moderate specific activity is sufficient for such applications.
5. *Radiotracer applications*. There is also a multitude of radiotracer applications in nuclear medicine. Here small activities of radioisotopes (^{14}C, ^{32}P, etc.) are given to the patient, then a sample (blood, urine, etc.) is taken, and the measured radiotracer activities allow quantifying metabolic activity or detect malfunctions. Usually low specific activities are sufficient.
6. *PET isotopes*. For completeness, it should be noted that even the PET isotope ^{18}F can be produced in nuclear reactors. Here, a two-step reaction is exploited in targets containing ^{6}Li and ^{16}O (Li_2CO_3 or LiOH respectively) where the energetic tritons (2.74 MeV) liberated in the ^{6}Li(n,α)t reaction can induce exothermal ^{16}O(t,n)^{18}F reactions. Batch yields up to 10 GBq ^{18}F have been reported [12].
7. *External beam radiation therapy* (EBRT). Intense sources containing long-lived emitters of high-energy γ-rays are used to provide collimated γ-ray beams (“γ-knife”) for EBRT (belonging formally to the discipline of radiology and not to nuclear medicine). This alternative to Bremsstrahlung from dedicated electron accelerators is frequently used in emerging countries. The most used isotopes are ^{60}Co (“cobalt bomb”) and ^{192}Ir. Specific activity is not decisive for this application; therefore medium flux reactors can be used for the production.
8. *Brachytherapy*. In brachytherapy (also formally part of radiology), sealed radiation sources are introduced into the body (e.g., in or around a tumor) or close to the body (e.g., for skin and eye treatment). One distinguishes between high dose rate (HDR) and low dose rate (LDR) brachytherapy.
   (a) In *HDR brachytherapy*, the required dose is given in short time (minutes) by an intense source (e.g., 370 GBq of ^{192}Ir). Very compact sources (e.g., metal disks or wires) are required for this application; hence the source needs a relatively high specific activity. ^{60}Co, ^{153}Gd, ^{169}Yb, and ^{192}Ir activated by long irradiations in high flux reactors are employed for this purpose. Also,

^{166}Ho and ^{188}Re are used, either in liquid form at high concentration to fill balloon catheters for intravascular applications or applied as ''cream'' or ''patch,'' to treat skin cancer. ^{188}Re at high concentration is conveniently obtained from ^{188}W generators with high specific activity.

(b) In *LDR brachytherapy*, the source remains for a longer time in the body, from days to weeks or even permanently. To avoid unnecessary dose outside the treatment area, emitters of short-range radiation (X-rays, low-energy γ-rays, and β-particles) are used. Longer-lived isotopes such as ^{103}Pd, ^{125}I, ^{131}Cs, or ^{32}P are fixed in the form of needles (''seeds'') or tubes (''stents''). A particular application is the treatment of liver cancer. Here small radioactive spheres containing ^{90}Y or ^{166}Ho or colloids containing ^{188}Re are directly injected into the hepatic artery supplying the tumor, thus destroying the ''supply lines'' of the latter. A low-to-moderate specific activity is usually sufficient for LDR brachytherapy.

17.3.2 Radioisotope Production in Accelerators

17.3.2.1 Accelerators

Accelerators for radionuclide production can be classified into four different classes with respect to the projectile energy and the type of available projectile [13]. The first class corresponds to a single particle (proton or deuteron) accelerator with a maximum incident energy available equal to 10 MeV. The second class corresponds to accelerators allowing the acceleration up to 20 MeV of protons and/or deuterons. The third class corresponds to multi-particle accelerators with a limit of 50 MeV. Finally, the last class corresponds to powerful accelerators with projectile energy higher than 50 MeV.

Most of the machines of class I and II are cyclotrons. These first two classes are known as *biomedical accelerators*. They are well suited for the production of ^{18}F as well as many of the β^+ emitters of interest. The great advantage of this kind of accelerators is that they are easy to manage and because of the low-energy, only one or a few reaction channels are open, which simplify the processing of the targets after irradiation. This kind of machine can be bought from several manufacturers worldwide along with all the associated equipments necessary to produce the radiopharmaceutical.

The last two classes correspond to accelerators that are not only used for radionuclide production but often also for other purposes such as nuclear or particle physics research. Many of the accelerators of class 4 are linear accelerators LINACs shared with physics research programs. For example, part of the beam provided by the injector of the LANSCE (Los Alamos Neutron Science Center) spallation neutron source is used to produce isotopes with 100 MeV protons. Recently, 70 MeV cyclotrons have been developed by several companies (Figure 17.3). The main advantage of this type of accelerator when compared to LINACs is the compactness and the limited number of operators needed. A high energy accelerator is needed for some of the radioisotopes of interest (^{82}Sr, e.g., which is the mother nucleus of ^{82}Rb used in cardiology) and allows to explore alternative production routes such as (p,2p) for ^{67}Cu production or spallation [14]. Finally, it is also possible to

Figure 17.3 View of the ARRONAX cyclotron in its main cave. It accelerates protons and α-particles up to 70 MeV. The cyclotron has a diameter of 4 m and a total height (including the ion sources) of 4 m. Two proton beams can be extracted simultaneously and directed to one of the six experimental vaults [16]. (Photo C. Huet, GIP ARRONAX)

produce a high flux of secondary neutrons that can be used for activation [15].

17.3.2.2 Radioisotope Production with Charged Particles

Besides the fact that proton-rich isotopes are most easily produced using accelerators, the use of charged particles as projectiles allows to have usually a target and a product that are different chemical elements, enabling a chemical separation of the product from the target.

The main parameter that is involved in the production of a radioisotope is the cross section associated to the nuclear reaction involved in the production mechanism. The cross section represents the probability for the reaction to take place and is expressed in barn (1 b is equal to 10^{-24} cm^2). For low-energy-charged particle induced reactions, cross sections evolve from few millibarn to hundreds of millibarn depending on the isotope and projectile of interest. Data about many cross sections can be found in the literature (EXFOR database). The reaction cross section evolves with the energy of the projectile, and it is necessary to define the best projectile energy in order to optimize the production yield of a radioisotope through a given reaction channel. In Figure 17.4, the ^{211}At production cross section via the ^{209}Bi(α,2n)^{211}At reaction is presented as a function of the energy. This curve is also called *excitation function*. The cross section value is non zero above $E = 20.7$ MeV, which corresponds to the energy threshold of the reaction. This means that below this threshold value, no ^{211}At can be produced by sending α-particles on a natural bismuth target. The energy threshold gives the lower possible energy to produce the radionuclide of interest using the chosen production route. The cross section reaches then a broad maximum of about 1 b between 28 and 32 MeV, and decreases to an asymptotic value at high energy. This curve indicates that with this nuclear reaction, it is better to use projectiles with incident energy around 30 MeV.

The produced activity, A in becquerel, at the end of bombardment (EOB) can be estimated from the following relation:

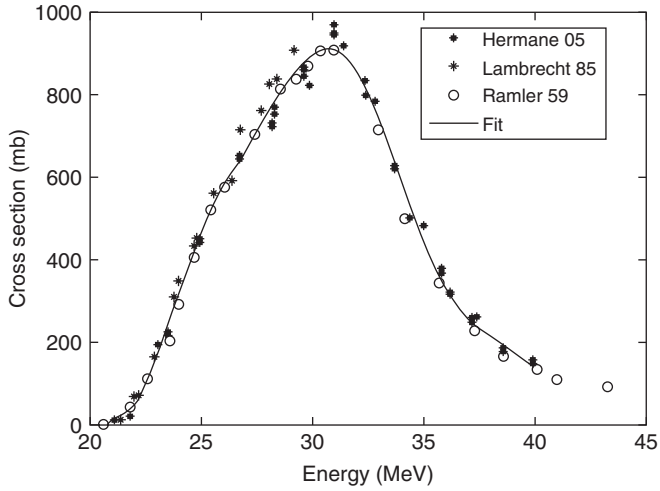


Figure 17.4 ^{211}At experimental production cross section via the ^{209}Bi(α,2n) reaction.

$$A = \Phi_0 \frac{\rho \times \chi}{M} N_A (1 - \exp(-\lambda t)) \times \int_{E_{min}}^{E_{max}} \frac{\sigma(E)}{\frac{dE}{dx}} dE \quad (17.4)$$

where Φ_0 represents the projectile flux (particle $s^{-1}\,cm^{-2}$), ρ the target density ($g\,cm^{-3}$), χ the enrichment, M the molar mass of the target, N_A the Avogadro number, λ the radioactive decay constant of the product (s^{-1}), $\sigma(E)$ the production cross section (cm^2), dE/dx the specific energy loss of the projectiles in the target material ($MeV\,cm^{-1}$), and t the irradiation time (s). E_{max} corresponds to the energy of the projectiles when entering the target, whereas E_{min} corresponds to the energy of the projectiles leaving the target. E_{min} is directly connected to the target thickness. If the projectile is stopped within the target, E_{min} is equal to 0.

From the previous formula, one can see different ways to increase the production yield.

One can increase the irradiation time. As it is involved through an exponential term, the production yield is almost proportional to the time when the irradiation time is lower than the half-life of the desired radionuclide, and it saturates after few half-lives of the desired radionuclide. In most cases, in accelerators, targets are irradiated no longer than one half-life of the desired isotope.

One can increase the number of target atoms by choosing the most appropriate chemical form (e.g., rubidium can be used as RbCl or Rb metal) or/and by using enriched material. In this latter case, it is necessary to recover the target material after the extraction and purification process to reduce the production cost. The use of enriched material also allows us to limit the production of contaminants from side reactions during the irradiation and to increase the specific activity of the final product.

One can increase the beam intensity on target (the projectile flux). Such increase is strongly constrained by the thermal properties of the target. Indeed, the beam interacts with the target mostly through ionization, and loses energy by heating up

the target. The power (in watts) dissipated by a beam in heating up a target is directly obtained by multiplying the energy loss in the target (in megaelectronvolts) by the beam intensity (in microamperes). For example, with the ARRONAX cyclotron, 350 μA proton beam at 70 MeV can deposit 24.5 kW over few square centimeters [16]. It is then of a great importance to be able to remove a large part of this energy by a proper cooling of the targets. Associated to an elevation of the temperature are phase changes and density changes. In the first case, the target integrity may be lost, whereas in the second case, the production yield may decrease because density is also one parameter of the production formula.

One must choose the appropriate reaction channel, beam energy, and target thickness (energy range). This is made by looking at the cross section values. It is important to consider the maximum value of the cross section but also its evolution with energy because when using thick targets, one integrates the production over the energy range of the projectile. For example, in some cases, deuteron-induced reaction will end up with higher production yields than protons for the same maximum value of the cross section. This is linked to the shape of the cross section as function of the projectile energy which is broader in the deuteron case. Correct choices of the beam energy and target thickness also have an impact on the level of contaminants that are created during irradiation through the different competing reaction channels. Whatever be the contaminant, it is better to try to avoid its production during irradiation, instead of putting effort to remove it during the post-irradiation chemical treatment of the target. In case of isotopic impurities, chemical separation is of no help, and these atoms will have an impact on the specific activity or radionuclide purity of the final product. In some cases, when the half-lives of the isotopes are very different, it is possible to use a decay time between the end of irradiation and the chemical processing to remove unacceptable short-lived contaminants.

The optimal production conditions correspond to a compromise because these different parameters are not independent. As an example of the optimization of irradiation parameters, we can again use the case of ^{211}At production. The interaction of the projectile with the bismuth target can create several astatine isotopes depending on the projectile energy. ^{209}At ($T_{1/2} = 5.4$ h) has a threshold of 35.9 MeV, which is above the maximum production cross section of ^{211}At. Limiting the beam energy allows us to avoid the production of this isotope with a limited effect on the ^{211}At production yield. ^{212}At ($T_{1/2} = 0.3$ s) is not a problem because it has a short half-life and at the moment of the chemical separation, all nuclei will have decayed to ^{208}Bi. Only ^{210}At ($T_{1/2} = 8.1$ h) is of a great concern because its half-life is close to that of ^{211}At, and its production starts before that of ^{211}At reaches its maximum value (energy threshold = 28.6 MeV). It decays mainly by electron capture (99.8%) to ^{210}Po, an α-emitter with 138 days half-life, which, as a bone seeker, causes great radiation safety issues. Therefore, the beam energy is usually limited to about 28 MeV [17].

The irradiation time is usually few hours, and there is no need of enrichment because bismuth is mono-isotopic.

Low-energy reactions have usually narrow excitation functions that populate just a single product isotope. With rising projectile energy, more reaction channels are opened, hence leading to a mixture of different product nuclei. Thus, the gain

in usable target thickness and higher product yield has to be paid by a reduction in selectivity. When the wanted product is (much) longer-lived than unwanted side-products of the same element, a sufficient radioisotopic purity can be achieved simply by a chemical separation combined with an appropriate delay between production and use (e.g., spallation production of ^{67}Cu). Moreover, the use of generators provides sufficiently pure products that are barely affected by coproduced side-products, even if the latter are longer-lived (e.g., ^{82}Sr/^{85}Sr).

In the extreme case of very high projectile energy (gigaelectronvolt protons), very thick targets (tens of grams per square meter) can be irradiated, and via fragmentation, spallation, and fission reactions virtually all isotopes from hydrogen up to the target element are produced simultaneously. Still, with a combination of mass separation and chemical separation, monoisotopic samples can be prepared. In the so-called isotope separation on-line (ISOL) method, even very short-lived isotopes are accessible [18]. Although not yet used for large-scale production, this method provides, for example, at the ISOLDE facility at CERN a selection of over 1000 different radioisotopes in non-carrier-added quality, thereof several with very promising decay properties for nuclear medicine applications. Thus, unusual radioisotopes are already available for radiochemical and preclinical R&D, while dedicated production methods at smaller accelerators are being developed. An example is the unique matched quadruplet of terbium isotopes where ISOLDE provides the PET isotope ^{152}Tb, the SPECT isotope ^{155}Tb, and the α-emitter ^{149}Tb. These can be labeled to the same vector guaranteeing identical in vivo behavior, and the therapeutic efficacy of ^{149}Tb can be directly compared to that of the reactor-produced β^-/Auger emitter ^{161}Tb [19].

17.3.2.3 Prominent Radioisotopes Produced with Accelerators

Prominent PET, SPECT, and therapy isotopes produced in accelerators are shown in Tables 17.5–17.7. The natural abundance of the target isotopes is given to demonstrate the varying cost and importance of enriched targets. For example, ^{68}Ge and ^{82}Sr are produced with non-enriched natural targets. Typical E_{max} and E_{min} are given to cover a good part of the excitation function. Obviously, projectiles with energy lower than E_{max} can also be used when compromising the achievable yields.

17.3.2.4 Targetry

For medical applications, it is necessary to produce large quantities of radionuclides on a regular basis under good quality control. The irradiation system must be designed for that purpose as well as for the processing laboratories. Depending on the physical nature of the target, different solutions can be considered [20].

In the cases of liquid targets (e.g., production of ^{18}F) and gas targets (e.g., production of ^{11}C), internal targets inserted between the pole shoes of a cyclotron are used most of the time, leading to compact facilities without beam lines. Transportation of activated material after irradiation is quite straightforward, and uses tubes to bring activated targets to the shielded cells (hot cells) for processing. The target material is placed in a container during irradiation. The material of the container must be studied in order not to interact chemically with the different species created during the irradiation. A thin window separates the target material from the vacuum. During irradiation, the container

Table 17.5 Positron emitters produced in accelerators with decay properties and common production paths.

Product isotope	Half-life	$E(\beta^+)$ average (MeV)	$E(\beta^+)$ maximum (MeV)	B.R. β^+ (%)	Applications	Target isotope	Natural abundance (%)	Reaction	Excitation curve	
									E_{max} (MeV)	E_{min} (MeV)
PET isotopes (direct production)										
^{11}C	20.4 m	0.96	0.39	99.8	Acetate, choline, and so on	^{14}N	99.6	p,α	20	4
^{13}N	10.0 m	1.20	0.49	99.8	NH_3 perfusion	^{16}O	99.8	p,α	20	7
^{15}O	122 s	1.73	0.74	99.9	H_2O perfusion	^{15}N	0.36	p,n	12	4
						^{14}N	99.6	d,n	12	3
^{18}F	110 m	0.63	0.25	96.7	FDG, F^-, F-DOPA, F-MISO, and so on	^{18}O	0.2	p,n	20	2.5
^{44}Sc	3.92 h	1.47	0.63	94.3	Peptides, antibodies	^{44}Ca	2.1	p,n	20	5
^{64}Cu	12.7 h	0.65	0.28	17.6	ATSM, peptides, antibodies	^{64}Ni	0.93	p,n	20	4
^{86}Y	14.7 h	1.55	0.66	31.9	Peptides, antibodies	^{86}Sr	9.9	p,n	16	7

^{89}Zr	78.4 h	0.90	0.40	22.7	Peptides, antibodies			^{89}Y	100	p,n	20	5
^{124}I	99.6 h	2.14	0.82	22.8	Antibodies			^{124}Te	4.7	p,n	16	6
PET isotopes (generator)						**Generator isotope**	**Half-life**					
^{44}Sc	3.92 h	1.47	0.63	94.3	Peptides, antibodies	^{44m}Sc	2.44 d	^{44}Ca	2.1	d,2n		
^{44}Sc	3.92 h	1.47	0.63	94.3	Peptides, antibodies	^{44}Ti	60 a	^{45}Sc	100	p,2n	35	14
^{62}Cu	9.7 m	2.93	1.31	97.4	ATSM	^{62}Zn	9.1 h	^{63}Cu	69	p,2n	30	18
^{68}Ga	67.6 m	1.90	0.83	89.1	Peptides, bone scan, and so on	^{68}Ge	271 d	^{69}Ga	60	p,2n	35	13
^{82}Rb	76 s	0.80	0.36	21.1	Rb^+ perfusion	^{82}Sr	25.3 d	^{85}Rb	72	p,4n	100	40

Table 17.6 SPECT isotopes produced in accelerators with decay properties and common production paths.

Product isotope	Half-life	$E\gamma$ (keV)	B.R. (%)	Applications	Target isotope	Natural abundance (%)	Reaction	Excitation curve	
								E_{max} (MeV)	E_{min} (MeV)
SPECT isotopes (direct production)									
^{67}Ga	78.3 h	93	42	Ga^{3+} for inflammation scan	^{68}Zn	19	p,2n	40	13
^{99m}Tc	6.01 h	141	89	Various, see text	^{100}Mo	9.7	p,2n	24	9
^{111}In	67.4 h	245	94	Peptides, antibodies, and so on	^{112}Cd	24.1	p,2n	30	12
		171	91	In-oxine for white blood cell scan					
^{123}I	13.2 h	159	83	Thyroid, MIBG, antibodies, and so on	^{123}Te	0.89	p,n	15	8
					^{124}Xe	0.10	p,2n > EC	35	17
^{201}TI	73.1 h	70	59	Perfusion	^{203}TI	29.5	p,3n > EC	36	20
SPECT isotopes (generator)					**Generator isotope**	**Half-life**			
^{81m}Kr	13.1 s	190	68	Lung ventilation	^{81}Rb	4.58 h	^{82}Kr	11.6	p,2n	30	15
^{99m}Tc	6.01 h	141	89	See text	^{99}Mo	65.9 h	^{100}Mo	9.7	p,pn	70	20

Table 17.7 Decay properties of therapy isotopes produced in accelerators (see columns to the right), reactors (Tables 17.2 and 17.4), and from generators (^{212}Bi, ^{213}Bi, ^{223}Ra, ^{225}Ac), respectively.

β⁻ (and conversion electron) emitter											
Isotope	Half-life (d)	β range average (mm)	E β⁻ (MeV)	β range maximum (mm)	E_γ (keV)	I_γ (%)	Target isotope	Natural abundance (%)	Reaction	Excitation curve	
										E_{max} (MeV)	E_{min} (MeV)
^{47}Sc	3.35	0.32	0.60	2.4	159	68	^{48}Ti	73.7	p,2p	≫100	20
^{67}Cu	2.58	0.26	0.56	2.3	185	49	^{68}Zn	19.0	p,2p	≫100	26
^{89}Sr	50.5	2.3	1.50	7.8	–	0	–	–	–	–	–
^{90}Y	2.67	4.4	2.28	12	–	0	–	–	–	–	–
^{117m}Sn	13.6	0.23	0.16	0.3	159	86	^{116}Cd	7.5	α,3n	50	24
^{131}I	8.02	0.39	0.81	3.6	364	82	–	–	–	–	–
^{153}Sm	1.94	0.55	0.81	3.6	103	29	–	–	–	–	–
^{161}Tb	6.9	0.29	0.59	2.4	75	10	–	–	–	–	–
^{166}Ho	1.12	2.8	1.86	10	81	6.6	–	–	–	–	–
^{169}Er	9.4	0.14	0.35	1.1	–	0	–	–	–	–	–
^{177}Lu	6.65	0.23	0.50	1.9	208/113	10/6	–	–	–	–	–
^{186}Re	3.72	1.1	1.07	5.2	137	9.5	^{186}W	28.4	p,n	20	7
^{188}Re	0.71	3.3	2.12	12	155	15.6	–	–	–	–	–
^{199}Au	3.14	0.10	0.45	1.6	158	40	–	–	–	–	–

is cooled using water. Often, the container window is cooled using helium gas.

In the case of solid targets (e.g., production of ^{64}Cu), external targets are used most of the time, requiring an external target station and even, in some cases, beam lines. Transportation of the target from the irradiation station to the hot cells is made remotely using, in many cases, a pneumatic system. Targets can be made of layers deposited by electroplating or physical vapor deposition (evaporation under vacuum), encapsulated pressed powder, or foils. Targets are cooled during irradiation. In some cases, the water cooling is only on the back of the target, whereas helium cooling can be applied on the front of the target. In other cases, when the thermal properties of the material are not very good, the target is cooled on all its faces using water. In this latter case, a window is used to separate the vacuum from the cooling water. It is also necessary to take into account the interaction of the beam with the water to define the projectile energy to be used. In order to further increase the beam current on target, it can be interesting to tilt the target with respect to the beam axis. This results in a thinner target thickness and in a larger interacting area between the beam and the target material. The first point allows having a better thermal conduction by reducing the path, whereas the second point reduces the power density on target. Both effects are favorable for the heat dissipation from the target.

After the irradiation, extraction and purification of the radionuclide of interest must be done under good manufacturing practice (GMP). In order to be able to use a product in humans, sterility and freedom of pyrogens must be ensured. This requires working in a clean and controlled environment [21].

17.4 Preclinical Radiolabeling and Biodistribution Validation

Radiopharmaceuticals, in the same manner as any other drugs, must pass a series of tests before they may be authorized for human use. Once they are qualified in terms of purity and stability, they are assessed for safety and efficacy. At this stage, there is no predictive *in vitro* method to demonstrate that the radiopharmaceutical will be well tolerated and will efficiently target radioactivity to the expected sites of action, either for imaging or for therapy. Thus, before clinical trials, the candidate radiopharmaceutical is tested in animal models, even though animal models cannot be fully predictive [22]. However, they can reveal non anticipated toxicity, inappropriate distribution in a living animal, or a lack of efficacy.

In practice, there are two stages in these preclinical studies. One is the assessment of performance, aiming at selecting candidates with some probability of efficacy for imaging or therapy. This stage primarily interests the academic laboratory or industry that develops the new radiopharmaceutical, and methods are fairly flexible. The second is the assessment of safety and toxicity. It involves series of tests that are more standardized and must be run according to good laboratory practice to be acceptable by regulatory authorities, which are primarily concerned by the safety of clinical trials. These safety and toxicity assessments usually involve *in vitro* testing for mutagenesis and *in vivo* testing in at least two animal species, one rodent (mouse or rat) and one bigger animal (dog or non-human primate). They include repeated administration of high doses of the unlabeled molecule over weeks or months while monitoring the

animal health, followed by sacrifice and extensive pathological examination [23]. Of course, when prior knowledge of any risks is associated with a specific drug category, these risks are specifically addressed in additional animal testing. In general, adverse side effects come from unexpected properties of the drug, and then the idea is to test it in a sufficient number of animals, well above the corresponding cumulated dose that will be administered to human beings, over a period of time long enough for potential side effects to be discovered. New regulations allow for a more limited toxicology evaluation when the total dose of drugs to be used in humans does not exceed 100 μg. This may be the case for diagnostic imaging radiopharmaceuticals or for early clinical assessment of new therapeutics (phase 0 clinical trials).

For radiopharmaceuticals, the radioactivity is generally the major source of hazard. However, the risk is considered marginal for diagnostic applications with low LET emissions and relatively low activities. On the basis of animal biodistribution studies, it is generally possible to evaluate tissue absorbed doses and to show that they are well below doses at risk of secondary effects. For therapeutic radiopharmaceuticals, in addition to dosimetry considerations, the maximum tolerated dose must be assessed in rodents or, if possible, in more relevant animal models, but this assessment may be conducted in academic laboratories not necessarily complying with good laboratory practice. From these preclinical studies, a dose considered as safe, both from the chemical and radiological point of view, is determined as the first, lower dose that will be used in the first administration of the radiopharmaceutical in humans.

Preclinical testing also involves a pharmacokinetic evaluation of absorption (when the drug is not given by direct intravenous injection), distribution (often limited to a blood pharmacokinetic analysis that shows to what extent the drug distributes out of blood volume), metabolism, and elimination (ADME – absorption, distribution, metabolism, and elimination) [24]. Most often radiopharmaceuticals are injected intravenously, but radioactivity may be measured easily in all tissues, and distribution (called *biodistribution*) studies are usually conducted in great detail. Metabolism and elimination studies are also conducted. It is obviously important to know how the radiopharmaceutical will be metabolized. However, since the active part of the radiopharmaceutical is the radioactivity, studies concentrate on monitoring the radioactivity. The elimination of radioactivity is also very important. It occurs both by excretion (in urine, feces, or other routes) and by radioactive decay.

For radiopharmaceuticals, this pharmacokinetic assessment is very important because there are clear relationships between irradiation doses delivered to the various tissues and risks of toxicity, but also potential efficacy, when the goal is to deliver high irradiation doses to tumor cells. In cardiology, experimental heart disease may be pharmacologically or surgically caused, or heart transplantation rejection may be studied by grafting a second, incompatible, heart in rats. Similarly, in neurology, a pathologic situation similar to human diseases, particularly neurodegenerative diseases, may be created by pharmacological or surgical destruction of particular areas of the animal brain. In oncology, human tumor grafting in immuno-compromised animals, most often nude or severe combined immunodeficiency (SCID) mice, is by far the most frequent approach. Grafts may be either sub-cutaneous (a very simple, but rather

artificial model) or orthotopic (which means that tumor cells are inoculated in the tissue of origin). Residual disease, that is, the presence of tumor cells in small numbers after surgery, as often observed in ovarian or bladder cancers, may be mimicked by injection of tumor cells in the peritoneum cavity or in the bladder. In that case, the radiopharmaceutical may be injected locally for therapy. Alternatively, the tumors may be of mouse origin and administered in immuno-competent animals. This is by far a more physiological situation, but several radiopharmaceuticals, particularly radiolabeled antibodies, are specific of human tissues. Animal models of spontaneous diseases are also developed, either by specific food regimen, chronic pharmacologic treatment, or genetic manipulation (knocking out tumor suppressor genes, knocking in tumor susceptibility genes, or oncogenes). These models go a big step further in similarity with human diseases, but are more difficult to control and more expensive. Finally models of truly spontaneous disease begin to be studied, such as dogs or cats with spontaneous tumors presented for veterinary care. Thus a wide variety of models are available for preclinical testing, but of course, none provides any certainty that a given drug, selected from its performance in such models, will be safe and efficacious in humans.

Diagnostic radiopharmaceuticals are expected to bind to pathological tissues (e.g., tumors) or concentrate in body areas where a given physiologic condition occurs (e.g., blood perfusion, glucose metabolism, hypoxia). They are also expected to give good contrast in molecular imaging, that is, in SPECT or PET imaging. Then the activity uptake ratio between the area of interest and surrounding tissues is the most important factor. Preclinical studies aimed at evaluating this ratio and finding the best conditions in terms of injected dose and time between injection and imaging. The presence of normal or non-specific uptake in some tissues may not be a problem for the imaging application. For instance, in PET imaging with ^{18}FDG, the normal uptake in brain tissue is not considered a problem for imaging tumors, except brain tumors. The performance is estimated in terms of sensitivity and specificity.

For therapy, both efficacy and toxicity is determined primarily by the number of disintegrations in tumor and normal tissues. Thus high target uptake, good elimination of excess activity, compatibility between activity accretion rate in target tissues and radionuclide half-life, long residence time in target tissue, and absence of uptake in normal tissues are the expected characteristics of a good therapeutic candidate. These properties may be assessed by biodistribution studies, conducted at selected time intervals after activity administration, that allow the evaluation of contrast ratios for imaging and dosimetry for therapy.

In animal models, biodistribution studies may be conducted either by sacrificing animals at selected time intervals, tissue dissection, and counting. Such studies involve relatively large numbers of animals to achieve statistical significance. They may also be conducted by imaging. This second approach is becoming more and more popular although it involves expensive instrumentation (PET and SPECT cameras for small animals), and image quantification in small animals remains a challenge. Imaging brings at least two strong advantages: the need for smaller numbers of animals through the possibility of longitudinal studies.

Figure 17.5 and Figure 17.6 show the results of a biodistribution study of the

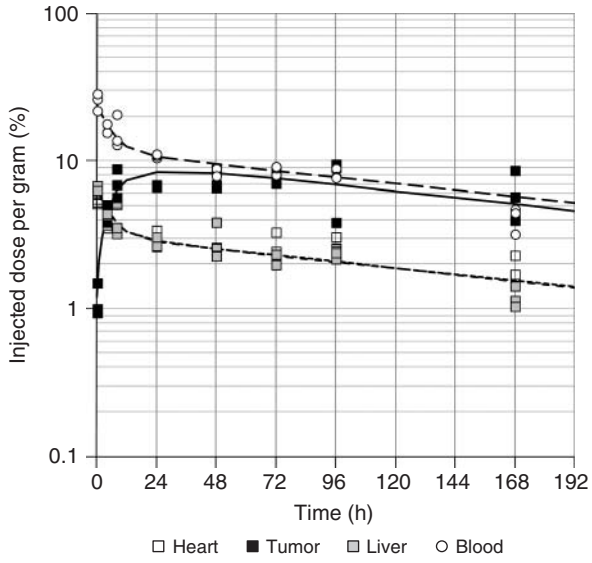


Figure 17.5 Biodistribution study of ^{125}I-labeled anti-CD138 antibody (B-B4) in nude mice xenografted with the triple-negative MDA-MB-468 breast cancer cell line. A multi-compartment pharmacokinetic analysis shows that liver and heart, as most organs, shows activity kinetics parallel to that of blood, indicating an absence of retention. By contrast, the activity builds up in the tumor over about two days then decays slowly. (Adapted from [25].)

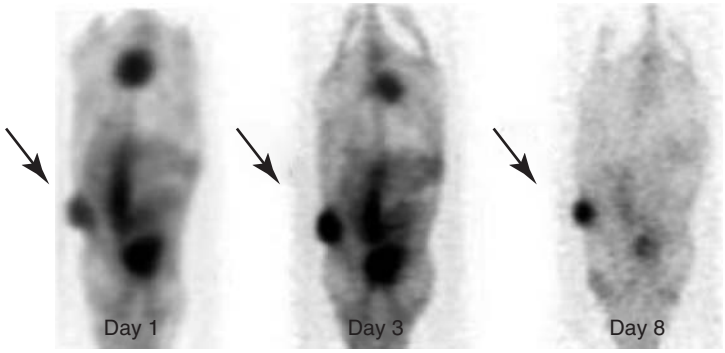


Figure 17.6 Serial PET imaging of a mouse harboring an MDA-MB-468 breast cancer xenograft after injection of ^{124}I-labeled anti-CD138 antibody (B-B4). Activity build up in the tumor between days 1 and 3 is clearly visible; while activity decays in all tissues from days 3 to 8, improved tumor to non-tumor tissue contrast is seen. (Adapted from [25].)

anti-CD138 monoclonal antibody labeled with ^{125}I (as a substitute for ^{131}I that may be used for RIT) and PET images recorded in the same model with the same antibody labeled with ^{124}I [25]. These studies allowed us to calculate absorbed doses in tumors and tissues of interest and to show that specific targeting may be achieved and irradiation of tumors could result in tumor regression, which was confirmed in a preclinical therapeutic experiment.

Preclinical therapy studies use the same animal models. First the maximum tolerated dose is determined. Then the radiopharmaceutical is injected after tumor inoculation at doses that do not exceed the maximum tolerated dose. Usually tumor growth is monitored in

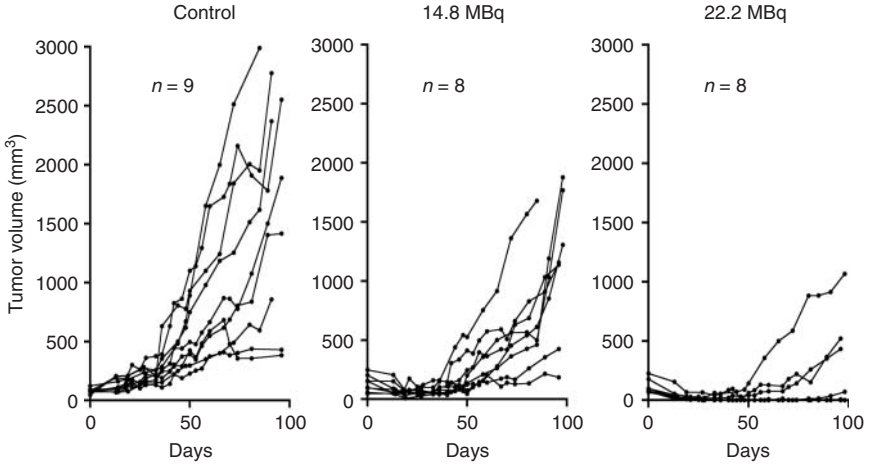


Figure 17.7 RIT of human breast cancer (MDA-MB-468) xenografts in nude mice with ^{131}I-labeled anti-CD138 antibody. Tumor cells (5×10^6) were grafted subcutaneously in the mouse flank. Mice were treated 24 days later when the tumor volume was 108 ± 55 mm^3. The maximum tolerated dose was found to be 22.2 MBq. Animals treated at the maximum tolerated dose and with a lower activity (14.8 MBq) clearly showed reduced tumor progression as compared to untreated controls. (Adapted from [25].)

groups of treated animals and in control groups that receive a sham treatment. This tumor growth may be monitored by various techniques, from serial caliper measurement of their size (sub-cutaneous tumors), to bioluminescence imaging when tumors are made to express firefly luciferase, or by MRI. When the treatment is efficacious, tumors shrink and may become undetectable for long periods of time (Figure 17.7). Follow-up must span over several months because recurrence is often observed. Toxicity is also monitored by weighing the animals and, when possible, by monitoring blood cell counts or specific biomarkers.

When preclinical testing is conducted according to well-designed protocols and provides a positive assessment of efficacy, chances of success in clinical trials are very much increased and risks are minimized for patients. For radiopharmaceuticals, these preclinical studies involve comprehensive studies of the distribution of radioactivity in the animal bodies. This kind of information cannot be obtained without animal experiments. However, differences between animals and human may always result in toxicity or a lack of efficacy.

17.5 Pre-clinical and Clinical Dosimetry

An important concern associated to the use of radionuclides is the amount of radiation delivered to the different organs of patients. In this context, dosimetry is an essential tool that aims at quantifying the amount of radiation absorbed in volumes of tissues.

17.5.1 Internal Dosimetry Methods

Methods to calculate absorbed doses in the context of *in vivo* administrations of

non-sealed radionuclides are referred as *internal dosimetry methods*. These methods follow a general scheme defined by the MIRD (medical internal radiation dose) committee of the Society of Nuclear Medicine [26]. It states that the absorbed dose within a target region is equal to the total number of nuclear transformation in a source region multiplied by the dose delivered to the target region per nuclear transformation within the source region. If numerous sources are in the vicinity of the target, the total absorbed dose to the target is equal to the sum of the different sources contributions. One of the strengths of this schema is that it can be applied to very variable scales: from the organ level to the sub-cellular level.

17.5.1.1 General Concept

The *absorbed dose* $D(r_T)$ is defined as the mean energy imparted to target region r_T per unit region mass [27]. The time-dependent rate at which the absorbed dose is delivered $\dot{D}(r_T, t)$ to target region r_T from radioactive material distributed uniformly within a source region r_S at time t after administration is given by [28]

$$\dot{D}(r_T, t) = \sum_{r_S} A(r_S, t) S(r_T \leftarrow r_S, t) \tag{17.5}$$

where $A(r_S,t)$ is the time-dependant activity of the radiopharmaceutical in source region r_S and $S(r_T \leftarrow r_S, t)$ is the radionuclide-specific quantity representing the mean absorbed dose rate to target region r_T at time t after administration per unit activity present in source region r_S.

17.5.1.2 Quantification of Activity

Activity quantification aims at measuring $\widetilde{A}(r_S, T_D)$, the time-integrated activity in source regions. Traditionally, measurements of activity levels within organs/tumors are carried out at different times after the injection of the radiopharmaceutical. From these measurements, the time–activity curve is derived, from which the area under curve gives the time-integrated activity.

For *in vivo* pre-clinical studies, the traditional technique of activity quantification consists in the sacrifice of the animal and the count of activity within the different organs/tumors, using a gamma-counter. This method provides rather accurate results, despite the issue of the unknown volume of blood content in the different tissues samples. Autoradiography is also used that consists in the production of an image resulting from the exposure of a photographic film or a scintillator associated to a CCD camera to sections of the organs/tumors. Activity is derived from the optical density on the film, or from the pixel value on an image. Both methods are invasive methods and imply the sacrifice of different groups of animals at various times after injection to derive the time-activity curves in the organs or tumors of interest.

A non-invasive method that is used both in pre-clinical and clinical studies is based on emission imaging devices such as SPECT and PET machines. Although the activity quantification is mostly straightforward, many biases impair the relation between the pixel value on the images (i.e., counts of photons) and the concentration of activity in the tissue of interest. Those biases can be classified in three groups according to their origin:

1. **Biases connected to patients**
   The movements of the whole-body or the internal movement of organs due to respiration are important issues for imaging. The time of acquisition is large compared to organ or patient

movements; this results in the apparition of blur in the image. Different motion correction algorithms are applied on respiratory-gated SPECT or PET to correct for this effect [29].

2. **Biases due to physics phenomena**
   Two main phenomena can impair the image quality and limit the accuracy of activity quantification: attenuation and scattering.
   Attenuation represents the decrease in the fluence of photons, while they travel through tissues. Underestimation of the activity occurs as a function of the depth of emission. Attenuation correction is based on the use of an attenuation map, that is, the spatial distribution of attenuation coefficients that can be obtained using a CT image of the patient in hybrid systems or by the use of rotating rod sources.
   Rayleigh and Compton scattering in the tissues changes the direction of photons. For the Compton scattering, there is also a decrease in their energy. Scattering can be important (up to 50% [30]), and it induces a degradation of contrast and spatial resolution on the images. Two types of scattering corrections are used: the first type aimed at eliminating scattered photons in each acquired projection, the second type aimed at repositioning the scattered photons by using iterative reconstruction methods and by modeling the response function of scattered events.
3. **Limitations of the detectors and signal processing**
   The last source of biases for quantification in SPECT and PET imaging is connected to the technological limits of the imaging device and the detectors.
   In SPECT imaging, the collimator, which aims at reducing the quantity of detected scattered photons, is a source of degradation for the spatial resolution.
   The dead time of the detectors used may induce a loss of counts when activities in the patient are important, leading to a highly inaccurate activity quantification.
   A great concern in imaging is the partial volume effect. This term refers to two distinct phenomena: a blurring induced by the finite spatial resolution of the detectors and the fact that the image of the activity distribution is sampled on a voxel grid. The result of these effects is that intensity values in images are not representative of the activity distribution in tissues. Multiple approaches to correct for this phenomenon are currently under development (see [31] for a review). Owing to the lack of commercially available robust correction methods in emission imaging, the activity quantification on clinical SPECT and PET images remains in some cases relatively inaccurate. It is a critical point to improve in order to achieve reliable dosimetry for patients. However, some research groups, which have developed their own correction methods, have locally reached a good accuracy, with, for instance, an error close to 10% for SPECT imaging [32].

17.5.1.3 Calculations of S Values

The quantity $S(r_T \leftarrow r_S)$ is specific to the radionuclide, the tissue composition and the spatial relationship of r_S and r_T, and finally the composition of regions located in between the sources and the target. S values are generally calculated using one of the three following techniques: analytic,

point-kernel (or its variant voxel-kernel), or Monte Carlo methods.

Analytic methods are based on the use of geometric relations and simple physics hypotheses, generally assuming that emitted particles travel following a straight line and deposit their energy continuously along their range [33]. Energy deposits are computed using empirical functions or published tables of stopping powers. Nowadays, they are mostly used for calculations at the cellular scale [34].

The dose point-kernel approach consists in the calculation of the radial distribution of absorbed dose around an isotropic point source of radiation in an infinite homogeneous medium (typically water). Convolution of the geometry of the decays within the source with the dose point-kernel function provides the distribution of absorbed dose in the target region.

The Monte Carlo method is a numerical solution to a macroscopic system through simulation of its microscopic interactions. A solution is determined by random sampling of the relationships, or the microscopic interactions, until the result converges. The Monte Carlo calculations represent actually the most refined method to calculate S values. Any geometry can be taken into account, and the physics processes used to determine the range and the energy deposits of the particles are the most accurate of the three described methods. The only drawback of this method is that it is much more time-consuming than analytic or point-kernel methods. Numerous Monte Carlo codes are available for dosimetry.

Source and target regions defined in the MIRD schema can represent whole organs, tissue subregions, voxelized structures, or cellular compartments. S values calculated for numerous radionuclides are available for standardized geometries through published tables or commercial software. Anthropomorphic phantoms representing human beings of different ages and genders were used to derive S values for organs [35]. Those S values were calculated using a Monte Carlo code, and are now available in a wide-spread software: OLINDA/EXM. For organs, which possess well-defined structural subunits and for which non-uniform activity uptake is frequently observed, the information of the mean dose delivered to the whole organ is less relevant to assess toxicity. S values were then derived for suborgan regions. At the cellular scale, tables of S values are available for monoenergetic electrons and α-particles and for β- and α-emitters [36]. Source and target regions are the cell nucleus, cytoplasm, cell surface, or whole cell. Calculations were carried out for variable radii and cell sizes. For pre-clinical studies, tables of S values are available for different models of mice and rats. Lately, with the development of new radiopharmaceuticals using short-range particle emitters, numerous ''small-scale'' dosimetry models were developed for normal organs. Those models provide S values at the cellular scale.

17.5.2 Patient-Specific Dosimetry

Therapeutic treatments in nuclear medicine are administered with a similar activity for all patients or with activities proportional to the weight or to the body surface of patients. Numerous studies demonstrated great variability among patients in their normal tissue and tumor uptake, and rates at which the activity is leaving those tissues. Intuitively, this leads to the fact that the majority of patients are undertreated, that is, receive a lower than optimal activity, because the nominal activity to administer is derived from the

patient who first overcomes toxicity in the phase I of a clinical trial. In this context, different studies or reviews have concluded for the need for patient-specific dosimetry in nuclear medicine therapy [37]. Patient-specific dosimetry refers to the acquisition of spatial and temporal distributions of activity for each patient and dosimetry calculations based on the patient's own anatomy. In detail, it includes different steps of image processing and dosimetry calculations. Individual patient dosimetry has two major interests:

- a radiobiological interest
  - to increase the knowledge of clinical radionuclide radiobiology, in part with the aim of developing new approaches and regimens.
- a clinical interest
  - In clinical trials:
    - to establish individual maximum tolerated and minimum effective absorbed doses in the process of phase I and phase II of clinical trials
    - to objectively compare the dose–response results of different radionuclide therapies, either between different patients or between different radiopharmaceuticals, as well as to perform comparisons with the results routinely obtained with external radiotherapy.
  - In daily routine to establish a dose–response relation to predict tumor response and normal organ toxicity on the basis of pre-therapy dosimetry (i.e., injection of a tracer activity and adjustment of the injected activity based on the observed biodistribution).

Although the rationale for individual patient dosimetry is well-known, there is no consensus within the clinical community on the interest of the pre-treatment dosimetry in the clinical routine. Different arguments suggested that calculations involved are difficult and expensive, that the quantification and dosimetry process are not standardized among the institutions, and finally that all this complexity prevented institutions from getting a clear dose–tissue response [37]. This last point maybe the most important: dosimetry can be useful for patients' care only if the calculated absorbed dose is related to biological effects [38].

17.5.3 Dose–Effect Relationship – Radiobiology

Whereas dose–effect relationships are relatively well defined and daily used for external radiotherapy – where the dose delivered to the tumor volume is controlled and well characterized – they are less clear in nuclear medicine. Different factors could explain those difficulties in nuclear medicine, and the numerous attempts to establish a dose–effect relationship whether they were successful or unfruitful say much about those different reasons. The first important element in order to demonstrate a dose–effect relationship is to calculate *accurate* absorbed doses. Lately, efforts to improve activity quantification protocols and the development of complex and individualized Monte Carlo dosimetry calculations have proved to be efficient to better understand dose–effect relationships in radionuclide therapy.

A second point is to define new dosimetric parameters that take into account radiobiological parameters such as dose rate, radionuclide decay, and tumor cell repair time.

Finally, there are still unsolved issues that can impair the demonstration of dose–effect relationship: the choice of a proper biological end-point and the requirement for reliable quantitative parameters to measure the treatment response and maybe more important, in the case of treatment toxicity, the influence of prior treatments on biological response [39].

17.6 Clinical Validation of Radiopharmaceuticals

17.6.1 PET Imaging

Even if aforementioned innovative positron-emitting radionuclides are not used in clinical routine practice yet, a lot of arguments in the literature suggest their major interest, especially in oncology. Functional imaging with SPECT or PET is used to characterize biologic processes at the molecular level. Most of PET imaging is performed using ^{18}FDG, which has a proven value for the detection of metastatic and relapsing disease and early therapy response evaluation. However, ^{18}FDG is not a specific tumor tracer. Consequently, new tracers are needed to specifically characterize tumor processes in the whole body, and predict tumor behavior and response to treatment. As previously mentioned, several biological markers are currently evaluated for such a specific approach, mainly peptides and antibodies. Several innovative radionuclides have been evaluated in patients: ^{68}Ga for labeling of small molecules such as peptides or antibody fragments, ^{89}Zr and ^{124}I (long-lived positron-emitting radionuclides) for labeling of large immunoglobulin molecules in immuno-PET, and ^{64}Cu for labeling of a broader range of molecular sizes.

The recent introduction of PET/CT using ^{68}Ga-DOTA-peptide (DOTA: 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid) for the evaluation of NETs has significantly improved the diagnostic work-up, previously based only on conventional imaging modalities such as ultrasound, CT, MRI, and somatostatin receptor (SSR) somatostatin receptor scintigraphy (SRS) with the same peptide labeled with ^{111}In. More recently, different ^{68}Ga-DOTA-peptides (DOTATOC, DOTANOC, and DOTATATE) specifically binding to SSR overexpressed on the surface of NET cells have been evaluated in the clinical setting for either NET diagnosis or PRRT. The major difference among these compounds relies on a slightly different affinity to SSR subtypes. ^{68}Ga-DOTANOC, compared with similar diagnostic compounds, was also reported to present a favorable dosimetry. PET/CT with ^{68}Ga-DOTA-peptides showed a higher sensitivity for the detection of well-differentiated NET than other imaging procedures (particularly CT and SRS) [40]. Thus, the clinical role of PET ^{68}Ga-peptide in the management of NET has been clearly validated.

PET using other carrier molecules labeled with ^{68}Ga has also been evaluated in various clinical settings with promising results. The human epidermal growth factor receptor 2 (HER2) is overexpressed in several cancer types characterized by aggressive tumors and decreased survival. Determining the presence of HER2 in potentially responsive patients is increasingly important given the growing arsenal

of HER2-targeted therapies [41]. Diagnosis of HER2 overexpression is recommended for newly diagnosed breast carcinomas, but discordance in HER2 expression between the primary tumor and metastases complicates diagnosis of recurrent disease. Thus, a molecular imaging with an agent showing a high specificity and affinity for HER2 might be highly beneficial for improved prediction of therapy efficacy and monitoring of therapy response.

Results of several clinical studies on the potential interest of immuno-PET using ^{89}Zr or ^{124}I have been reported in the literature. Very promising results have been obtained using the chimeric antibody cG250 directed to carbonic anhydrase-IX antigen (CAIX), which is over-expressed in clear-cell renal carcinomas [6]. Divgi *et al.* assessed whether ^{124}I-cG250 PET could help distinguish clear-cell renal carcinoma, the most common and aggressive renal tumor and a benign tumor in the situation of incidental discovery of a renal mass. The authors concluded that ^{124}I-cG250 PET can accurately identify clear-cell renal carcinoma that requires a surgical resection.

^{64}Cu is a promising radionuclide for hypoxia imaging using ^{64}Cu-diacetyl-bis(N4-methylthiosemicarbazone) (^{64}Cu-ATSM) [42]. Hypoxia has been recognized as a major obstacle to effective radiotherapy and chemotherapy in different types of tumor. Copper-ATSM PET has been shown to distinguish patients likely and unlikely to respond to conventional therapies for cancers of the lung, uterine cervix, and head and neck cancer.

17.6.2 Molecular Therapy

For the past 20 years, molecular radiotherapy tremendously developed using antibody- and peptide-radionuclide conjugates for RIT and PRRT. This development gave rise to a new concept of Theranostics, which consists, for a cancer patient, in first receiving a diagnostic activity of an antibody or a peptide labeled with a radionuclide appropriate for SPECT or preferably PET imaging. If good antibody or peptide localization is observed at the tumor sites, a therapeutic activity of the same antibody or peptide labeled with a therapeutic (β^- or α emitting) radionuclide is injected aiming at inducing curative effects.

17.6.2.1 Radioimmunotherapy (RIT)

The first RIT clinical studies started in the 1980s and failed to demonstrate an efficacy. This failure was associated with several problems, including the large tumor burden of patients. It has been later shown that there is an inverse relationship between tumor size and tumor uptake. So a very low tumor uptake was observed in large tumors resulting in a very low tumor absorbed dose insufficient to have a response. Another problem was the murine origin of monoclonal antibodies, which prevented repeated administration to patients. This problem has been subsequently overcome by the advent of humanized and fully human antibodies.

The first clinical efficacy of RIT was observed in the 1990s when an anti-CD20 antibody was labeled with ^{131}I (^{131}I-tositumomab, Bexxar®) and evaluated in patients with non-Hodgkin's lymphoma. This type of cancer is inherently sensitive to radiation, and then a relatively low tumor dose can induce a clinical response. It has been clearly demonstrated that ^{131}I-tositumomab is an effective treatment for relapsed or refractory follicular lymphoma with a toxicity profile similar to those of second-line chemotherapy regimens. It has been approved by

the FDA in the USA in 2003. Another anti-CD20 antibody has been labeled with ^{90}Y-ibritumomab tiuxetan or Zevalin®. It was approved in March 2002 for the treatment of adults with relapsed or refractory low-grade, follicular B-cell lymphoma, and in September 2009 for the treatment of patients with previously untreated follicular non-Hodgkin's lymphoma who achieve a partial or complete response to first-line chemotherapy [43].

Following this demonstration of efficacy in radiosensitive tumors, the same efficacy is expected in the more frequent situation of solid tumors that are more resistant to radiation. A first demonstration of such efficacy was published in 2005 in the situation of consolidation after salvage resection of liver metastases of colorectal carcinoma. Indeed, after surgical resection, the surgeon can consider that he resected all what was visible but cannot exclude the persistence of invisible residual microscopic metastases. It is highly probable that this first demonstration will be followed by other ones in different types of cancer.

For the future, efficacy of RIT could still be enhanced by the use of α-particle-emitting radionuclides such as ^{213}Bi or ^{211}At. The LET of α particles is much higher than that of β particles, and the irradiated tumor cells have a limited capacity to repair DNA damages. Moreover, cell death may occur after a single decay traversing the nucleus. Consequently, α-immunotherapy has the potential great interest to have a very efficient killing effect of isolated tumor cells while limiting damage to surrounding normal tissues. A first proof-of-concept clinical study for systemic targeted α particle immunotherapy was conducted at the beginning of the 2000s in patients with refractory acute myelogenous leukemia and demonstrated the safety, feasibility, and some therapeutic effects [44]. Another proof-of-concept clinical study for infusion in surgically created resection cavity with an ^{211}At-labeled antibody was conducted in patients with recurrent brain tumor and also demonstrated the safety, feasibility, and some promising antitumor benefit [45].

17.6.2.2 Peptide Receptor Radionuclide Therapy (PRRT)

Molecular therapy using radiolabeled peptides has become an important topic both in Europe and the USA [46] owing to its clinical efficacy documented in a number of studies. Thousands of patients have been treated successfully with this approach, using mainly somatostatin analogs labeled with ^{90}Y or ^{177}Lu. Quite interestingly, tumor response rates and progression-free survival compare favorably to the limited number of alternative treatment modalities. A survival gain of several years from time of diagnosis has been observed. Clinical studies with other peptide analogs are ongoing in several cancer types.

Glossary

Bq (becquerel): SI unit of activity. 1 Bq = 1 decay per second.

Brachytherapy: Introduction of closed sources that are emitting (short-range) ionizing radiation into the body or close to the body. Formally a discipline closer to radiology than to nuclear medicine.

Ci (curie): A unit of activity that is still used in nuclear medicine. 1 Ci = 37 GBq, 1 mCi = 37 MBq.

CT: Computed tomography. Several X-ray transmission scans are taken at different angles rotated around a single axis and numerically combined to a 3D

view. This provides a morphological image of the body.

Lesion: Any abnormality in the tissue of an organism.

MRI: Nuclear magnetic resonance imaging provides a morphological image with better soft tissue contrast than CT and no radiation burden. However, it cannot be used directly for attenuation correction of PET or SPECT.

Multi-modality imaging: Combination of different imaging modalities such as PET/CT, PET/MR, SPECT/CT, and ultrasound.

Nuclear medicine: Medical discipline using open radiation sources in the human body for diagnostic or therapeutic applications.

PET: In positron emission tomography, the emitted positrons annihilate, emitting two 511 keV γ-rays emitted back to back. Coincident detection provides the line of response. Detection of many such events allows computing the 3D distribution of the positron emitters, that is, functional imaging.

PET/CT and SPECT/CT: Combination of PET and CT or SPECT and CT into a single device for improved combination of functional and morphological images. The CT scan helps also for the attenuation correction.

PRRT: Peptide receptor radionuclide therapy uses therapeutic radionuclides linked to peptides that target specific peptide receptors that are overexpressed on the cancer cells.

RECIST: Response evaluation criteria in solid tumors is a set of published rules that define when cancer patients improve (''complete or partial response''), stay the same (''stable disease''), or worsen (''progressive disease'') during treatments.

RIT: Radioimmunotherapy uses therapeutic radionuclides linked to antibodies that target specific antigens that are overexpressed on the cancer cells.

SCID mice: Mice with severe combined immunodeficiency have an immune system that cannot efficiently fight infections or reject tumors and transplants. They are well suited as model systems for the study of disease in mammals.

Scintigraphy: It is a functional imaging technique providing 2D planar images of a γ-emitter distributed in the body registered with a gamma camera. Position resolution is achieved via suitable collimators.

SPECT: Single photon emission computed tomography. 2D scintigraphic images from different directions are combined to compute a 3D image for functional imaging with improved spatial resolution.

Theranostics: A modality of personalized medicine where individual patients are first tested with a diagnostic procedure (e.g., SPECT or PET) to judge the probable response to a therapy before the latter is applied. This assures that only probable responders are treated, thus improving the ratio of overall success versus side-effects.

References

1. Sörensen, J. (2012) How does the patient benefit from clinical PET? *Theranostics*, **2**, 427.
2. Even-Sapir, E. *et al.* (2006) The detection of bone metastases in patients with high-risk prostate cancer: 99mTc-MDP planar bone scintigraphy, single- and multi-field-of-view SPECT, 18F-fluoride PET, and 18F-fluoride PET/CT. *J. Nucl. Med.*, **47**, 287.
3. Gonzalez-Angulo, A.M., Hennessy, B.T., and Mills, G.B. (2010) Future of personalized medicine in oncology: a systems biology approach. *J. Clin. Oncol.*, **28**, 2777.

4. Schwaiger, M. and Wester, H.J. (2011) How many PET tracers do we need? *J. Nucl. Med.*, **52** (Suppl. 2), 36S.
5. Pfeifer, A. *et al.* (2012) Clinical PET of neuroendocrine tumors using 64Cu-DOTATATE: first-in-humans study. *J. Nucl. Med.*, **53**, 1207.
6. Divgi, C.R. *et al.* (2007) Preoperative characterisation of clear-cell renal carcinoma using iodine-124-labelled antibody chimeric G250 (124I-cG250) and PET in patients with renal masses: a phase I trial. *Lancet Oncol.*, **8**, 304.
7. Kayani, I. *et al.* (2008) Functional imaging of neuroendocrine tumors with combined PET/CT using 68Ga-DOTATATE (DOTA-DPhe1,Tyr3-octreotate) and 18F-FDG. *Cancer*, **112**, 2447.
8. Barbet, J. and Chatal, J.F. (2011) The best radionuclide for radioimmunotherapy of small tumors: beta- or alpha-emitter? *Eur. J. Nucl. Med. Mol. Imaging*, **38**, 271.
9. Nuclear Energy Agency (2011) The Supply of Medical Radioisotopes: The Path to Reliability. Report no. 6985, NEA.
10. Kassis, A.I. (2011) Molecular and cellular radiobiological effects of Auger emitting radionuclides. *Radiat. Prot. Dosimetry*, **143**, 241.
11. Lewington, V.J. (2005) Bone-seeking radionuclides for therapy. *J. Nucl. Med.*, **46**, 38S.
12. Vera Ruiz, H. (1988) Report of an IAEA Consultants' Meeting on fluorine 18: reactor production and utilization. *Appl. Radiat. Isot.*, **39**, 31.
13. IAEA (2009) Cyclotron Produced Radionuclides: Principles and Practice. Technical Report Series 465, IAEA.
14. Horlock, P.L. *et al.* (1981) The preparation of a rubidium-82 radionuclide generator. *J. Radioanal. Chem.*, **64**, 257.
15. Abbas, K., Buono, S., and Burgio, N. (2009) Development of an accelerator driven neutron activator for medical radioisotope production. *Nucl. Instrum. Methods A*, **601**, 223.
16. Haddad, F. *et al.* (2008) ARRONAX, a high-energy and high-intensity cyclotron for nuclear medicine. *Eur. J. Nucl. Med. Mol. Imaging*, **35**, 1377.
17. Helmeke, H.J., Mahnke, E., Schaardt, U., and Knapp, W.H. (2004) External targets for the production of 211At–review and status of the target development at the Hannover cyclotron. *Z. Med. Phys.*, **14**, 195.
18. Köster, U. for the ISOLDE Collaboration (2001) ISOLDE target and ion source chemistry. *Radiochim. Acta*, **89**, 749.
19. Müller, C. *et al.* (2012) A unique matched quadruplet of terbium radioisotopes for PET and SPECT and for α- and β^--radionuclide therapy: an in vivo proof-of-concept study with a new receptor-targeted folate derivative. *J. Nucl. Med.*, **53**, 1952.
20. IAEA (2009) Cyclotron Produced Radionuclides: Physical Characteristics and Production Method, Technical Report Series 468, IAEA.
21. IAEA (2009) Cyclotron Produced Radionuclides: Guidelines for Setting up a Facility, Technical Report Series 471, IAEA.
22. Voskoglou-Nomikos, T., Pater, J.L., and Seymour, L. (2003) Clinical predictive value of the *in vitro* cell line, human xenograft, and mouse allograft preclinical cancer models. *Clin. Cancer Res.*, **9**, 4227.
23. Baldrick, P. (2008) Safety evaluation to support first-in-man investigations II: toxicology studies. *Regul. Toxicol. Pharmacol.*, **51**, 237.
24. Dalvie, D. (2000) Recent advances in the applications of radioisotopes in drug metabolism, toxicology and pharmacokinetics. *Curr. Pharm. Des.*, **6**, 1009.
25. Rousseau, C. *et al.* (2011) Syndecan-1 antigen, a promising new target for triple-negative breast cancer immuno-PET and radioimmunotherapy. A preclinical study on MDA-MB-468 xenograft tumors. *EJNMMI Res.*, **1**, 20.
26. Loevinger, R., Budinger, T.F., and Watson, E.E. (1991) *MIRD Primer for Absorbed Dose Calculations*, Society of Nuclear Medicine, Reston, VA.
27. ICRU (2011) Fundamental Quantities and Units for Ionizing Radiation. Report 85, ICRU.
28. Bolch, W.E. *et al.* (1999) MIRD pamphlet No. 17: the dosimetry of nonuniform activity distributions – radionuclide S values at the voxel level. Medical Internal Radiation Dose Committee. *J. Nucl. Med.*, **40**, 11S.
29. Nehmeh, S.A. and Erdi, Y.E. (2008) Respiratory motion in positron emission tomography/computed tomography: a review. *Semin. Nucl. Med.*, **38**, 167.

30. Buvat, I. (2007) Quantification in emission tomography: challenges, solutions, and performance. *Nucl. Instrum. Methods A*, **571**, 10.
31. Erlandsson, K. *et al.* (2012) A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys. Med. Biol.*, **57**, R119.
32. Dewaraja, Y.K. *et al.* (2012) MIRD pamphlet No. 23: quantitative SPECT for patient-specific 3-dimensional dosimetry in internal radionuclide therapy. *J. Nucl. Med.*, **53**, 1310.
33. Goddu, S.M. *et al.* (1994) Cellular dosimetry: absorbed fractions for monoenergetic electron and alpha particle sources and S-values for radionuclides uniformly distributed in different cell compartments. *J. Nucl. Med.*, **35**, 303.
34. Chouin, N. *et al.* (2009) Evidence of extranuclear cell sensitivity to alpha-particle radiation using a microdosimetric model. I. Presentation and validation of a microdosimetric model. *Radiat. Res.*, **171**, 657.
35. Cristy, M. and Eckerman, K. (1987) Specific Absorbed Fractions of Energy at Various Ages from Internal Photon Sources. I. Methods. Report ORNL/TM-8381/V1, Oak Ridge National Laboratory.
36. Howell, R.W. *et al.* (1997) *MIRD Cellular S Values: Self-absorbed Dose Per Unit Cumulated Activity for Selected Radionuclides and Monoenergetic Electron and Alpha Particle Emitters Incorporated into Different Cell Compartments*, Society of Nuclear Medicine, Reston, VA.
37. Stabin, M.G. (2008) Update: the case for patient-specific dosimetry in radionuclide therapy. *Cancer Biother. Radiopharm.*, **23**, 273.
38. Sgouros, G. (2007) Toward patient-friendly cell-level dosimetry. *J. Nucl. Med.*, **48**, 496.
39. Sgouros, G. and Hobbs, R.F. (2012) Patient-specific dosimetry, radiobiology, and the previously-treated patient, in *Therapeutic Nuclear Medicine* (ed. R.P. Baum), Springer.
40. Gabriel, M. *et al.* (2007) 68Ga-DOTA-Tyr3-octreotide PET in neuroendocrine tumors: comparison with somatostatin receptor scintigraphy and CT. *J. Nucl. Med.*, **48**, 508.
41. Murphy, C.G. and Modi, S. (2009) HER2 breast cancer therapies: a review. *Biologics*, **3**, 289.
42. Bourgeois, M. *et al.* (2011) Contribution of [64Cu]-ATSM PET in molecular imaging of tumour hypoxia compared to classical [18F]-MISO-a selected review. *Nucl. Med. Rev. Cent. East. Eur.*, **14**, 90.
43. Morschhauser, F. *et al.* (2013) ^{90}Yttrium-Ibritumomab Tiuxetan Consolidation of First Remission in Advanced-Stage Follicular Non-Hodgkin Lymphoma: Updated Results After a Median Follow-Up of 7.3 Years From the International, Randomized, Phase III First-Line Indolent Trial. *J. Clin. Oncol.*, **31**, 1977.
44. Jurcic, J.G. *et al.* (2002) Targeted α particle immunotherapy for myeloid leukemia. *Blood*, **100**, 1233.
45. Zalutsky, M.R. *et al.* (2008) Clinical experience with α-particle-emitting 211At: treatment of recurrent brain tumor patients with 211At-labeled chimeric antitenascin monoclonal antibody 81C6. *J. Nucl. Med.*, **49**, 30.
46. Ambrosini, V. *et al.* (2011) Radiopeptide imaging and therapy in Europe. *J. Nucl. Med.*, **52**, 42s.

18
Cancer Therapy with Ion Beams

Gerhard Kraft

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

18.1 Introduction

Shortly after the discovery of X-rays in 1896 by W.C. Roentgen, radiotherapy developed very fast, and it became evident that any tissue could be destroyed using radiation doses high enough. But X-rays do not differentiate between normal and malignant tissue to an extent that would yield the tumor inactive with moderate side effects. Therefore, the dose had to be focused on the planned target volume and normal tissue spared. To reach this goal, cross-fire techniques are used. Until the Second World War, the classical X-ray tubes operated with a few hundred volts were the main radiation sources for external radiotherapy and produced dose distribution with a steep decay in the depth where the initial X-ray intensity was reduced to one half after 3–5 cm. This dose distribution was improved first by using million-electronvolt γ-rays from Co γ-sources and later on bremsstrahlung produced by electron linacs in an energy range between 2 and 20 MeV electron energy.

For high-energy X-rays, the dose maximum is shifted a few centimeters below the skin, and the exponential decay in depth becomes less steep, allowing greater tumor doses. Although sophisticated techniques of cross-fire irradiation combined with intensity modulation such as the intensity-modulated radiation therapy (IMRT) are applied, the tolerance of the normal tissue is the limiting factor, and in many cases, the dose is too low to eradicate the tumor completely.

In 1946, Robert Wilson measured the thickness of the shielding material necessary for protection against proton beams at the new cyclotron at Berkeley. Wilson confirmed that for ions, the energy deposition increased with penetration depth to a maximum value at the end of range as found in 1905 by W. Bragg for α particles, and Wilson realized that ion irradiation could be very beneficial for radiation therapy of deep-seated tumors. In contrast to photons, particle beams exhibit a low dose in the entrance channel but high doses at the end in the tumor. However, his first publication in the medical journal *Radiology* did not generate any resonance from the medical community, and it took nearly another decade until the first patient was treated at Berkeley with protons in 1955, followed by helium treatment in 1957 and heavier ions in 1975.

Other locations such as Uppsala (1957), Harvard (1961), and Dubna (1967) followed some time later [1].

All these treatments were performed at accelerators originally installed for nuclear research and then partially modified for

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

therapy applications. The main problem to be solved for the treatment of tumors was to adapt the primary small pencil-like beam to an extended volume of a real tumor. First, the tumor size and its location in the body had to be determined with high accuracy, which is now possible using the modern imaging procedures such as CT (computed tomography), MRI, or PET (positron emission tomography). Second, the beam has to be enlarged in three dimensions to cover the target volume, which was done in analogy to the field shaping in conventional photon therapy. It was not possible at that time to distribute the beam over the volume using magnetic deflection. But greater precision was also not requested because medical diagnostics at that time could not deliver the adequate planning data.

In these accelerators, a fast energy variation was not important for physics experiments, and therefore, the beam energy could be only varied very slowly. Consequently, energy variation and modulation for therapy were done by inserting material of various thickness and shapes in the beam path. These techniques resulted in dose distributions superior to those from the conventional therapy at these former times, but they are still used in most of the older ion-beam therapy facilities [2]. At present, IMRT and other conformal photon therapies are, in many cases, of at least of the same quality as passive beam modulations for protons.

To improve target conformity, an active beam delivery in three dimensions was introduced for carbon therapy at GSI in 1997 and for protons in two dimensions by PSI, Switzerland. In the active 3D technique, the beam is scanned over the target volume laterally by a fast magnetic scanning system, and layers of different depths are addressed by the beam extracted with different energy levels from the accelerator. Active scanning produces a target conformity that is given by the physical limits, that is, lateral and longitudinal scattering of the ions, which is, depending on penetration depth, in the order of 1–2 mm for carbon ions and three times larger for protons [3].

Using beam scanning techniques, an ultimate dose distribution can be produced with a high and conformal dose to the tumor but low dose to the normal tissue around.

For the heavier carbon ions, the high dose in the target volume is potentiated by an increase of the relative biological effectiveness (RBE) because of the large dose in the individual ion tracks that reduces the capacity for repairing radiation damage produced in the DNA. Clinical RBE values can be in the order of 2–3; but in contrast to neutrons, the RBE elevation is a differential effect: for neutrons, RBE is constant over the penetration depth, but for carbon ions, it is low in the entrance channel and increases strongly toward the end of the ion range, giving another boost of inactivation to the tumor tissue while sparing the normal tissue before the tumor.

In general, RBE is a complex function of many parameters such as particle energy and energy loss at the physical side and repair capacity and size of the cell nucleus at the biological side. In order to include the RBE variation in treatment planning, the local effect model (LEM) was introduced, which enables to calculate the biological effective dose for each location of the affected tissue (tumor) and the normal tissue [3].

The final innovation in this field was the *in vivo* control of the carbon beam using PET techniques. A minor fraction of the primary carbon-12 ions undergo nuclear reactions such as neutron stripping, which

yields the carbon isotopes 11 and 10 with significant intensity. These isotopes decay under positron emission and can be monitored from outside by their coincident decay of the 511 keV γ-quanta. Because the radioactive carbon isotopes have long lifetimes, they decay mainly when stopped and indicate the location of the Bragg maximum. The positron decay of the radioactive isotopes can be measured during and shortly after the irradiation, and the correct range distribution of the primary carbon beam can be determined [4].

In summary, the novel ion-beam therapy techniques include not only beam scanning but also *in vivo* PET verification and treatment planning including potentiation of the biological effective dose. Using all these techniques together, it was possible to lower the dose to the normal tissue or to increase the target or both. In consequence, greater tumor control rates (80% or better) with lower side effects could be reached. These results together with similar results at NIRS (National Institute for Radiological Science), Japan, stimulated the construction of medical dedicated carbon facilities. In Japan, at Hyogo and Gunma, carbon facilities went in operation; at Heidelberg and Pavia, patients are treated in combined proton/carbon units. Other facilities in Wiener Neustadt, Shanghai, and Marburg are under construction. A facility at Kiel was deconstructed for financial reasons before it went into operation. But at the running facilities, the clinical results are still extremely good.

18.2 Physical Advantages of Ion Beams for Therapy

The main advantage for all ion beams independent from atomic number is the inverse depth dose distribution compared to photon beams [3]. Because of the stochastic interaction of the photons with the penetrated matter, the deposited dose decreases exponentially with depth. For greater photon energies, the exponential decay becomes less steep, and because of the forward scattering of the created Compton electrons, a dose maximum is built up in the tissue up to a depth of 1-3 cm depending on the photon energy followed by a shallow decay (Figure 18.1).

In contrast, beams of charged particles exhibit an inverse dose profile: the deposited dose is small at high energies because of the $1/E$ dependence of the interaction cross section with the electrons according to the Bethe–Bloch formula. But with decreasing energy, the cross section, and consequently the dose deposition, increases to the so-called Bragg maximum before the charge of the ions is diminished by electron capture before the particle stops.

For a beam of many particles, the statistics of the energy loss processes causes range straggling of the individual ions, which yields an increasing width of the Bragg peak but a smaller height of maximum (Figure 18.2).

The general shape of the Bragg curves is strongly influenced by the second process, namely, the nuclear fragmentation of the projectile ions. At therapy energies, the target nucleus can undergo nuclear reactions and all the lighter isotopes between carbon and protons are produced, including lighter carbon isotopes that result from the stripping of neutrons from the stabile carbon. These isotopes are positron emitters and can be used to trace the beams *in vivo*. Because the long lifetimes, they decay when they are stopped in the carbon Bragg maximum where they can be detected via PET (see later discussion).

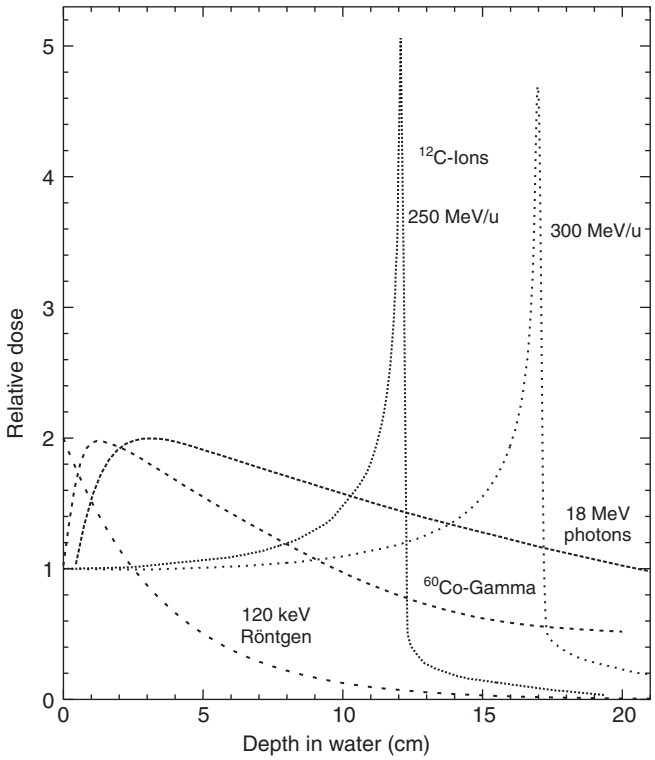


Figure 18.1 Comparison of the depth dose distribution of X-rays and particles. Photon curves of different energies show an exponential decrease with depth after an initial "buildup" effect that is caused by the forward-scattered Compton electrons. For a deep-seated tumor, the photon dose in the normal tissue before the target is larger than the tumor dose. Ions such as carbon have an inverse dose profile: small dose in the entrance and large dose in the end in the Bragg maximum. This dose maximum can be shifted over the target volume, changing the particle energy. Becasue of the inverse dose profile, the dose in the target volume can be greater than that in the normal tissue.

Because of the smaller atomic numbers, the range of the other nuclear fragments becomes larger and a part of the dose is transferred as a tail behind the Bragg maximum. This dose tail depends on the atomic number of the projectiles. The fragmentation increases with atomic number and energy. For protons, there is no tail because they cannot decay as lighter fragments. For ions heavier than carbon, such as neon or argon, the fragmentation process becomes too dominating and the advantages of inverse dose profile and greater biological effectiveness disappears for larger penetration depths. Therefore, the fragmentation process determines also the choice of the ion used for therapy, and carbon is a reasonable compromise to achieve good conformity.

Another important physical parameter for therapy that determines the overall conformity is the lateral and longitudinal scattering of the ions. In general, a steep falloff, that is, sharp gradients, is favorable because in many cases critical organs, which are close to the tumor volume, have

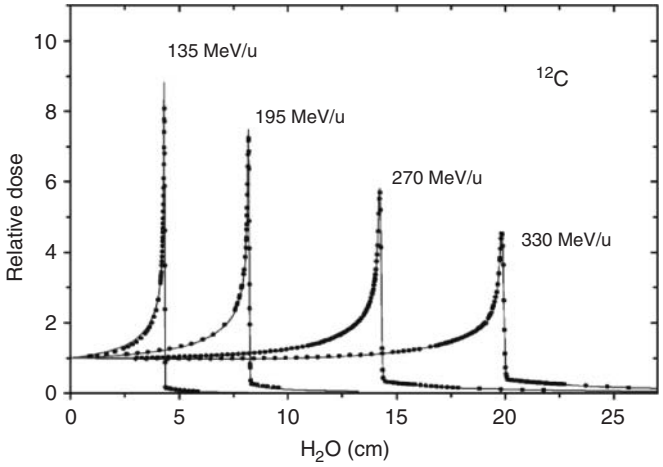


Figure 18.2 Measured Bragg curves of carbon ions in water. The Bragg maximum decreases in height and increases in half width because of the multiple scattering, and the dose tail caused by nuclear fragmentation becomes larger at greater energies and larger penetration depth of the carbon ions.

to be protected. At present, the diagnostic tools such as CT, MRI, and PET have a resolution of 1–3 mm, which can be used by the physicians for planning. Therefore, a dose gradient of 1–2 mm is desirable for the dose delivery but sharper gradients are mostly not necessary. In contrary, soft gradients are in a few times of interest, for instance, when fields are patched to a larger field. But soft gradients can also be produced by beam scanning using intensity modulation.

18.3 Passive and Active Beam Delivery Systems

In the past, most proton and heavy ion facilities used passive dose-forming techniques that are very close to the dose-shaping systems in photon therapy. Using these techniques, the primary monoenergetic and sharp "pencil" beam is widened laterally by complex scattering systems that produce a profile with a flat top (Figure 18.3). The outer contours of this profile are adjusted to the maximum tumor contours by apertures in front of the patient.

In addition, the monoenergetic beam is modulated in depth with ridge filters where the range of the beam is reduced to the most proximal part of the tumor by the large absorption in the thick part of the material in the peaks of the ridges. The most distal range originates from particles penetrating through the ridge filters at the thinnest parts. Finally, the initial energy of the ions can be reduced using appropriate absorbers or range shifters. With this technique, the primary monoenergetic pencil beam can be distributed in depth and shaped to the tumor volume. But with the passive shaping methods, the high dose volumes have a finite distal and a finite proximal border and steep walls at the side representing a cylindrical volume that covers the complete target volume. When there are sensitive structures behind the

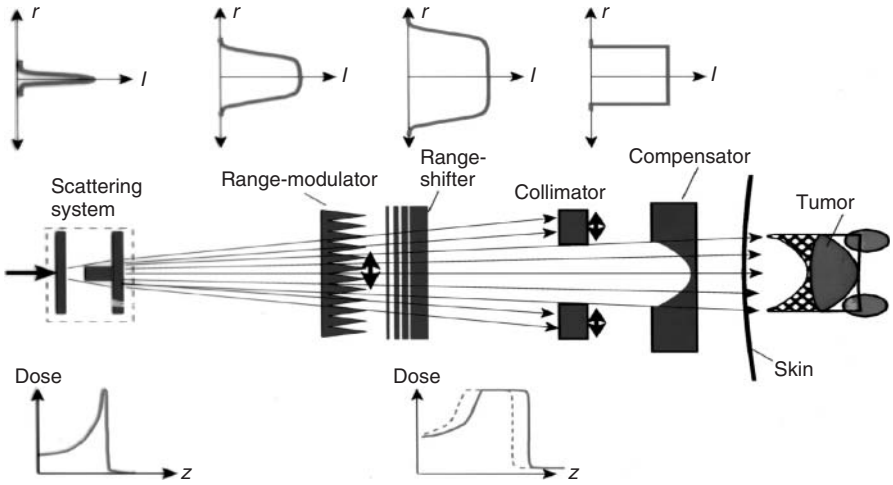


Figure 18.3 Principles of passive beam forming system. The pristine beam is laterally enlarged by a complex multistage scattering system. The outer contours of the beam spot are then shaped by apertures according to the target contours. For depth variation, ridge filters are used. Finally, critical structures at the distal side can be spared using absorbers (boli) placed before the patient.

tumor (at the distal side), local absorbers can be positioned over the patient's skin that shift the dose distribution at these spots to the proximal parts.

In general, with these passive techniques, large treatment volumes enclosing the tumor are possible, which contain not only the target volume but also a large fraction of normal tissue. But the normal tissue in the target volume is inactivated with a similar effectiveness as the tumor, which limits the maximum dose that can be given to the tumor. In addition, the passive beam shaping devices in front of the patient are strong sources of neutrons that are mostly scattered in the forward direction, that is, into the patient's body. These neutrons can induce secondary tumors after treatment and should prevent from treating young patients with these techniques. The instrumentation for passive beam shaping is given in great detail in [2].

To improve the target conformity, at GSI and PSI, active beam shaping methods have been developed that exploit the fact that ions are charged particles that can be deflected laterally using fast magnets and by changing the range by an energy variation in the accelerator from pulse to pulse [4, 5]. These active beam delivery systems are very flexible because they can follow the outer contours of the tumor in 3D and do not use patient-specific hardware for beam shaping (Figure 18.4). They also produce a very clean beam with a minimum of neutron contamination. But active beam delivery requires a greater effort in planning and beam control, both requiring large computer capacities that were not available in the beginning of particle therapy.

In the practice of active beam delivery, the target volume is divided into slices of equal particle range and each slice is covered by a net of pixels that are irradiated

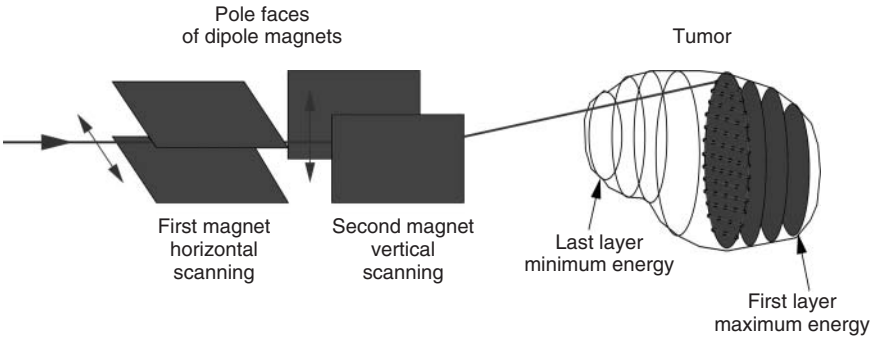


Figure 18.4 Principles of an active beam forming system. The target volume is dissected in slices of equal particle range, and a fine pencil beam is guided over each slice by fast magnetic deflection. The range is varied from pulse to pulse by an energy variation in the accelerator.

sequentially in a rasterlike pattern. For the irradiation, two different strategies have been developed under the names ''voxel scanning'' and ''raster scanning.''

In the voxel scan technique, the centers of the individual voxels were one half widths FWHM (full width at half-maximum) of the lateral beam profile apart yielding in the ideal case a flat dose distribution over the irradiated area. The voxels were irradiated separately from each other by turning off the intensity when the beam is moved from one to the next beam position along a horizontal line. The following line again was one half width separated from the first one, and the energy layers were also one half width of the Bragg peak apart. Typical beam widths were 5-10 mm, so a few thousand voxels had to be planned and irradiated independently for 1 l target volume.

With the voxel method, a homogeneous dose profile over an irregularly shaped target volume could be obtained under the condition that the beam profile would stay constant and the center of the voxels will be hit correctly; deviations from the required positions result in local hot and cold spots. The time to irradiate a volume is limited by time needed to sample the dose at each voxel plus the time needed for the turning off and turning on procedure [5].

The raster scan method started from the idea of moving the beam continuously over an isoenergy slice with varying but controlled speed according to the particle coverage needed at each spot. In then turned out that for practical reasons, the path has to be ''digitized,'' that is, divided into small parts or pixels where a constant dose or particle fluency has to be assigned for each pixel. The beam is then moved from one pixel to the next when the planned dose was administered. But during this movement, the beam is not turned off and the particle fluence in the transition time is accounted for in the next pixel. By optimization, it was found that the distance between two pixels should be not more than one-third of the FWHM. Larger distances resulted in inhomogeneities, whereas smaller distances increased the number of pixels to be planned without changing the quality. About 20–50 000 pixels are usually needed to treat a tumor depending on its size, usually a few hundred milliliters, and the beam diameter, which is again in

the order of 6–10 mm. It was found that the raster method is more robust against beam instabilities in time and in position and is faster in its application [4].

18.4 Quality Control and Patient Flow

The limiting factor for fast irradiation is quality control, that is, the online measurement of the intensity and location of the beam in the detectors in front of the patient. A combination of an ionization chamber with a multiple wire chamber gives the actual intensity and position of the beam. In the 3D delivery system at GSI, the intensity was read out every 10 μ sec and the location every 100 μ sec. At least three measurements have to be made for each pixel to decide with some statistical significance whether the irradiation was correct within the given limits.

The irradiation of a complete target volume takes a few minutes depending on the number of pixels and isoenergy layers.

For a clinical facility, the irradiation time is important because it determines the number of patients that can be treated. Usually, the dose to a tumor is given in daily fractions: 30 for conventional photon or proton treatment and 20 for carbon ions. But even shorter fractionation seems to be feasible for carbon therapy. The treatment of 1000 patients a year corresponds to 20 000 fractions or 70 fractions each day assuming 300 working days a year. This is possible only when the irradiation time for one fraction is less than 10 min, which is the critical number for the commercialization of the system, and because of the high investment for particle therapy, more than 1000 patients are needed for amortization.

Another important aspect is the application of focused treatment to moving targets, that is, tumors that move because of breathing or heartbeat. This is a problem of all conform irradiations not only for particle but also for photons. When the irradiated volume gets close to the borders of the tumor, any motion shifts the tumor at least partially out of the target field. In addition, the scanning structure will interfere with the motion frequency and produce inhomogeneities called *interplay effects*.

To compensate for motion, different strategies have been proposed and used:

1. **Synchronization (gating)**: The target volume is only irradiated in the exhale phase of the breathing cycle where the internal geometry is mostly the same. Gating gives good results but prolongs the irradiation time.
2. **Multiple painting**: The target volume is enlarged and multiply irradiated with a smaller dose. This is compensated for inhomogeneities but destroys the good conformity and prolongs the exposure time.
3. **Tracking**: The scanning is corrected for the actual motion of the target. This is not only the most promising technique but also the most challenging because it requires fast monitoring of the motion.

The best method to irradiate moving organs is to shorten the total treatment of one fraction to such a short time (<1 min) that the motion can be stopped by breath-holding techniques. Such systems are under way in NIRS and PSI.

18.5 PET Detection of Instable Isotopes is Used for Quality Control

The nuclear fragmentation and production of instable isotopes results in a dose tail

beyond the Bragg maximum, which is not wanted. But the nuclear fragmentation has also a positive effect when the resulting isotopes are positron (β^+) emitters. In case of a carbon beam, there are two possibilities of producing β^+ emitters. First is by the loss of one or two neutrons from the carbon, yielding ^{11}C or ^{10}C with lifetimes of 20 min and 19 s, respectively. As the neutron stripping does not change the velocity, the remaining carbon ions have nearly the same range and stop close to the Bragg maximum of the stable carbon ions. Because of their long lifetimes compared to the stopping time, they decay mostly at the end of their range forming a maximum (Figure 18.5). By monitoring their decay via PET, the range of the primary particles can be measured *in vivo* without applying an additional dose.

The second process that produces β^+ emitters is the activation of the target atoms, mainly the production of ^{15}O with a lifetime of 122 s. As the reaction cross section is relatively constant over the complete energy range, the target activation and the produced isotopes are homogeneously distributed over the total range without a peak at the end. But using a special analysis, it is also possible to verify the particle range. This second method can be used for all ions, as well as for protons and other ions lighter than carbon that do not produce β^+ emitters by fragmentation. But the target activity helps determine the quality of the irradiation *in vivo*.

From the therapeutically relevant heavy ions, the carbon ions have the simplest β^+ spectrum consisting of two carbon isotopes. For the heavier ions, the spectra are more complex and contain many different β^+ lines, which make the analysis more difficult. But in all cases, it is not possible to recalculate the dose from the measured PET distribution. In practice, the expected β^+ distribution is calculated from the planned dose, and the measured PET distribution is then compared to the precalculated one. From such a comparison, discrepancies in the order of 2 mm in range distribution could be detected. The online PET control was routine at the GSI carbon therapy where a limited-angle PET camera registered the β^+ activity during and a short time after the patient exposure. Because of the limited space available during exposure, offline PET analysis is presently under research where the patient immediately after exposure is transferred to a commercial full-ring PET scanner and the residual β^+ activity is scanned. Offline PET is installed at the Heidelberg facility [6].

18.6 Radiobiological Advantages of Carbon Ions

The main reason to use ions heavier than protons in tumor therapy is the elevated RBE that can potentiate the damage especially to otherwise highly resistant tumors [7].

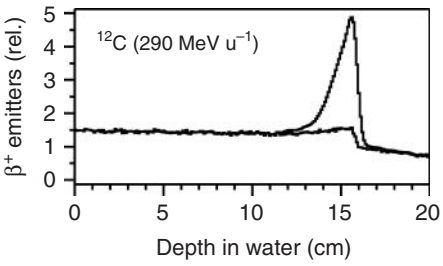


Figure 18.5 β^+ Activity of a 290 MeV u^{-1} carbon beam. The upper line represents the activity of the carbon-10 and carbon-11 ions showing a maximum at the position of the Bragg maximum. The continuous plateau originates from target activation of mainly ^{15}O isotopes adapted from [6].

It was recognized very early that the biological response to the same physical dose changes when radiation of different quality, for instance, α particles or ions, is used. To account for this change, the RBE was very empirically introduced as the ratio of X-ray dose as the standard radiation to the test radiation, necessary to produce the same effect. For sparsely ionizing radiation such as X-rays and γ-rays, the RBE = 1; for densely ionizing radiation, the RBE > 1 and can reach clinical values up to 2–4. The RBE is a biological consequence of high ionization densities. Because of a very efficient cellular repair system, X-ray damage to cells can be repaired at low doses, yielding a shoulder in the dose–response curve (Figure 18.6). For ions, the repair is suppressed to a large extent and the dose–effect curve becomes more linear when linear energy transfer (LET) increases.

But RBE is a complex function of many parameters. First, RBE depends on the effect level because of the nonlinearity of the X-ray dose–effect curve, which is taken as a reference. The RBE of low effect levels (80% survival) is greater than that for high effect levels (10% survival). At extremely high doses, both the curves are parallel to each other and RBE comes close to 1. On the biological side, RBE depends mainly on the radiosensitivity of the cells: for highly radio-resistant cells and tumors, cell killing is more potentiated as for radio-sensitive cells, and RBE values up to 3 have been clinically verified for radio-resistant tumors.

Second, RBE depends on the energy and the energy loss of the particles, which is usually classified in radiobiology as LET and measured in kilo electronvolts per micrometer (keV/μm)(Figure 18.7).

When the carbon beam enters the tissue with high energy, the LET is small and RBE is close to 1. Then with decreasing energy, the LET increases, yielding greater dose and greater RBE values. At the end of the range, both LET and RBE increase to their largest values in the Bragg maximum.

For other ions, these functional dependencies are less favorable for therapy. For the heavier ions, the RBE increases already at the entrance channel in the normal tissue and drops to smaller values at the position of the Bragg peak, causing severe damage to the normal tissue in front of

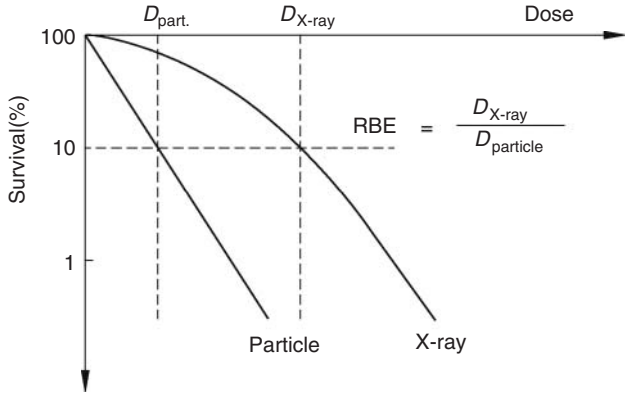


Figure 18.6 Schematic dose–response curves and the definition of the relative biological effectiveness (RBE) as the ratio between the doses of the standard radiation and the test radiation to produce the same effect.

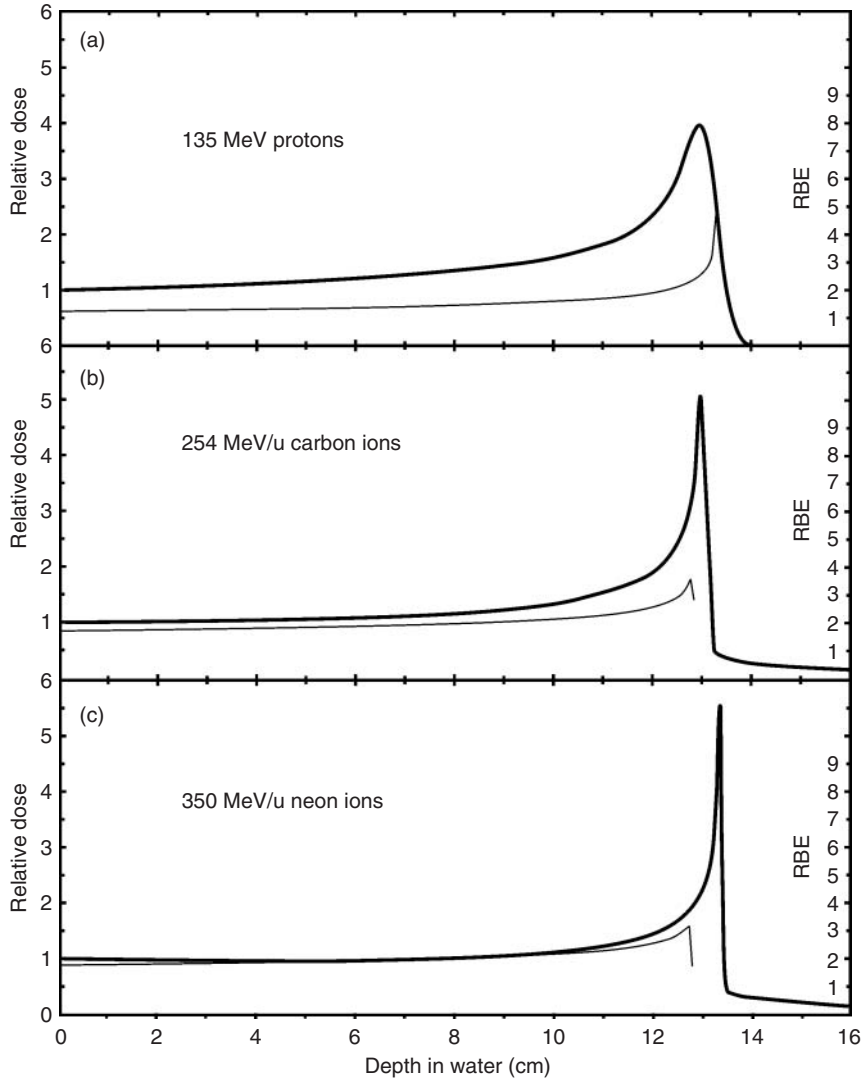


Figure 18.7 Comparison of the RBE (right scale) and the dose deposition (left scale) as a function of penetration depth for different ions. For protons, RBE is elevated at the last few millimeters beyond the Bragg maximum; for the heavy neon ions, the RBE maximum occurs before the Bragg maximum. But for carbon ions, the RBE maximum coincides with the Bragg maximum potentiating the high dose by an elevated RBE.

the tumor. For protons, the RBE increases only at the last few millimeters of the proton range where the dose of the Bragg maximum already starts to decreases. Only for carbon ions the range of increased RBE values coincides with the high dose of the Bragg peak. Consequently, the strategy of a successful tumor therapy is to restrict the range of high local dose and RBE stringently to the tumor volume and to allow only radiation of low biological effectiveness to hit normal tissue in the entrance channel.

But this strategy also implies that the dose application system restricts the irradiated volume strictly to the tumor and that the normal tissues are spared.

The RBE dependencies from LET can be explained on the basis of the ions track structure and the properties of the biological target molecule: by far, the largest and the most important molecule for cell killing is the double-stranded DNA. Because its integrity is so important for biological evolution, it is protected by a very efficient intracellular repair system and the genetic information is redundantly encoded in both strands. To create a non-repairable lesion, it is necessary to produce multiply damaged sides, for instance, clustered double-strand breaks within the distance of a few DNA bases. Therefore, many ionization events in an area in the order of a few nanometers have to be active.

The **primary interactions of photons** are photoeffect or the Compton effect and pair production yielding isolated ionization events and electrons that can later produce secondary ionization events. These are single events and the mean free distance between two ionization events in the slowing down of an electron is larger than the molecular dimension of the DNA. But at the end of the electron range, that is, at energies of about 100 eV, the main free path length reaches a value of 1 nm, which is in the order of the distance of the two DNA strands. At this range, one electron can produce a double-strand break. To produce multiply damaged sides, the total numbers of the electrons, that is, the total exposure dose, has to be increased to reach some areas of randomly overlapping electron tracks of low energies.

In particle tracks, the primary ion interacts mainly with the target electrons and produces a track of liberated electrons along the path of the primary ion. The number of liberated electrons depends on the energy loss of the ions and is largest at the end of the range in the Bragg maximum. The diameter of the track is given by the highest energy of the electrons, which is given – according to billiard ball kinetics – by the mass ratio between both and decreases with the ion energy. The radial dose distribution inside the track is inversely proportional to the distance from the center of the track. Thus, for light ions at the end of the range, many low-energy electrons are liberated that are biologically very efficient. For the greater ion energies before, the electron density is not large enough to produce multiply damaged sides in a single track. This is, however, the case for heavier ions that have a greater energy loss from the beginning, and therefore, there is a large amount of unwanted biological damage in the particle entrance in the normal tissue in front of the tumor.

At the biological side, the processing of the multiply damaged sides differs based on the cell type, the cell proliferation status, and the organization in the corresponding tissue, as well as on other environmental conditions such as oxygen pressure. Therefore, it is not possible to produce a theoretical description of RBE on the basis of physical parameters only – as it has been tried many times-the

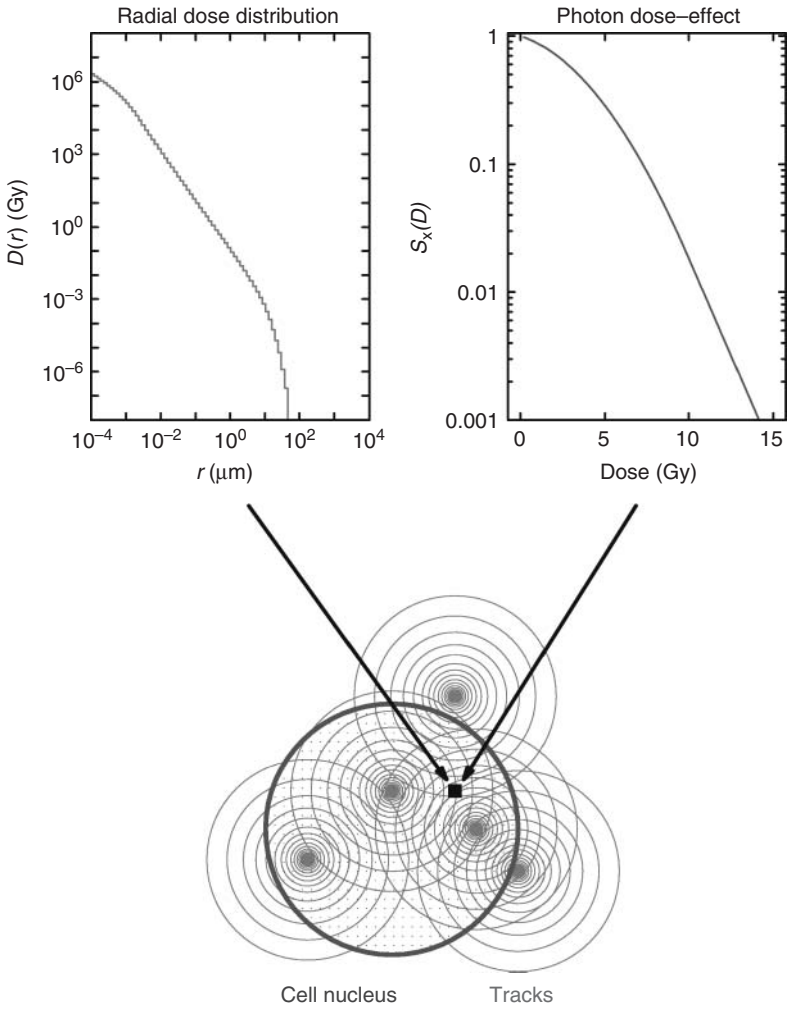


Figure 18.8 Illustration of the basic principles of the local effect model. The cell nucleus is penetrated by a number of ion tracks according to the macroscopic dose, and the local dose distribution is calculated as superposition of the tracks. For the dose of each nuclear pixel, the expected biological effect is then taken from an X-ray dose–effect curve.

repair processing of the specific biological target also has to be taken into account, as it is the case in the LEM Figure 18.8.

The combination of biological repair and local physical dose has been handled in a heuristic way in the LEM [8]. First, the structure of the ion track is calculated according to its energy loss and its energy, which gives the total radial dose distribution (Figure 18.8). Then the cell nucleus is overlaid with many tracks according to the proposed macroscopic dose and is divided into small volume elements, voxels, selected in such a way that the dose inside the voxel stays constant within a small percentage. Finally, for each individual voxel, the expectation values of biological response are taken according to an X-ray dose–effect curve taken for the same cells or tissue. This X-ray response contains the repair capacity of the specific tissue or cell. For therapy planning, the X-ray response curve has to be taken from the experience of the same type of tumors treated with X-rays in patients. From many dose points, a dose effect and hence the RBE can be calculated.

It is not correct to transfer the cellular RBE values to therapy: cells in a Petri dish behave different than that in an organized tissue *in vivo*. The RBE values taken for treatment planning have to be calculated from the experience of X-ray radiotherapy. This has been performed with great success at the GSI experimental therapy on the basis of the LEM and is now applied at the Heidelberg Ion-Beam Therapy (HIT) Center and at other places.

Meanwhile, the LEM has been improved and additional effects have been added, such as the diffusion of chemical radicals in the track, the interaction of single-strand breaks to form a double-strand break, the energy dependence of the width of the track center, and finally, the interaction of double-strand breaks to form a cluster. The improved final version, LEM IV, is able to describe not only the average tumor inactivation but also the expected effects in the normal tissue that is affected by irradiation in the entrance channel.

18.7 Treatment Planning

Treatment planning for particle therapy starts with the same procedures as in conventional therapy. In an appropriate CT image, the target volume has to be delineated slice by slice and the critical organs have to be delineated by the physician and medical physicist. Then the entrance channels are selected and the CT is turned into the beam direction (beam's eye view). In the following steps, particle planning differs from conventional therapy because it is a 3D planning: for each voxel, there should be a separate calculation for the number of stopping ions that are necessary to reach a homogeneous dose or – including the local RBE values – a homogeneous biological effect. For this procedure, target volume is dissected in slices of equal energy. Starting from the most distal one, each isoenergy slice is then covered with a net of pixels along the continuous path of the scanned beam. Then the most distal slice is covered homogeneously with particles. For all the following slices, the preirradiation dose by the penetrating ions of the more distal sides has to be subtracted from the nominal dose. Then the same procedure can be repeated as before, and the slice is covered by particles in a nonhomogeneous way to reach in the sum, the wanted effect. This procedure becomes more complex when the RBE has to be incorporated. The local RBE values have to be calculated pixel by pixel, according to the radiation field and the repair capacity of the tumor tissue. Because RBE depends on dose, the

corresponding local RBE values have to be recalculated in the optimization procedure when the local dose is changed during the optimization process. Therefore, the biological optimization becomes more elaborate than the first step, the physical optimization. Although when all necessary RBE values corresponding to the various parameter – sets concerning particle's energy and atomic numbers – are calculated before and stored in lookup tables, the treatment planning procedure is time consuming. But for heavy ion therapy, it is the main key for clinical success [9].

18.8 Clinical Results and the International Situation

Until now, worldwide approximately 80 000 patients have been treated with particle beams, mostly protons, and 9000 patients with carbon ions. These treatments were very successful and high tumor control rates in terms of 5 year disease-free survival have been reached. But even with these large numbers, it is difficult to compare the data with conventional therapy because they have been accumulated over several decades, while the techniques in both arms, conventional and particles, have improved significantly. In both cases, there was better conformity in case of photons using IMRT and particles using beam scanning, frequently called *intensity-modulated particle therapy* (*IMPT*).

In the past, in particle therapy, the majority of the patients have been treated using passive beam delivery system that was superior to the contemporaneous photon technique but is not so in comparison to the present IMRT and other advanced photon techniques [10].

In addition, the clinical outcome of particle therapy has not been confirmed in randomized studies. Also other advanced therapies have never been subjected to randomized studies. When sufficient evidence was gained that local tumor control and the exposure of the normal tissue were superior, particle therapy was accepted as the better treatment without randomized studies. This point has been reached for particle therapy using conformal beam delivery. In Figure 18.9, therapy results of the patients treated with carbon ions are compared with those treated with other therapy modalities. Carbon ion treatment provides higher tumor-targeted doses with tolerable side effects, thus yielding 5 year tumor control rates of 100%, which is perfect.

But the largest clinical experience exists at NIRS Chiba, Japan, that operates the Heavy Ion Medical Accelerator at Chiba HIMAC since 1994 [11]. This complex consists of two synchrotrons on top of each other and accelerates mainly carbon ions for therapy, although it could be used for ions as heavy as argon. Worldwide, most of the patients treated with heavy ions, that is, carbon ions, were patients of NIRS. As the passive beam preparation is less perfect in conformity, the dose to the normal tissue is greater than that for scanned beams. But as a great advantage, the passive method can also be used to treat tumor volumes that are moving because of respiration, such as tumors in the lungs and liver. In consequence, many tumor sites all over the body have been treated successfully at NIRS, especially prostate, liver, head and neck, and lungs, as well as brain, uterus, and others. Recently, NIRS has constructed a new irradiation area where a few patients have been treated with scanned beams. In this method, the treatment of a complete field could be reduced to less than 1 min, which in future would enable to treat moving tumors with breath-holding techniques.

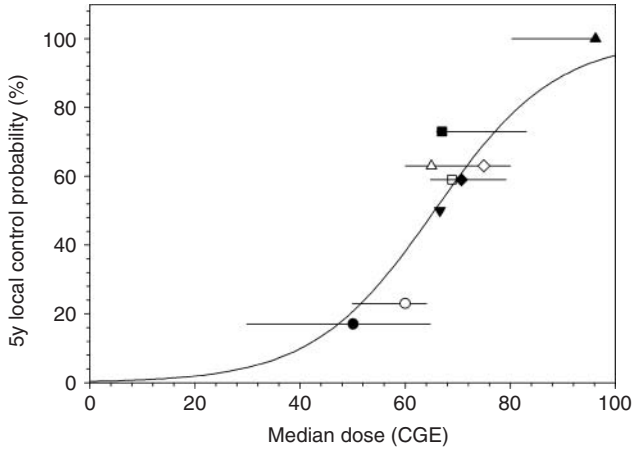


Figure 18.9 Tumor control rate of chordomas in the head as a function of the total dose CGE, cobalt gray equivalent. The low control data between 20 and 30% are X-ray treatments. Higher control rates at around 70% are reached with intensity-modulated photon or proton beams .The 100% point was reached using scanned carbon beam.

As a consequence of the clinical success of NIRS, two centers at Gunma and Hyogo are treating patients since 2002 and 2010, respectively, with carbon beams and the HIMAT facility at Tosu is under construction. The fifth carbon facility, the Ion-beam Radiation Oncology Center in Kanagawa (i-ROCK) is in preparation.

Because of the good clinical results of the GSI pilot therapy, new projects for particle therapy based on scanned beams have evolved in Europe: in 2009, the HIT Center started carbon ion and proton treatment at two horizontal fixed beams. A rotating beam applicator, the so-called gantry, was set in operation in 2012 and allows patient treatment under all angles. At the HIT Center, the good results from the GSI pilot project could be confirmed. In addition, new tumor entities such as prostate and liver could be irradiated with scanned beams.

In 2004, the Siemens company went in the field of heavy particle therapy on the basis of exclusive licenses of the GSI patents. The beam applicator, that is, the scanning system at the HIT Center, was made by Siemens. Then, parallely, new facilities at the university hospitals at Marburg and Kiel, as well as at Shanghai, were constructed.

In all these facilities, horizontal beam applicators were combined with oblique beams at 45° to have a greater freedom in the selection of the entrance ports. At present, the Marburg facility is ready for clinical routine but is used only as a test facility for the Shanghai project. In 2011, Siemens decided to withdraw completely from radiation therapy, conventional as well as particles. Consequently, the Kiel facility was not finished and is presently dismantled. But the facilities at Shanghai and hopefully Marburg will start clinical operation in 2013.

At Pavia, Italy, the Centro Nazionale di Adroterapia Oncologica, CNAO, has started its operation in 2010 first with proton beams, and carbon treatments in 2013. The CNAO facility is the result

of a Proton Ion Medical Machine Study (PIMMS) at CERN where the master layout of a therapy center was studied in detail combined with an injector manufactured by GSI.

Similar to the CNAO facility, the MedAustron project at Wiener Neustadt is constructing a particle therapy facility for proton and carbon beam treatment. MedAustron will have four exposure areas, one dedicated mainly for research and the other three for clinical use. It is supposed to start patient treatment in 2015.

18.9 Conclusions

In summary, the use of heavy particles in therapy has a long and successful history: proton therapy using passive and, recently, active beam modulation has become a standard especially for the treatment of children because of the lower dose contamination to the normal tissue than that in conventional therapy. Proton facilities have spread worldwide but are mostly concentrated in the United States.

The carbon therapy, especially when given with scanning techniques, yields another step of conformity and efficiency. Treatment with carbon is most advanced in Japan where three carbon therapy centers are in clinical operation, one more center is under construction and another one is in the planning stage. In Europe, two carbon/proton therapy centers are in clinical operation and two are under construction and in the testing phase. There are several additional projects, but they are not yet decided.

In general, it has been proven that particle therapy helps many patients. Therefore, the numbers of these facilities will grow, but rather slowly because of the great investment required for such facilities.

Glossary

Bragg Maximum: Maximum of the energy loss of particles close to the end of their range, the height, and range position depend on the charge and mass of the ions and the stopping material.

Bremsstrahlung: The electromagnetic radiation (bremsstrahlung) from high-energy electrons (5–20 MeV) stopped in a heavy target (tungsten/tantalum) is used in modern photon therapy as the radiation source.

Spread Out Bragg Peak/Extended Bragg Peak: For tumor therapy, the dimensions of the Bragg maximum of a few millimeters FWHM have to be enlarged to the size of the target volume either by passive modulators (spread out Bragg peak, SOBP) or by active systems (extended Bragg peak, EBP).

CT, computed tomogram: Several X-ray images of a patient are used to reconstruct the 3D images of the inner parts of the patient's body.

Dose: Energy deposited in matter, the unit is gray; $1\ \mathrm{Gy} = 1\ \mathrm{J\ kg^{-1}}$.

DNA: Deoxyribonucleic acid (DNA) is the carrier of the genetic information in the biological cells.

Heavy/Light Ions: In radiobiology and radiotherapy, all ions heavier than protons are called *heavy ions* because of their significantly different biological action.

Hadron Therapy: Originally termed to include neutron-, pion-, proton, and heavy ion therapy, but now it is mainly used by physicians for the heavy ion therapy in contrast to proton therapy. Also, CERN scientists like to use the term *hadron therapy* for obvious reasons.

IMRT, intensity-modulated radiotherapy: A high-energy photon beam (electron

bremsstrahlung) is given with different intensities from different entrance ports (5–10) to the target volume.

IMPT: The same as IMRT but uses particle beams, see: raster scan, voxel scan.

LEM, local effect model: A biological model to calculate the response of different cells or tissues to radiation of different quality such as ions or neutrons.

LET, linear energy transfer: The energy transfered to the biological target. It is equal to the energy loss as used in physics when all electrons stop in the biological target.

Normal tissue: The nonmalignant tissue around the tumor, which might be not healthy but affected from other diseases or burdens, for instance, a smokers's lung, which is not healthy but normal.

Raster scanning: In order to irradiate an extended target volume, a pencil beam is moved in a raster like pattern over each iso-energy slice of the target volume without turning the beam off.

RBE: The relative biological effectiveness is defined as the ratio of the dose of particle irradiation to the dose of X-rays necessary to achieve the same biological effect.

Target volume: The volume that is considered by the physician to be irradiated; it is larger than the tumor volume.

Voxel scanning: The target volume is divided in small volume element ''voxels'' that are irradiated sequentially. The beam is turned off when moved from one to the next voxel.

References

1. Particle Therapy Co-operative Group An Organization for those Interested in Proton, Light ion and Heavy Charged Particle Radiotherapy, *ptcog.web.psi.ch/ptcentres.html* (accessed 14 February 2013). In this reference the last updated list of all particle therapy centers and patient statistics is given.
2. Chu, W.T. *et al.* (1993) *Rev. Sci. Instrum.*, **64**, 2055–2122.
3. Kraft, G. (2000) *Prog. Part. Nucl. Phys.*, **45**, 473–544.
4. Haberer, T. *et al.* (1993) *Nucl. Instrum. Methods Phys. Res.*, **A 330**, 296–305.
5. Blattmann, H. *et al.* (1995) *Prog. Radio-Oncol.*, **V**, 347–352.
6. Parodi, K. *et al.* (2008) *Nucl. Instrum. Methods Phys. Res.*, **A 591**, 282–6.
7. Weyrather, W.K. *et al.* (2004) *Radiother. Oncol.*, **73** (Suppl. 2), 161–169.
8. Elsässer, T. *et al.* (2010) *Int. J. Radiat. Oncol. Biol. Phys.*, **78** (4), 1177–1183.
9. Jaekel, O. *et al.* (2001) *Phys. Med. Biol.*, **46**, 1101–1116.
10. Schulz-Ertner, D. and Tsujii, H. (2007) *J. Clin. Oncol.*, **25** (8), 953–964.
11. Kanai, T. *et al.* (1999) *Int. J. Radiat. Oncol. Biol. Phys*, **44** (1), 201–210.

Further Readings

Brada, M., Pijls-Johannesma, M., and Ruysscher, D.D. (2007) Proton therapy in clinical practice: current clinical evidence. *J. Clin. Oncol.*, **25** (8), 965–970.

Durante, M. and Loeffler, J.S. (2010) Charged particles in radiation oncology. *Nat. Rev. Clin. Oncol.*, **7** (1), 37–43.

Greco, C. and Wolden, S. (2007) Current status of radiotherapy with proton and light ion beams. *Cancer*, **109** (7), 1227–1238.

Lodge, M., Pijls-Johannesma, M., Stirk, L., Munro, A.J., Ruysscher, D.D., and Jefferson, T. (2007) A systematic literature review of the clinical and cost-effectiveness of hadron therapy in cancer. *Radiother Oncol.*, **83** (2), 110–122.

Schardt, D., Elsässer, T., and Schulz-Ertner, D. (2010) Heavy-ion tumor therapy: physical and radiobiological benefits. *Rev. Mod. Phys.*, **82** (1), 383–425.

Part C
Nuclear Power

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

19
The Physics of Nuclear Power from Fission and Fusion

Weston M. Stacey

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

19.1
Introduction

The objective of this chapter is to provide a brief overview of the physics of nuclear fission reactors, which provide nuclear power today, and of magnetically confined fusion plasmas, which will provide nuclear power in the future, as an introduction to this section on nuclear power. The organization of this chapter will follow that of my recent textbooks [1, 2] on these subjects, specific chapters of which will be referenced for those wishing more details (e.g., Chapter 3 of [1] will be indicated [1(3)]). Comprehensive developments or reviews of some specific subjects will be referenced, but this chapter is not intended to provide a comprehensive review of the fields covered. Major historical books that have been landmarks in the two fields are listed in the section on Further Reading.

19.2
Binding Energy [1(1)],[2(1)]

The actual mass of the stable atomic nuclei (m_Z^A) is not the sum of the masses of the Z protons corresponding to the atomic number protons and the masses of the $A-Z$ neutrons corresponding to the atomic mass neutrons. Rather, there is a mass defect $\Delta m_Z^A \equiv [Zm_\mathrm{p} + (A-Z)m_\mathrm{n}] - m_Z^A$ and a *binding energy* per nucleon $\mathrm{BE}/A \equiv \Delta m_Z^A c^2/A$, which is plotted in Figure 19.1 for the stable nuclei. Any process that causes nuclides to be converted to other nuclides with greater BE/A results in the conversion of mass into energy. Clearly, the fission of a heavy mass nucleus into two intermediate-mass nuclei converts mass into energy (about 200 MeV), and the fusion of two very low mass nuclei into a somewhat heavier nucleus, which decays immediately, converts mass into energy [17.6 MeV in the case of D + T fusing to form He (3.5 MeV) and a neutron (14.1 MeV)]. The controlled production, recovery, and conversion to electricity of this energy constitutes the applied nuclear physics discipline of *nuclear engineering*.

19.3
Physics of Neutron Chain Fission Reactors

19.3.1
Neutron-Induced Fission [1(1,2)]

While some of the heavier nuclides (e.g., $^{240}\mathrm{Pu}_{94}$) fission spontaneously, the rate is small. However, the excited energy levels of the heavier nuclides are low-lying, and neutron capture can form an excited state that frequently de-excites by fission.

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

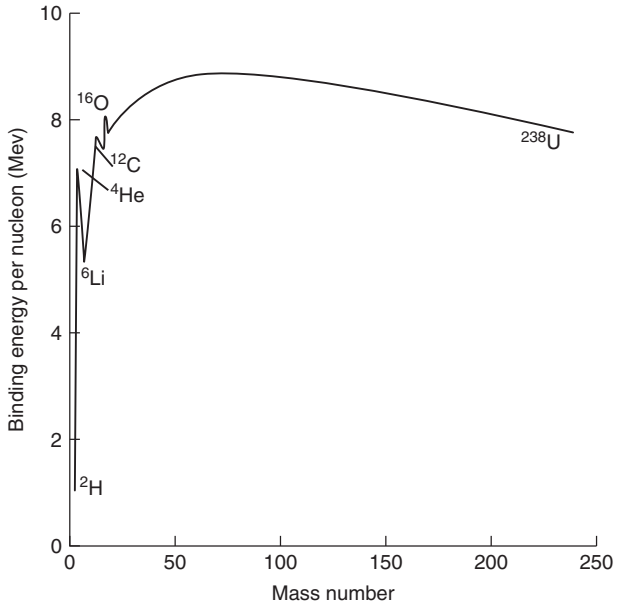


Figure 19.1 Binding energy per nucleon for the stable isotopes.

The odd-mass nuclei (e.g., $^{233}U_{92}$, $^{235}U_{92}$, $^{239}Pu_{94}$, $^{241}Pu_{94}$) are particularly low-lying and can be excited by the capture of "thermal" neutrons with energies of only about 0.025 eV, but the even-mass nuclides (e.g., $^{238}U_{92}$, $^{240}Pu_{94}$) normally require the capture of an energetic (>10 keV level) neutron to excite a higher energy level that will de-excite by fission. Most of the energy released in fission is in the form of the kinetic energy of the recoil nuclei formed in fission, and is deposited as heat in the fuel element within millimeters of the fission event, but the ≈10% of the energy that is in the form of fast neutrons and γ-rays is deposited in the surrounding (10–100 cm) material.

The fission process produces two intermediate-mass nuclei, distributed over $70 < A < 160$ with peak yields of nuclides with $90 < A < 100$ and $130 < A < 150$, two or three neutrons that are distributed in energy between $0 < E < \approx 6$ MeV but with a peak at about 1 MeV, and γ-radiation. As the ratio of neutrons to protons in the nucleus of the stable isotopes increases with A, these intermediate-mass *fission products* are "neutron rich," and will subsequently decay to a stable isotope by emitting electrons, neutrons, α-particles, and γ-rays. This radioactive decay is a source of *after-heat* that persists even after the fission reactions have ended, and the presence of the *delayed neutrons* has a profound effect on the controllability of nuclear reactors.

19.3.2 Neutron Moderation [1(1,4,10,12)]

The fission neutrons are born predominantly with energies in the megaelectronvolts range, but are quickly moderated into the 100 keV range by inelastic scattering and then are further moderated in energy by elastic scattering to an extent depending on the material composition. Elastic scattering kinematics yields a relationship between the

average logarithmic energy loss and the atomic number A of the material $\xi \equiv \langle \ln E_{\text{before}} - \ln E_{\text{after}} \rangle = 1 - [(A-1)^2/2A]\ln(A+1/A-1)$. For mixtures, a composite ξ can be constructed by weighting with the atom density and elastic scattering cross section of each constituent. The number of collisions required to elastically moderate a fission neutron from $E_0 = 2$ MeV to E_1 is then $\# = \ln(E_0/E_1)/\xi$.

In order to moderate a fission neutron from 2 MeV to $E_1 = 1$ eV, the upper bound of the thermal energy range, the number of collisions required is 16 for H_2O, 20 for D_2O, 43 for He, 91 for C, 171 for Na, 411 for Fe, and 1730 for ^{238}U. Of course, the effectiveness of a material as a moderator also depends on the likelihood of a neutron having an elastic scattering collision instead of an absorption reaction. The product of ξ and the ratio of the elastic scattering probability to the absorption probability, known as the *moderating ratio*, is 71 for H_2O, 5670 for D_2O, 83 for He, 192 for C, 1134 for Na, 35 for Fe, and 0.009 for ^{238}U.

One classification of nuclear reactors is according to the neutron energy spectrum. Those reactors with water coolant or gas coolant and graphite moderating material generally have neutron spectra that are predominantly *thermal* ($E < 1$ eV), while those reactors that are cooled by sodium or liquid lead-bismuth have neutron spectra that are predominantly *fast* (1 keV $< E <$ 1 MeV).

19.3.3
Neutron Transport [1(3,9)]

The fission neutrons have speeds on the order of 10^7 m s^{-1}, which would be reduced to 10^{3-4} m s^{-1} if the neutron is moderated to the thermal energy range. The elastic scattering collisions that cause the moderation in energy also cause a change in direction, with the consequence that the neutrons also execute a random-walk transport process that takes them out of the fuel region in which the fission occurred and into other fuel and non-fuel regions with different moderating properties. The root-mean-square distance that a fission neutron travels before being absorbed is about 15 cm in H_2O, 425 cm in D_2O, and 150 cm in graphite, with most of this displacement taking place before the neutron is thermalized. Taking into account the juxtaposition of fuel, structure, and moderator/cooling regions in nuclear reactors, the root mean square distance a neutron travels from fission to absorption is about 15 cm in a water-cooled thermal reactor; about 50 cm in a gas-cooled, carbon-moderated thermal reactor; and about 14 cm in a sodium-cooled fast reactor.

Elaborate and sophisticated computation methods [1(9)] have been developed to simulate the transport of neutrons.

19.3.4
Neutron Fission Chain Reaction [1(2,6)]

Two or three neutrons are produced per fission event, so that on average $\nu \approx 2.5 - 2.8$, depending on the fissionable species and the energy of the neutron causing the fission event (hence fast or thermal spectrum), neutrons are produced in a fission event. If, on average, exactly one of these neutrons causes another fission, then there is a *critical* self-sustained neutron chain fission reaction, and the number of fission neutrons remains constant in time. If, on average, fewer than one neutron from a fission event causes another fission, the reactor neutron population decays exponentially, and the reactor is said to be *subcritical*. On the other hand, if more than one neutron, on average,

from a fission event survives to cause another fission event, the reactor neutron population increases exponentially, and the reactor is said to be *supercritical*.

The probability that a fission neutron causes another fission is designated as $\Sigma_f \equiv N\sigma_f$, the total probability that a fission neutron is absorbed in a nucleus as $\Sigma_a \equiv N\sigma_a$, and the probability that a fission neutron leaks from the system as $\Sigma_{leak} \equiv N\sigma_{leak}$. The quantity N is the density of nuclei; the σ_f and σ_a are microscopic nuclear reaction cross sections for fission and for absorption (fission + radiative capture); σ_{leak} represents the leakage probability; and a suitable average over nuclear species, space, and energy is implied. The total mean-free-path for a fission neutron to either be absorbed or lost is $\lambda_{mfp} \equiv [1/(\Sigma_a + \Sigma_{leak})]$; so the lifetime of a neutron in the system is $l_n \equiv \lambda_{mfp}/\upsilon = 1/\upsilon(\Sigma_a + \Sigma_{leak}) \approx 1/\upsilon\nu\Sigma_f$ for a system near critical. The criticality of the system can be expressed in terms of the multiplication constant $k \equiv \nu\Sigma_f/(\Sigma_a + \Sigma_{leak})$. For $k < 1$, the reactor is subcritical, and in the absence of an external source, the neutron population will decrease in time; for $k = 1$ exactly, the reactor is critical, and the neutron population in the reactor will remain constant in the absence of an external source; and for $k > 1$ the reactor is supercritical, and the neutron population in the reactor will increase over time.

All of the heavy mass nuclei mentioned earlier, including uranium, have a significant probability of fissioning during the absorption of a high-energy neutron, but only the odd-A heavy mass nuclei have a significant probability of fissioning during the absorption of a thermal neutron, and these nuclei are referred to as *fissile*. The fissionable material in most existing nuclear reactors is $^{235}U_{92}$, which constitutes 0.72% of natural uranium and which will fission with thermal or fast neutrons. The uranium is "enriched" in $^{35}U_{92}$ to 3-5% for most thermal spectrum "commercial" reactors and to higher concentrations for other reactors.

$^{238}U_{92}$ and $^{232}Th_{90}$ are referred to as *fertile* materials because neutron transmutation of these nuclei into fissile nuclei occurs in nuclear reactors. Neutron capture in $^{238}U_{92}$ produces $^{239}Np_{93}$, which β-decays into $^{239}Pu_{94}$ within a couple of days. Neutron capture in $^{239}Pu_{94}$ produces $^{240}Pu_{94}$, and neutron capture in $^{240}Pu_{94}$ produces $^{241}Pu_{94}$. This neutron capture/decay chain continues to produce the transuranic (TRU) isotopes (higher plutonium, americium, curium). By a similar transmutation chain, $^{233}U_{92}$ can be produced from $^{232}Th_{90}$.

The expression for the multiplication constant can be factored into a form that more clearly displays the physics of the neutron chain fission reactor.

$$\begin{aligned} k &\equiv \frac{\nu\Sigma_f}{(\Sigma_a + \Sigma_{leak})} \\ &= \left(\frac{\nu\Sigma_f}{\Sigma_a^F}\right)\left(\frac{\Sigma_a^F}{\Sigma_a}\right)\left[\frac{1}{(1 + \Sigma_{leak}/\Sigma_a)}\right] \\ &\equiv \eta f P_{nl} \end{aligned} \tag{19.1}$$

The term η is the number of neutrons produced per neutron absorbed in the fissile fuel. The term f is the ratio of the rate that neutrons are absorbed in the fissile fuel to the rate neutrons are absorbed in all materials or leak out of the reactor; it is a *fuel utilization factor*. The term P_{nl} is the ratio of the rate at which neutrons are absorbed by all materials to the rate at which they are lost by absorption plus leakage, or a *non-leakage probability*. [The fact that some fission occurs in the non-fissile fuel can be taken into account by adding a multiplicative factor ε to

Eq. (19.1).] The value of η and f depend on the nuclear cross sections for the reactor materials, and the value of P_{nl} depends on the geometry and nuclear cross sections.

19.3.5 Nuclear Reaction Cross Sections [1(1)]

The factor η for the principal fissile nuclei are shown in Figure 19.2. In the thermal reactor energy range ($E < 1$ eV) $\eta_{U233} > \eta_{U235} > \eta_{Pu239}$, but in the fast reactor energy range ($E > 5 \times 10^4$ eV) $\eta_{Pu239} > \eta_{U233} > \eta_{U235}$.

The fuel utilization factor f depends on the relative values of the absorption cross sections for the different materials (fuel, structure, fission products, coolant, moderator, etc.) in the reactor. The low energy absorption cross sections for the main fissile fuel isotopes, some of the higher Pu isotopes, two important fission products (xenon and samarium), and the *burnable poison* gadolinium are shown in Figure 19.3.

The strong variations in the epithermal energy region ($1 < E < 15$ eV) in Figure 19.2 and Figure 19.3 are due to absorption resonances associated with the absorbed neutron energy being just the right magnitude to excite a low-lying excited state of the compound nucleus.

19.3.6 Nuclear Reactor Dynamics [1(5,16)]

Taking into account that a fraction β of the ν neutrons produced by fission are not available until a later time $t_{delay} \equiv \lambda^{-1}$ when the unstable fission product *delayed neutron precursor* C has radioactively decayed, the balance equations governing the dynamics of the neutron population $n(t)$ and of the population of the unstable fission products $C(t)$ that decay to release a delayed neutron can be written as

$$\begin{aligned}\frac{dn(t)}{dt} &= \left[\nu\Sigma_f(1-\beta) - \left(\Sigma_a + \Sigma_{leak}\right)\right]\upsilon n(t) \\ &\quad + \lambda C(t) + S(t) = \frac{(\rho-\beta)}{l_n} n(t) \\ &\quad + \lambda C(t) + S(t) \frac{dC(t)}{dt} \\ &= \beta\nu\Sigma_f \upsilon n - \lambda C(t) = \frac{\beta}{l_n} n(t) - \lambda C(t)\end{aligned} \tag{19.2}$$

where $\rho = (\nu\Sigma_f - \Sigma_a - \Sigma_{leak})/\nu\Sigma_f = (k-1)/k$ is the *reactivity*, and S is any external source of neutrons.

For a subcritical system ($k < 1, \rho < 0$) with a constant external source S_0, the asymptotic solution to Eq. (19.2) is $n_0 = S_0 k l_n/(1-k)$.

The asymptotic solution for a critical system ($k = 1, \rho = 0$) with no external source is $n(t) = n_0, C(t) = C_0 = (\beta/l_n\lambda)n_0$, where the value of n_0 depends on how the system was brought from its initial shutdown state to the particular critical state. For parameters typical of a uranium-fueled thermal reactor, this last relation yields $C_0 \approx 1600 n_0$;that is, the population of delayed neutron precursors is much larger than the neutron population, indicating that the rate of decay of the former will determine the dynamics of the latter under normal operation.

The general solution of Eq. (19.2) in the absence of an external source is

$$\begin{aligned}n(t) &= n_0\left[\frac{\rho}{\rho-\beta}\exp\left(\frac{\rho-\beta}{l_n}t\right)\right. \\ &\quad \left. -\frac{\beta}{\rho-\beta}\exp\left(-\frac{\lambda\rho}{\rho-\beta}t\right)\right] \\ C(t) &= C_0\left[\frac{\rho\beta}{(\rho-\beta)^2}\exp\left(\frac{\rho-\beta}{l_n}t\right)\right. \\ &\quad \left. +\frac{\beta}{l_n\lambda}\exp\left(-\frac{\lambda\rho}{\rho-\beta}t\right)\right]\end{aligned} \tag{19.3}$$

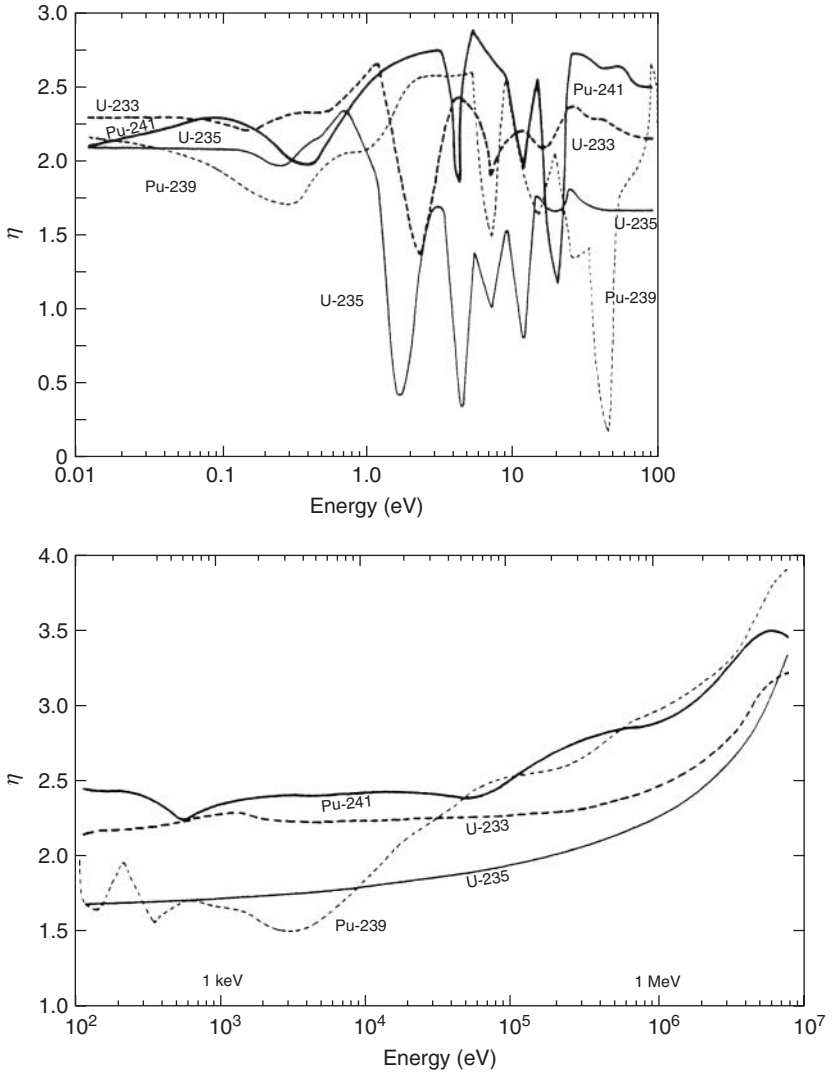


Figure 19.2 η for the principal fissile nuclides.

Representative values of the neutron lifetime l_n are 10^{-4} to 10^{-5} s in thermal reactors (where neutrons must undergo enough collisions to slow down from the megaelectronvolts range to the less than electronvolts range) and about 10^{-6} s for fast reactors. The delayed neutron fractions β are about 0.0026 for ^{233}U, 0.0065 for ^{235}U, 0.0064 for ^{239}Pu, 0.0054 for ^{241}Pu, 0.020 for ^{232}Th, 0.016 for ^{238}U, and 0.0029 for ^{240}Pu. The decay constants for the delayed neutron precursor fission products vary depending on the isotope, but are in the range of 0.012–4.0 s^{-1}.

If $\rho < \beta$, the first exponent in Eq. (19.3) is large and negative, and the first terms rapidly decay away, representing a *prompt jump* in the neutron population when the

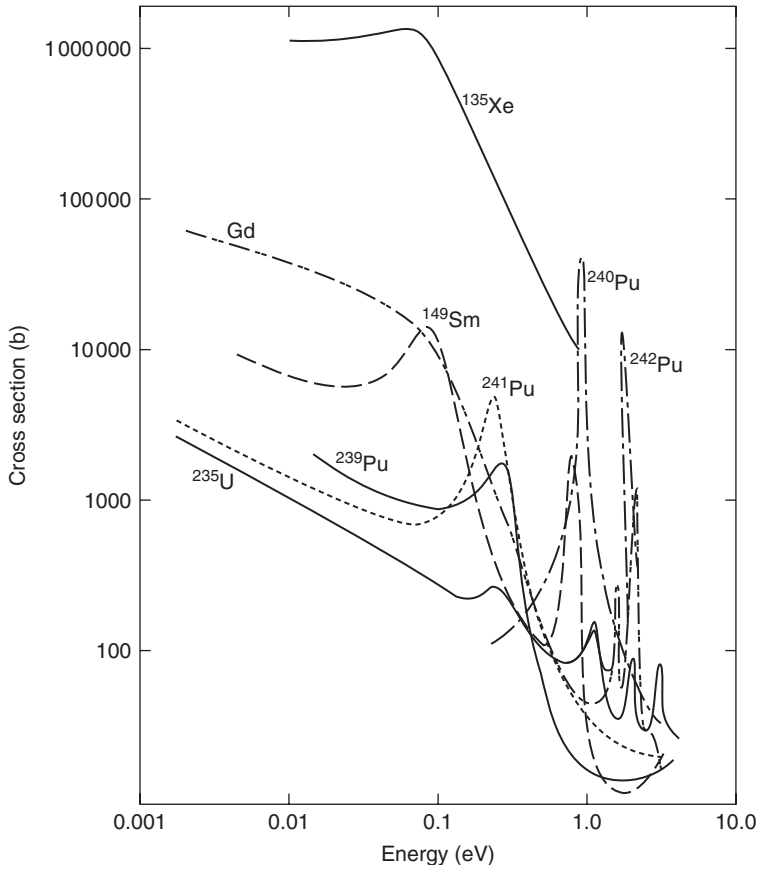


Figure 19.3 Low-energy absorption (fission + capture) cross sections for important nuclides.

reactivity changes from 0 to $\rho < \beta$. The subsequent dynamics is determined by the second term, which exhibits a long-time exponential growth ($\beta > \rho > 0$) or decay ($\rho < 0$) on the delayed neutron precursor decay time constant (λ^{-1}) timescale. As this timescale is on the order of 10 s, there is time to detect and correct or to compensate any problem causing an unintended reactivity change before a large change in the neutron population (and power level) results.

However, when $\rho > \beta$, the situation is quite different. The exponent in the first terms in Eq. (19.3) is now positive and, because of the small value of the neutron lifetime l_n, quite large, and the neutron population and power level increase exponentially on a very short period because the neutron balance is supercritical on the prompt neutrons alone (the first term) and does not need the neutrons provided by the much slower decay of the delayed neutron precursors (the second term) in order to rapidly increase the neutron population and power level. Fortunately, there are several physical mechanisms by which the increased temperatures associated with an increase in power level produce negative *feedback reactivity*,

which decreases, or even reverses, the rate at which the neutron population is growing.

19.3.7 Feedback Reactivity [1(4,5,11)] [3, 4]

As any increase in fission energy is deposited in the fuel element, the fuel temperature responds immediately to any change in fission rate. Many of the fertile isotopes in the fuel ($^{238}U_{92}$, $^{240}Pu_{94}$) have large absorption resonances in the epi-thermal region just above thermal (Figure 19.3), and both the fissile and fertile fuel isotopes have many excited states that can be formed by absorption of neutrons in the range $10–10^4$ eV.

The critical requirement for the formation of a compound nucleus excited state is that the relative energy of the neutron and the nucleus (i.e., the neutron energy in the center-of-mass system) and the difference in binding energy between the nucleus and the compound nucleus rather precisely agrees with the difference in energy levels between the ground and excited states of the compound nucleus. When the fuel lattice is heated, the thermal motion of fuel nuclei results in a wider range of neutron energies corresponding to the correct relative energy with respect to the nucleus for the formation of the excited state of the compound nucleus, in effect *Doppler broadening* the resonance absorption cross section in the lab frame. As neutrons moderate in finite energy loss steps, this broadening of the energy range in which neutrons in the fuel region will be absorbed increases the absorption rate.

If the absorption is primarily neutron capture (not fission), as it is for the low-energy resonances of the fertile materials, then this Doppler broadening reactivity is negative in thermal reactors. Some of the absorption resonances at higher energy in the fissile fuel isotopes are mostly fission reactions; so the Doppler broadening reactivity in fast reactors can be positive or negative, depending on the neutron energy spectrum.

About 1/10 s after an increase in fission energy deposition in the fuel elements, this energy reaches the surrounding coolant, increasing the coolant temperature, thereby decreasing the coolant density and thus decreasing the absorption rate of neutrons in the coolant, which produces a positive reactivity contribution. If the coolant also serves as the moderator, the decrease in moderator density decreases the probability that a neutron will scatter across the broadened resonance absorption energy interval and diffuse into the fuel region and be absorbed, which is a negative reactivity contribution in a thermal reactor but could be of either sign in a fast reactor, depending on the energy spectrum of the neutrons. Reduction in moderator density allows neutrons to diffuse further before being absorbed, thus increasing the leakage probability, which is a negative reactivity contribution.

In a fast reactor, a reduction in sodium coolant density reduces the neutron moderation further and produces a "harder" (more energetic) neutron spectrum. Because the capture-to-fission ratio for the principle fissile nuclides decreases as the neutron energy increases, this hardening of the neutron spectrum contributes a positive reactivity contribution.

Within about a second after the increase in fission energy deposition in the fuel elements, the warmer coolant will transfer some of this energy to the structure, causing it to expand and contort. The overall expansion of the reactor means that a neutron must travel further to escape, which reduces the leakage rate and contributes

a positive reactivity contribution. Because the neutron distribution and *importance* is not uniform in a reactor, if the structural expansion moves the fuel outward into a region of lower neutron population and importance, this would produce a negative reactivity coefficient, and conversely an inward fuel expansion into a region of larger neutron population and importance would produce a positive reactivity contribution. (The importance of a neutron is related to the likelihood that it will subsequently produce a fission.)

For a drop in fission rate causing an immediate drop in fuel temperature, the same chain of events as described earlier occurs, but with decreases, rather than increases, in temperatures, and the resulting feedback reactivity contributions are of opposite signs.

Light water reactors (LWRs), which have negative temperature feedback reactivity characteristics, are capable of automatically adjusting the power output to ''follow the load.'' A demand for more or less power results in more or less heat being extracted from the heat removal system to produce more or less power, which in turn results in cooler or warmer coolant entering the core, producing a positive or negative reactivity contribution, which causes the fission power to increase or decrease, which in turn causes an increase or decrease in the coolant temperature, and so on until equilibrium conditions are established at the new power level set by the external demand for power.

Reactors can be designed to take advantage of feedback reactivity to shut down when an off-normal event takes place. The most convincing demonstration was in the Experimental Breeder Reactor II, a metal-fueled fast reactor that sat in a pool of sodium coolant [5]. The operators intentionally turned off the pumps that circulated the sodium through the reactor, leading to a large increase in the sodium temperature within the reactor, which subsequently caused thermal expansion of the fuel assembly (negative feedback reactivity) and a reduction in sodium density leading to increased leakage (negative feedback reactivity). The reactor shut itself down, and natural circulation of the sodium prevented damage to the fuel. In a separate experiment, the operators shut down the pumps in the intermediate heat removal loop, essentially removing the heat sink provided by the normal power production, and causing the heat to be dumped into the sodium pool, raising the inlet sodium temperature to the reactor and thereby producing negative feedback reactivity. Again, the reactor shut itself down without damage.

19.3.8 Fuel Burnup [1(6)]

At the beginning of life (BOL), a nuclear reactor is loaded with fresh uranium fuel enriched in $^{235}U_{92}$ (to 4-5% for a thermal LWR) loaded into fuel pins and clad in a structural material such as zircalloys. These fuel pins are loaded into numerous structural assemblies, with space between for coolant (H_2O for an LWR) to flow [1(7)].

The BOL fuel loading is more than what is needed for initial criticality in order to compensate for the negative reactivity due to the burnup of fuel and for the production of neutron-absorbing fission products, to compensate the negative reactivity *temperature defect* that is encountered in bringing the reactor (with a negative temperature reactivity feedback) from room to operating temperature, and to compensate the negative reactivity due to the transient buildup of the strongly neutron-absorbing fission product $^{135}Xe_{54}$ within 12 h after the reactor is shut down. Partially offsetting

these negative reactivity effects of burnup is the production of fissile fuel nuclei by the neutron *transmutation* of fertile fuel nuclei [1(6)].

In an LWR, this net excess positive reactivity necessary at BOL must be compensated by strong absorbers of thermal neutrons, such as gadolinium in the form of *burnable poison* pins placed in the fuel assemblies that will be depleted over the operating cycle, of an adjustable concentration of soluble boron poison in the coolant, or of control rods of boron, which may be withdrawn to increase reactivity. By clever design, these stratagems enable a typical LWR to remain critical for a year or two. For fast reactors, this period could be much longer because the production of fissile nuclei from fertile nuclei could more than offset the burnup of fissile nuclei and the production of fission products, owing to the difference in cross sections in fast and thermal neutron spectra. In either case, when it becomes no longer possible to maintain criticality, the reactor must be shut down and refueled.

The buildup of plutonium in a representative LWR is shown in Figure 19.4.

19.3.9 Spent Nuclear Fuel [1(6)]

The spent nuclear fuel (SNF) removed from a nuclear reactor contains (i) a significant

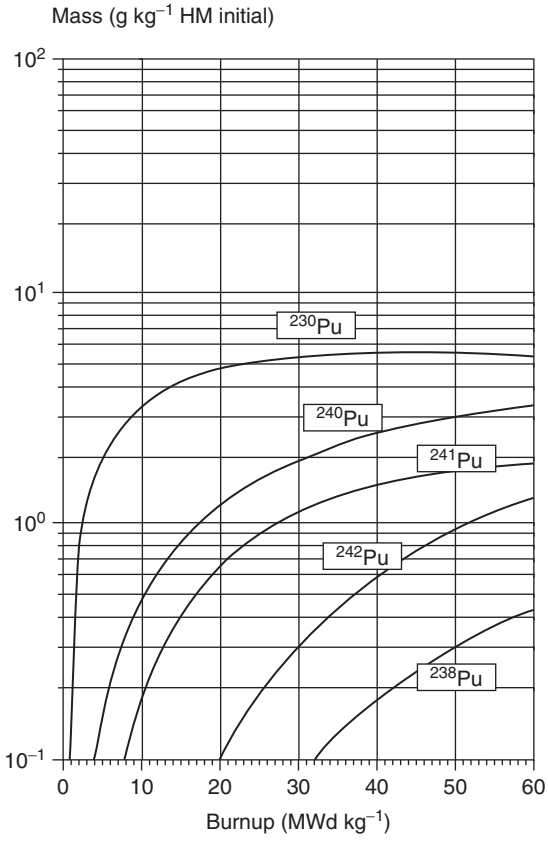
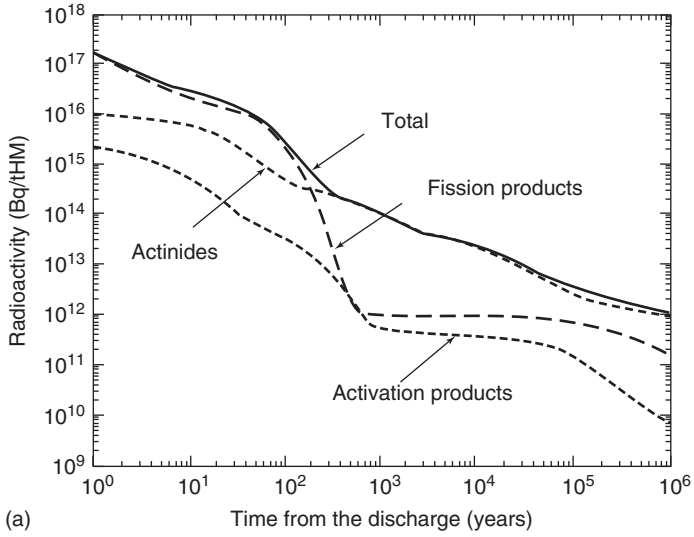


Figure 19.4 Buildup of Pu isotopes in 4 wt% enriched UO_2 in an LWR.



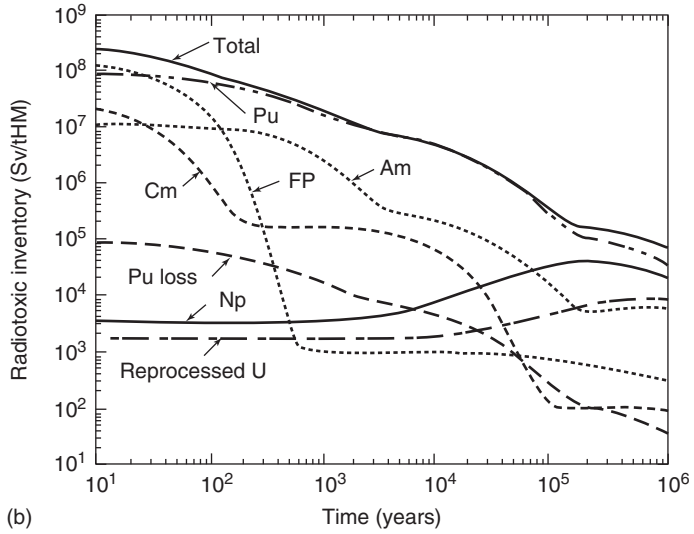


Figure 19.5 (a, b) Radiotoxic inventory of 3.7 wt% enriched UO_2 fuel irradiated to 45 GWd t^{-1} (1 Bq = 1 dps; 1 Sv = 100 rad).

fraction of the original fissile fuel inventory, (ii) the fissile nuclei that have been produced by neutron transmutation of fertile fuel nuclei, (iii) the intermediate-mass fission products that have been produced, and (iv) activated structural material. The calculated radiotoxicity and radioactive inventory of the various components of the SNF from a representative LWR are depicted in Figure 19.5. The short-term radioactivity is dominated by the fission products, but after 300–500 years,

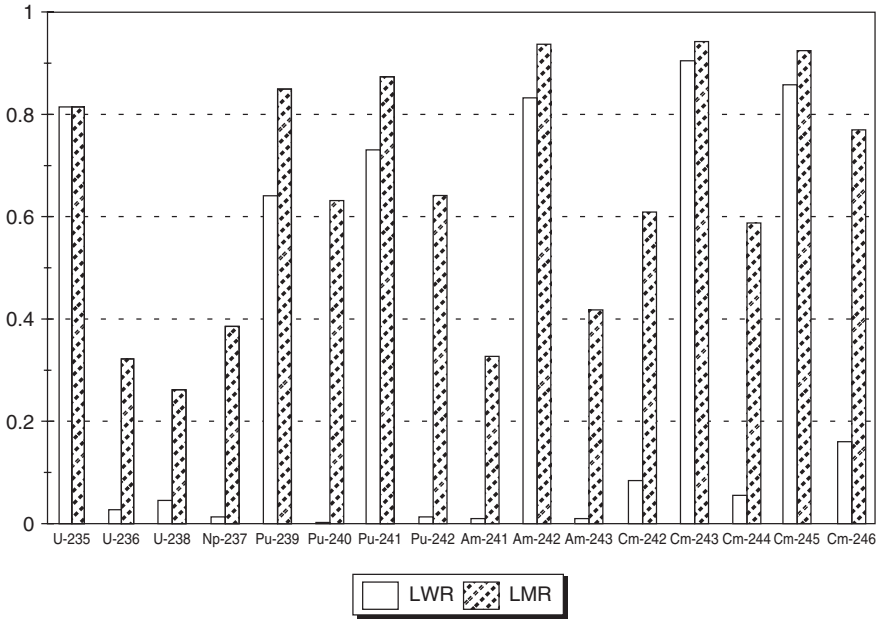


Figure 19.6 Probability of fission per neutron absorbed for actinide isotopes in fast and thermal neutron spectra.

only a few long half-life isotopes (^{99}Tc, ^{129}I, ^{135}Cs, ^{93}Zr, etc.) remain, and the long-term radioactivity is dominated by the TRU actinides (Pu, Np, Am, Cm, etc.) produced from the neutron transmutation of $^{238}U_{92}$. All of these nuclei are fissionable, much more efficiently in a fast neutron spectrum than in a thermal neutron spectrum, as indicated in Figure 19.6. The fast reactor technology necessary to utilize these TRUs in SNF as fuel has been developed but not yet implemented.

19.3.10 Operating Nuclear Power Reactors

As of May 2012, there are 433 nuclear power reactors operating worldwide, producing 371 GWe of electrical power, and 108 more are under construction to produce an additional 104 GWe [6]. The majority of these reactors are LWRs.

19.3.11 Closing the Nuclear Fuel Cycle

Most nuclear power reactors presently operate on the *once-through cycle* (*OTC*), in which the mined uranium fuel is enriched from 0.72% concentration in the fissile isotope $^{235}U_{92}$ in natural uranium to 3-5%, loaded into a reactor and "burned" for about 1.5 years, removed and placed in temporary storage on the reactor site pending permanent disposal. The envisioned permanent disposal has until recently been, and for countries other than the USA still is, permanent burial for hundreds of thousands of years in a *high-level waste repository* (HLWR). This OTC extracts about 1% of the potential energy content in the mined uranium. (Another similarly small fraction of the energy content can be extracted by separating the plutonium in the spent fuel and blending

it with uranium to use as fuel in a similar reactor.) A few more percent of the potential energy content is contained in the fissionable $^{235}U_{92}$ and TRUs remaining in the spent fuel, and the remaining 90+% is in the form of $^{238}U_{92}$, which has been set aside in the mining and enrichment processes or that remains in the spent fuel.

While substantial expansion of nuclear power is certainly the most technically credible option for emission-free electrical power production on the scale needed to meet the world's growing requirements, the accumulation of spent fuel is an immediate psychological, if not technical, impediment to the expansion of nuclear power. On the longer term, the utilization of only 1% of the uranium energy content will impose a fuel limitation later in this century. There is a technical solution to both of these problems – fissioning the TRUs in spent fuel and converting $^{238}U_{92}$ to fissionable TRUs in fast reactors – that has been developed on the laboratory scale [5], but that remains to be implemented on the necessary industrial scale.

19.4 Physics of Fusion Plasmas

19.4.1 Fusion Energy Release [2(1)]

The energy released in fusion reactions also results from an increase in the binding energy per nucleon for the reaction products relative to the reactants, albeit at the opposite end of the binding energy curve of Figure 19.1 from the fission reactions. Just as all the heavy-mass nuclei undergo neutron-induced fission, all the light nuclei undergo fusion if they come close enough together for the strong but short-range attractive strong nuclear force to overcome the longer-range Coulomb repulsive force. This is achieved in nature (the stars) and in terrestrial plasmas by heating the particles to temperatures on the order of 50 000 000 K or higher, and confining them long enough that a sufficient number of pairs of ions approach each other with sufficient relative speed to overcome the Coulomb repulsion and come close enough together to fuse under the attractive nuclear force. Because these particles will be randomly distributed in direction, the repulsive Coulomb force will prevent this from happening in all but a small fraction of the encounters. Averaging the interactions over Maxwellian particle distributions leads to cross sections for Coulomb repulsion winning out (scattering) and for the nuclear attractive force winning out (fusion), shown in Figure 19.7 for a deuterium–tritium (D-T) plasma.

The threshold temperature for overcoming the Coulomb force in a D-T plasma is about 4 keV (4.5×10^7 K), and the reaction products are a 14.1 MeV neutron and a 3.5 MeV α-particle. For a plasma consisting entirely of deuterium, the threshold temperature for fusion is 35 keV (4×10^8 K), and there are two equally likely reactions, one yielding a 2.5 MeV neutron and a 0.75 MeV $A = 3$ α-particle (^{3}He), and the other yielding a tritium ion and a proton, with combined energies of 4.0 MeV. For a plasma of D and ^{3}He, the threshold energy is 30 keV (3.5×10^8 K), and the reaction products are ^{4}He and a proton, with combined energies of 18.2 MeV. The cross sections for these reactions are a few order of magnitude less than the D-T cross section shown in Figure 19.7. There are other exoergic fusion reactions involving nuclei up through the peak of the BE/A curve in Figure 19.1. However, the vast majority of fusion research is focused

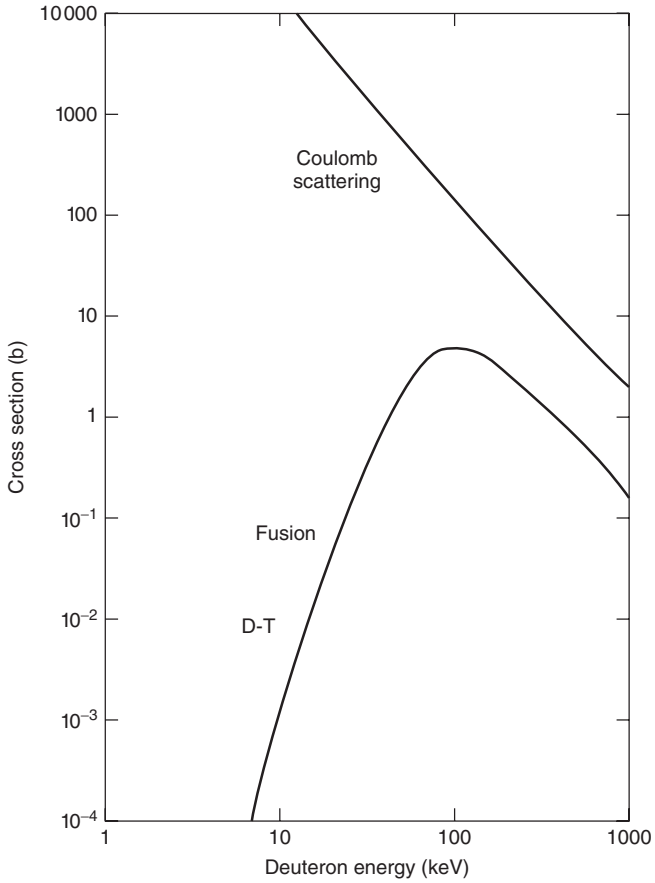


Figure 19.7 Fusion and Coulomb scattering cross sections ($1\,\text{b} = 10^{-24}\,\text{cm}^2$).

on the D-T reaction because of its lower threshold energy and larger cross section for fusion.

19.4.2 Plasma [2(1)]

At temperatures much less than those required for fusion, the atoms of deuterium, tritium, helium, and so on have become fully ionized and the resulting macroscopically neutral mixture of ions and electrons constitutes the fourth state of matter, a plasma. Using Poisson's equation to determine the electrostatic potential in such a mixture and the Maxwell-Boltzmann distribution to determine the distribution of charged particles in such a potential shows that the electrostatic potential from each ion or electron attenuates exponentially $\sim e^{-r/\lambda_D}$ with the *Debeye length* $\lambda_D(\text{m}) = \sqrt{(\varepsilon_0 kT/n_0 e^2)} = 7.43 \times 10^3 \sqrt{T(\text{eV})/n_0(\text{m}^{-3})}$, where n_0, T, e are the average density, temperature, and charge of the ion or electron; k is the Boltzmann constant; and ε_0 is the permittivity of a vacuum. The implication is that the interaction among particles separated by $r > \lambda_D$ must be treated collectively (e.g., in terms of the forces due to the electric

and magnetic fields they produce), and only those interactions at separations, $r < \lambda_D$, can be treated in terms of the Coulomb force acting between particles (i.e., as scattering events). A collection of charged particles with characteristic dimension $L \gg \lambda_D$ is a plasma if there are a large number of particles in a sphere of radius $r = \lambda_D$.

19.4.3 Magnetic Confinement of Plasmas [2(3)]

In extra-terrestrial plasmas, the charged particles are confined by gravitational forces. The size of the sun and stars give some idea of the mass (and size) required for gravitational confinement. As particles at the temperatures necessary for fusion clearly cannot be confined materially without immediately melting the wall, terrestrial confinement research has concentrated on confinement by magnetic, and to a much lesser extent electric, forces.

The Lorentz electromagnetic force acting on a moving charged particle is $\mathbf{F} = e(\mathbf{E} + \upsilon \times \mathbf{B})$. Magnetic confinement is based on the second term, the magnetic force on a moving charged particle in a magnetic field, which is orthogonal to both the direction of motion and the direction of the magnetic field. In a uniform, straight magnetic field, the Lorentz force produces a spiraling motion about the magnetic field lines with a Larmor- or gyroradius $r_L = \upsilon_{th}/|\Omega \equiv |\sqrt{2kT/m}/|eB/m|$. For deuterium ions and electrons at $T = 10\,\text{keV}$, $r_{LD}(m) = 10^{-2}/B(T)$ and $r_{Le}(m) = 3 \times 10^{-4}/B(T)$. For magnetic fields of a few Tesla or more, the deuterons and electrons spiral about the field lines with $r_{LD} < 1$ cm and $r_{Le} < 0.03$ cm. Thus, if field lines can be produced that everywhere remain within the desired confinement volume without intersecting any material wall, the ions and electrons will spiral along these field lines with only small excursions r_L, and the particles will remain within the larger confinement volume with characteristic gradient scale length $L \gg r_L$.

The most obvious way to create such a confining magnetic field configuration is to encircle a toroidal confinement chamber with a set of electromagnets that produce a toroidal field, as depicted in Figure 19.8. This is the basic idea behind the *stellarator* concept pioneered by

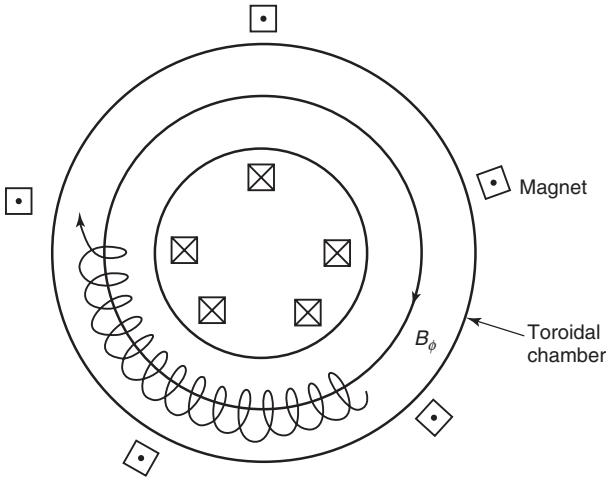


Figure 19.8 Closed toroidal confinement configuration.

Lyman Spitzer at Princeton, of the *tokamak* concept pioneered by Lev Artsimovich at Kurchatov Institute, of the toroidal extension of the *pinch concept* of George Thompson in the UK in the 1950s, and of many subsequent ''closed'' toroidal magnetic confinement concepts.

It goes without saying that things cannot be this simple, and indeed they are not. Such field lines are curved and there is a radial gradient in the magnetic field strength, which in a horizontal toroidal device both produce forces that drive ions and electrons up or down (depending on the current direction in the electromagnets) in opposite directions, thus producing a charge separation and a vertical electric field, which in turn produces a radially outward $\mathbf{E} \times \mathbf{B}$ drift of both ions and electrons onto the outer wall of the confinement chamber. This necessitates adding a poloidal magnetic field to cause the magnetic field lines themselves to spiral about the toroidal chamber in the short (poloidal) direction as well as to circulate around in the long (toroidal) direction, thereby providing a mechanism for distributing particles poloidally before they have time to drift far radially, effectively shorting out the charge buildup caused by the curvature and gradients in the magnetic field.

In the *tokamak* confinement concept, this poloidal magnetic field is produced by placing a solenoidal electromagnet inside the toroidal magnets (in the center of Figure 19.9) to act as the primary of a transformer by providing a time-varying magnetic flux that links the plasma in the confinement chamber (which acts as the secondary of the transformer) to produce a toroidally directed electric field

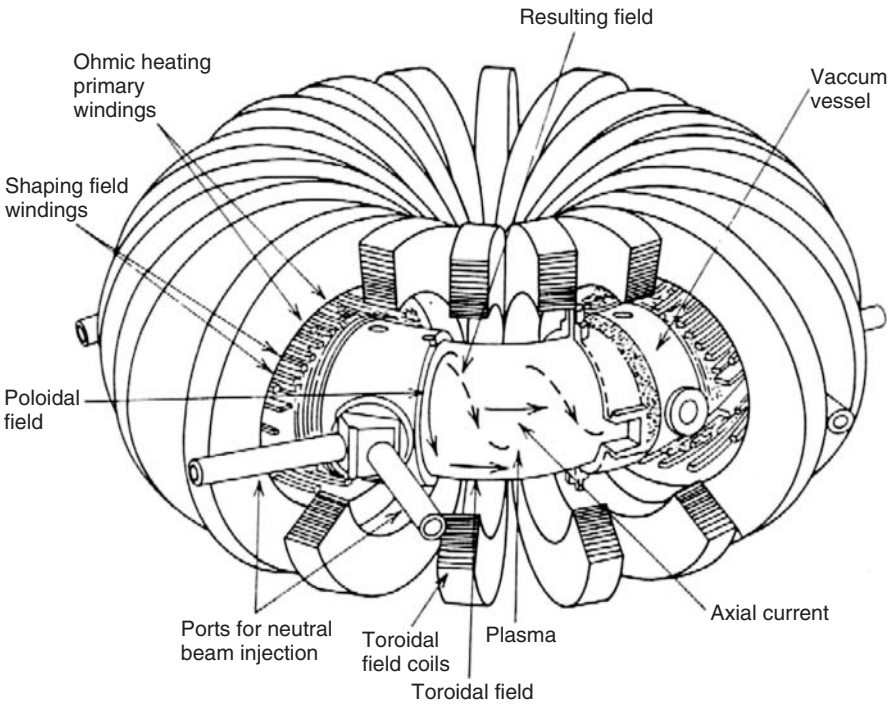


Figure 19.9 Tokamak magnetic confinement concept.

(Faraday's law) around the confinement chamber. This electric field drives a current around the plasma in the toroidal direction that produces a poloidal magnetic field (Ampere's law) in the short direction around the confinement chamber.

In the *stellarator*, the poloidal magnetic field is provided by currents flowing toroidally in electromagnets just outside the confinement chamber. While the stellarator concept is easier to describe, it is harder to implement because of the demanding requirements for precision in the magnetic fields. As a consequence, the tokamak concept has been developed to a much higher degree than the stellarator concept (the $nT\tau_E$ triple product discussed in Section 19.4.6 achieved in a large tokamak is about 40 times larger than the same triple product in a stellarator of similar size), but the stellarator has the potential advantages over the tokamak of inherent steady-state operation and no current flowing in the plasma and hence no disruptions, which cause practical problems in the tokamak, to which we will return.

It is also possible to magnetically confine plasmas in "magnetic well" configurations in which the field lines flow in and out of an "open" confinement volume. Conservation of the action integral over the periodic gyromotion of the particle spiral about the field line requires that the gyro-period average of the magnetic moment $\mu = mv_\perp^2/2B$ is a constant of the motion, and the fact that the magnetic field does no work on the ions and electrons requires that the kinetic energy $KE = m(v_\parallel^2 + v_\perp^2)/2 = \frac{1}{2}mv_\parallel^2 + \mu B$ is a constant of the motion, where $v_\parallel$ and $v_\perp$ are components of the particle velocity along the magnetic field line and in the spiral motion about the magnetic field line, respectively. Clearly, if B increases along the field line, then $v_\parallel$ must decrease, and there will be some particles that will not be able to follow the field lines as they exit the confinement chamber. This is the basic idea of what is known as *mirror confinement*, which has been developed far enough to make a judgment that it is unlikely to lead to an economical fusion reactor.

The production of fusion without confinement is also possible by heating the reactants to temperatures above the threshold for fusion. Such energetic particles would quickly disperse, and the production of a significant amount of fusion energy would require that there be an extraordinarily high initial density of such particles. This, of course, is how thermonuclear weapons work and is the basis of the concept of using lasers or heavy ion beams to compress 1–10 mm D-T pellets to densities of 10^{27}–10^{28} m^{-3} to produce a prodigious fusion energy production rate in the 10–100 ns required for explosive disassembly. This concept, known by the misnomer *inertial* confinement, is also relatively highly developed.

19.4.4 Plasma Heating [2(12)]

Plasmas carrying an internal current, such as the tokamak, are heated resistively, but as the resistivity scales $\sim T_e^{-3/2}$, this heating saturates with increasing temperature. Despite valiant efforts to design a tokamak plasma that could be heated resistively into a thermonuclear range of 8–10 keV, where the fusion self-heating would become sufficient for further heating, this does not seem to be possible; so auxiliary heating methods have been developed. The production and injection into the plasma of 40–100 keV beams of neutral deuterium, which was then converted by atomic collisions to ions that were confined

and collisionally transferred energy to the plasma ions and electrons, was successful in heating plasmas to the thermonuclear temperature range since 1978. This technique has been widely used for plasma heating, with beam energies of 80 to several 100 keV, but is thought not to be practical for future fusion reactors because of the high beam energy requirements for penetration and because of the size and complexity of the technology. Electromagnetic wave heating of the plasma in the ion cyclotron frequency range (0.025–0.1 GHz, radio waves) to the electron cyclotron frequency range (100–200 GHz, microwaves) has been developed subsequently and successfully applied to heat plasmas to the thermonuclear temperature range.

19.4.5 Equilibrium and Stability [2(6,8)]

Once the basic possibility of confinement in a given magnetic configuration is established, the next issue is the nature and stability of the equilibrium force and particle and energy balances [7, 8]. The lowest order force balance in a magnetically confined plasma is the balance between the pressure gradient and Lorentz forces $\nabla \mathbf{p} = \mathbf{j} \times \mathbf{B}$, which indicates that the equilibrium condition is one in which the magnetic field and the current within a plasma both lie in the isobaric surfaces, which are therefore called *flux surfaces*. It can be shown that the largest particle and energy fluxes within a magnetically confined plasma are along the magnetic field lines, that the next largest fluxes are smaller by order $\delta = r_L/L$ (L is a characteristic gradient scale length) also lie within the flux surfaces but orthogonal to the field lines, and that the particle and heat fluxes across the flux surfaces are smaller yet, of order δ^2. This parameter $\delta \ll 1$.

The system consisting of a non-uniform plasma confined in a finite volume by a non-uniform magnetic field configuration (produced mostly by currents in electromagnetics but also by currents in the plasma) is not a system in its lowest free energy state. This implies that rearrangements of the plasma and the magnetic configuration could lead to a lower free energy state for the plasma plus magnetic configuration, thereby making free energy available, which could be converted to plasma kinetic energy to drive plasma motion. The plasma conservation laws and the part of the magnetic field configuration produced by external electromagnets constitute constraints against such releases of free energy, but these constraints are not always sufficient.

Understanding and suppressing such spontaneous conversion of free energy to kinetic energy has been a major focus of magnetic confinement research from the outset, frequently in response to experimental surprises. These instabilities can be grouped into four categories: (i) low-frequency ($\omega(s^{-1}) \leq \Omega_{ion} = e_{ion} B/m_{ion} \simeq 10^8 B(T)$) magnetohydrodynamic (MHD) instabilities that are driven by large-scale displacements or contortions of the plasma and magnetic field, (ii) high-frequency ($\omega(s^{-1}) \geq (n_{ion} Z^2 e^2/\varepsilon_0 m_{ion})^{1/2} \simeq 10^{10}$) electrostatic instabilities driven by the energy released by the relaxation of pressure or density gradients, (iii) magnetic perturbations, and (iv) thermal instabilities.

The most limiting MHD instabilities are *ballooning modes* (which limit the plasma pressure achievable in tokamaks), surface and internal *kink modes* (which set the lower bound on the required magnetic field in tokamaks), *resistive wall modes* (which degrade energy confinement), and *tearing modes* (which produce a

disruptive collapse of the current channel in tokamaks). The *Troyon ballooning mode limit* on the ratio of plasma to magnetic pressures can be expressed as $\beta_{\max}(\%) = \beta_N I(\mathrm{MA})/a(m)B(T)$, where a and I are the plasma minor radius and toroidal current. Present tokamak experiments routinely achieve stable operation up to $\beta_N \approx 2.5-3.0$, and higher values have been reached. Stability against the surface and internal kink modes defines a required relationship between the plasma current and toroidal magnetic field in a tokamak plasma that can be expressed in terms of the *safety factor* $q(r) = r^2 B_\phi/\mu_0 R\int_0^r j_\phi(r')\mathrm{d}r'$ by the requirements $q(r) \geq 1$ everywhere and $q(a)/q(0) \geq 2.5-3.0$.

Magnetic field perturbations can effectively draw energy from the magnetic field produced by currents flowing in the plasma to drive instabilities. The most important of these is the *resistive tearing mode*, which can lead to the disruptive loss of plasma confinement if the toroidal plasma current is increased to the point that the radial location of the $q(r) = 2$ magnetic flux surface moves outward into the steep gradient in the current profile near the plasma edge. Avoidance of such *disruptions* places an upper limit on the plasma current.

The electrostatic instabilities drive various short-wavelength *drift mode fluctuations* that enhance the outward transport of particles and energy, and radial magnetic fluctuations similarly produce an outward transport of particles moving freely along the field lines.

The presence of fast α-particles in a fusion plasma introduces the possibility of destabilization of shear Alfven waves in toroidal geometry. Analysis indicates that there would be a continuum of strongly damped modes and a discrete toroidal Alfven eigenmode that would be driven unstable by particles with velocities near the Alfven velocity, which is $v_A \approx 10^7\,\mathrm{m\,s^{-1}}$ for anticipated fusion plasmas, in which the fusion α-particles are born with slightly greater speed $v_\alpha = 1.3 \times 10^7\,\mathrm{m\,s^{-1}}$.

The strong positive temperature dependence of the fusion cross section in the temperature range of interest introduces the possibility for power balance instabilities, but operating conditions can be chosen so that other stabilizing terms in the energy balance dominate.

19.4.6 Fusion Conditions [2(1,17)]

The energy balance equation in a fusion plasma can be written as

$$P_\Omega + P_\alpha + P_{\mathrm{aux}} - P_{\mathrm{tran}} - P_{\mathrm{rad}} = \frac{\mathrm{d}W}{\mathrm{d}t} \tag{19.4}$$

The first term represents the resistive heating (ηj^2), which is negligible in the thermonuclear temperature range, where the second, fusion α self-heating term $\left(P_\alpha = {}^1\!/_4\, n^2 \langle \sigma v \rangle_f U_\alpha\right)$ is dominant, and $U_\alpha = 3.5$ MeV is the energy of the fusion α-particles that are confined in the plasma. The third term is the auxiliary heating or non-inductive current drive power into the plasma. The fourth and fifth terms represent the power loss due to transport ($3nkT/\tau_E$) and radiation ($n^2 f_z L_z$), where τ_E is the *energy confinement time*, and f_z and L_z are the impurity ion concentration fraction and radiation emissivity due to bremsstrahlung, atomic line and recombination, and cyclotron rotation processes. Ignoring the resistive heating and defining the *plasma energy gain* $Q_p \equiv P_\alpha/P_{\mathrm{aux}}$, the requirement for a steady-state energy balance may be written as a requirement on the triple

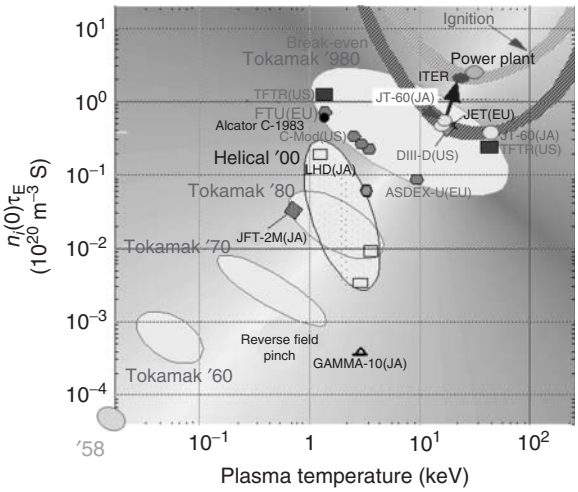


Figure 19.10 Progress in achieving the plasma conditions necessary for fusion.

product

$$nT\tau_E = \frac{1/3}{\left[\frac{1}{4}n^2\langle\sigma\upsilon\rangle_f U_\alpha(1+Q_p^{-1}) - \frac{L_z}{T^2}\right]} \tag{19.5}$$

The right-hand side of this equation is plotted as the thick curves in the upper right in Figure 19.10 for $Q_p = \infty$ (ignition) and for $Q_p = 1$. Progress in achieving this requirement in tokamaks, stellarators (helical devices), pinches, and mirrors (Gambal) are also shown. Clearly the temperatures needed for sustaining fusion have been achieved, and the necessary values of $n\tau_E$ are only a small step away for the tokamak.

For a practical fusion reactor, it is also necessary to achieve high power density $\mathrm{PD(MW\,m^{-3})} = (n^2/4)\langle\sigma\upsilon\rangle_f U_f = (\beta^2 B^4/16\mu_0^2)(\langle\sigma\upsilon\rangle_f/T^2)U_f = (\beta_N^2 I^2 B^2/16a^2\mu_0^2)(\langle\sigma\upsilon\rangle_f/T^2)U_f$ The quantity $\langle\sigma\upsilon\rangle_f/T^2 \sim$ const and maximum over $10 < T < 20$ keV, $U_f = 17.6$ MeV per fusion for D-T, and the Troyon ballooning mode β-limit $\beta \equiv nkT/(B^2/2\mu_0) \le \beta_N I/aB$ has been incorporated to display the dependence of high power density on plasma current, magnetic field, and the achievable β_N.

19.4.7 Energy Confinement and Transport [2(9,11)]

Energy confinement times, which are a measure of the energy loss rate from a plasma, can be determined experimentally from the measured energy content of the plasma and the known rate of replacing energy by external heating. Such confinement times have been measured for a wide range of plasma conditions on a large number of tokamaks worldwide, and have been correlated with plasma size, current, magnetic field, and other parameters. Such correlations do a good job of fitting the available worldwide tokamak database and are used to extrapolate to future tokamaks. In parallel, there is a major research focus on developing first-principle predictions of these energy confinement times.

Although particle and energy flows are much more rapid along the field lines in the isobaric flux surfaces, there are a number of mechanisms that produce particle and energy flows outward across the flux surfaces. Interspecies collisions produce random inward and outward displacements with characteristic dimensions

that are related to the excursion of the particle orbit from the flux surface due to gyromotion and drifts, the well-developed theory of which is referred to as *neoclassical transport* [9]. However, this theory underpredicts the transport rates inferred from experiments considerably for electrons and somewhat for ions, implying the presence of additional transport mechanisms.

The reigning paradigm for a supplemental transport mechanism are the short wavelength fluctuating drift wave instabilities and, to a lesser extent, the radial fluctuations in the magnetic field, which are collectively referred to as *anomalous transport* [10]. Semi-analytic expressions have been developed for diffusive transport coefficients associated with these various drift wave and magnetic fluctuations [11], and more recently numerical simulation of the growth and saturation of such *microinstabilites* and the resulting transport fluxes have been undertaken [12] in realistic models of the plasma magnetic geometry. Although certain modes have emerged as the most likely candidates, they are still work in progress.

In addition to these essentially diffusive transport mechanisms, there are several *non-diffusive transport mechanisms* [2(15)]. Analysis of the momentum transport equation indicates that the "radial" pressure gradient must be balanced by forces associated with the radial electric field and the Lorentz $V \times B$ forces, which gives rise to an expression for the radial particle flux driven by the pressure gradient (diffusive) plus a *pinch* term involving the electric field, the rotation velocities, and any external momentum input (non-diffusive). Moreover, near the plasma edge, some ions have orbits with excursions from the flux surface that cross the last closed flux surface, and particles on these orbit are lost from the plasma in a non-diffusive manner.

As of 2012, first principles prediction of energy transport and confinement remains a major research challenge of the international magnetic fusion program.

19.4.8 Non-Inductive Current Drive [2(12)]

The original tokamak concept of maintaining confinement by driving a plasma current (against resistive losses in the plasma) by transformer action made the tokamak an inherently pulsed device, because continued transformer action requires continued increase in the magnetic field in the transformer primary – the solenoidal electromagnet – and there is a technical limit on magnetic field strength. A repetitively pulsed device would have the disadvantages of thermal cycling, component fatigue, and a low availability for power production. In order to overcome these disadvantages, there has been a major focus in recent tokamak research on mechanisms for noninductively driving the toroidal plasma current.

The same techniques used for plasma heating can also be used for driving current, but with different efficiencies. The mechanism for driving current with neutral beam injection is the collisional momentum exchange of fast ions formed by the injected particles with the plasma ions and electrons. The driving of current by injection of electromagnetic waves is more subtle, but involves the acceleration of plasma electrons by resonant interactions with the wave and the heating of the plasma to increase the temperature and thereby reduce the resistivity.

19.4.9 Edge Plasma Phenomena [2(15)]

While most plasma physics research over the years has focused on the core plasma, there is growing evidence that processes occurring in the plasma edge have a profound influence on plasma performance, and the physics of the plasma edge is today an active area of research [13].

Of particular interest is the observation that a sharp increase in the value of the overall energy confinement time τ_E for tokamak plasmas is correlated with the formation of steep radial density and temperature gradients and a strongly negative radial electric field in the outer ~5% of the plasma. This improved confinement is referred to as *high-confinement mode* (H-mode), as distinguished from the *low-confinement mode* (L-mode) with more gradual edge density and temperature profiles.

However, these steep edge pressure gradients drive a *bootstrap current* in the edge plasma. Sustainment of this H-mode regime is limited by the interplay of MHD ballooning modes, driven by the steep pressure gradient but stabilized by the bootstrap current, and MHD kink modes, driven by the bootstrap current [14]. The resulting edge-localized modes (ELMs) cause the steep edge pressure profile to collapse, with a rapid transfer of the energy and particles in the outer ~5% of the plasma across the last closed magnetic flux surface onto open magnetic field lines that intersect a material surface in the "divertor" chamber, resulting in several percent of the plasma energy being dumped on the divertor target in a sharp burst. ELM control and mitigation is a major current research focus.

Even under normal operation, energetic plasma ions escape the confined plasma across the last closed magnetic flux surface onto open field lines that intersect a material wall at the divertor target. These ions are reflected or re-emitted from the wall as neutral atoms or molecules that diffuse out of the divertor chamber back into the main plasma chamber to refuel the plasma. The energetic ions incident on the divertor target also "sputter" atoms of the divertor target material, which also diffuse out of the divertor chamber and back into the main plasma chamber. Both the neutral plasma atoms and the sputtered "impurity" atoms undergo a variety of interactions (ionization, charge-exchange, line and recombination radiation, etc.) with the plasma ions and electrons in the divertor and in the confined plasma edge, cooling the plasma. The interaction of the incoming neutral plasma atoms and the impurity ions with the edge plasma ions and electrons can lead to *thermal*, or radiative, *instabilities* in the divertor and edge plasma power balances [15], the avoidance of which can set limits on the allowable edge plasma density.

One such instability is the multifaceted asymmetric radiation from edge (MARFE), a radiative condensation instability along the magnetic field lines in the edge plasma that produces a dense, cool, highly radiating, and poloidally localized region. An immediate consequence of MARFE formation is a reduction of plasma confinement from the H-mode level to the lower L-mode level; so the density limit for MARFE formation is a practical density limit for the edge plasma.

The influx into the edge plasma of a sufficient number of highly radiating impurity ions can produce a radial instability in the plasma power balance that results in a rapid collapse in the temperature profile from the edge inward. As the plasma resistivity scales as $\sim T_e^{-3/2}$, the

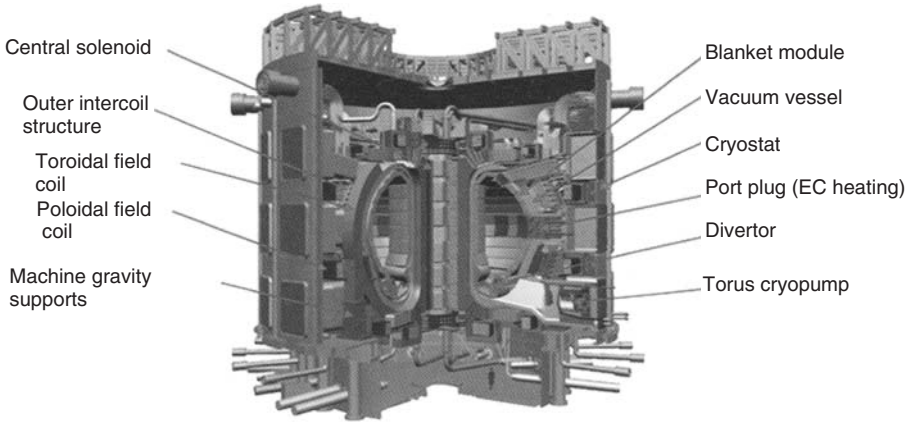


Figure 19.11 ITER design.

temperature profile collapses the current channel and causes the sharp gradient in the radial current profile to move radially inward. When this sharp gradient region reaches the $q(r) = 2$ magnetic flux surface, a $m = 2$ MHD *tearing mode* instability destroys plasma confinement, causing the plasma energy to be dumped on the chamber wall and the plasma current to jump to the chamber wall and surrounding structure [16]. Such an event is known as a *disruption*, avoidance or mitigation of which sets a practical limit on the plasma density.

19.4.10 ITER and Beyond [2(19)]

From 1979 to 1988, fusion scientists and engineers from Europe, Japan, the USA and the USSR joined together in the international tokamak reactor (INTOR) Workshop [17, 18], under IAEA auspices, to evaluate whether the tokamak concept was ready to advance to the experimental power reactor (EPR) stage, to define a design concept and the further R&D required for an EPR, and to identify and analyze the critical technical issues facing such a device. An affirmative conclusion was reached about readiness of the tokamak to move forward to the EPR stage, and at the 1985 Geneva Summit Conference General Secretary Gorbachev proposed to President Reagan that the two countries join together to design, construct, and operate a device based on the INTOR concept. As a result, the ITER (International Thermonuclear Experimental Reactor) project was formed in 1988, involving also Japan and Europe to perform the engineering design and the supporting R&D. Membership in the ITER project has expanded to include also China, India, and South Korea, and construction of the ITER device began in 2009 at the Cadarache site in the south of France (Figure 19.11).

ITER [17] is designed to achieve plasma performance parameters at the threshold of those needed for a practical fusion power reactor (including 500 MWth of fusion power and production of 10 times the amount of fusion power as the power used to maintain the plasma, or $Q_p = 10$) using reactor-relevant fusion technology. It will begin operation in the early 2020s.

After ITER, the path to fusion power is envisioned to include several national demonstration reactors followed by the first ''commercial'' reactors. This development

will be supported by further plasma physics theoretical and experimental development to optimize plasma performance parameters and by fusion technology and materials development and test facilities. Fusion power plants could become available in the second half of the century, depending on the commitment of governments to support their development.

19.5 Fusion–Fission Hybrids

The idea of using fusion neutrons to help close the nuclear fuel cycle (Section 19.3.11) was proposed by Sakharov [19] in the 1950s, and somewhat later by Bethe [20], for transmuting fissionable ^{239}Pu from fertile ^{238}U for fueling nuclear reactors. Although the availability of fissionable fuel will become a problem for the sustainable expansion of nuclear power sometime after mid-century, a more immediate impediment is the accumulation of increasing inventories of radioactive spent fuel at numerous reactor sites, and recent interest in the hybrid concept has focused more on "burning" the radioactive TRUs as fissionable fuel in *fusion–fission hybrid* (FFH) *burner reactors* [21].

A FFH is basically a nuclear reactor operated subcritical, with a D-T fusion neutron source. An FFH burner reactor based on the most highly developed fast reactor and (magnetic) fusion technologies [21] would consist of an annular metal-fuel, sodium-cooled fast reactor (e.g., of the IFR (integral fast reactor) type [5]), fueled with the TRU from SNF, surrounding a tokamak neutron source of the ITER type [17] but scaled down by two-third.

Such a FFH would have two advantages relative to a critical fast burner reactor. The same fuel could remain in the FFH to higher burnup than in a critical fast burner reactor because the reactivity loss due to fuel burnup and fission product buildup could be compensated by changing the fusion neutron source level to achieve deep burnup to the radiation damage limit in the FFH, whereas the requirement to maintain criticality would almost certainly result in less burnup before fuel removal for reprocessing in a critical reactor. Thus, an FFH burner reactor would require fewer recycling steps than a critical reactor of the same type to process a given batch of fuel to a given burnup. The amount of TRU that would eventually end up in a HLWR depends on the number of reprocessing steps (and the separation efficiency). Thus, relative to a critical fast burner reactor, the FFH would require fewer HLWRs and fewer fuel reprocessing and re-fabrication plants.

Because the *reactivity margin to prompt critical* is equal to the delayed neutron fraction β in a critical reactor (Section 19.3.6) and because β is about three times smaller for TRU fuel than for U fuel, it is probable that a critical burner reactor would be fueled with 20–25% TRU fuel and 75–80% uranium fuel in order to retain a reactivity margin of safety that was comparable to a U-fueled critical reactor. For an FFH, the reactivity margin to prompt critical is the degree of subcriticality, which would be orders of magnitude larger than β. So, an FFH burner reactor could safely be loaded with 100% TRU fuel, and four to five times fewer FFH than critical burner reactors would be needed to burn a given inventory of spent fuel.

Such a FFH would also have two disadvantages relative to a critical IFR burner reactor. It would be more complex and it would be more expensive. However, the real issue is the overall cost and complexity

of the entire system – burner reactors, fuel reprocessing and refabrication facilities, and HLWRs – and this has not been worked out yet.

As the fast reactor and fuel reprocessing and re-fabrication has already been demonstrated, albeit some only on a laboratory scale, and the ITER fusion physics and technology has been demonstrated on a laboratory scale and will be demonstrated on a reactor scale in the 2020s, a FFH burner reactor could be online in the second quarter of the present century.

Glossary

Afterheat: The energy released by the subsequent radioactive decay of unstable nuclei produced in the fission of heavy-mass nuclei.

Anomalous Transport: Transport of particles, momentum, and energy in a plasma due to fluctuations caused by microinstabilities.

Ballooning Mode: An MHD instability that causes the outboard part of the toroidal plasma surface to balloon outward (as same as a rubber tube would do if over-inflated).

Binding Energy: The energy equivalent of the mass defect between the sum of the masses of the nucleons in an atomic nucleus and the mass of the nucleus.

Bootstrap Current: The current driven in plasmas by plasma pressure gradients.

Breeder Reactor: A nuclear reactor in which more fissile nuclei are produced by the neutron transmutation of $^{238}U_{92}$ (or $^{232}Th_{90}$) than are fissioned.

Burnable Poison: A solid neutron-absorbing material that is put in a reactor for the purpose of controlling the reactivity due to fuel burnup.

Burner Reactor: A reactor fueled in total or in part by the fissionable transuranics in the spent fuel discharged from nuclear power reactors.

Critical: A condition in which the multiplication constant $k = 1$ in a nuclear reactor and the neutron population will remain constant in time.

Debeye Length: The very short distance over which the electrostatic potential from a point charge is attenuated by e^{-1} due to the many other charged particles present in a plasma.

Delayed Neutron: Neutron released several seconds after the fission event by the radioactive decay of the fission products.

Delayed Neutron Precursor: An unstable fission product that releases a delayed neutron at the time of decay.

Disruption: Massive loss of plasma confinement caused by MHD tearing modes.

Doppler Broadening: The increase with temperature of the range of neutron energies for which the formation of an excited state of a compound nucleus is energetically possible.

Drift Mode Fluctuations: Microinstabilities driven by the free energy liberated by the relaxation of thermodynamic gradients within the plasma.

ELM: Edge localized mode.

Energy Confinement Time (τ_E): The mean time required for energy to be lost from a confined plasma.

Fast Reactor: A nuclear reactor in which the majority of the neutrons have energies $E > 10\,\text{keV}$.

Feedback Reactivity: Reactivity produced by a change in the temperature of a reactor material.

Fertile Nucleus: A nucleus that can be transmuted into a fissile nucleus by the capture of a neutron.

Fissile Nucleus: A nucleus that has a large probability of fissioning when it absorbs a thermal neutron.

Fission Products: The intermediate-mass nuclei produced by the fission of a heavy-mass nucleus.

Flux Surfaces: Isobaric surfaces in the plasma on which both the magnetic fields line and the plasma currents lie.

Fuel Utilization Factor: The ratio of the rate that neutrons are absorbed in the fissile fuel to the total rate at which neutrons are absorbed by all materials present or leak from the reactor.

Fusion–Fission Hybrid (FFH): A subcritical reactor driven by a fusion neutron source.

High-Confinement Mode (H-Mode): A mode of tokamak plasma operation with large energy confinement time.

High Level Waste Repository (HLWR): A permanent, secured storage facility for high-level radioactive waste.

IFR: Integral fast reactor, a metal-fuel, sodium-cooled fast reactor with onsite fuel reprocessing and re-fabrication.

Importance: (i) Related to the likelihood that a neutron introduced at a certain location in a reactor will subsequently result in a fission. (ii) Mathematically, the adjoint function.

ITER: A 500 MWth tokamak experimental power reactor being constructed internationally by Europe, Japan, USA, Russia, China, India, and South Korea for operation in the 2020s.

Kink Modes: MHD instabilities that produce a twisting distortion of the magnetic flux surfaces.

Light Water Reactors (LWRs): The majority of nuclear power reactors operating in the world are either pressurized water reactors (PWRs) or boiling water reactors (BWRs), both of which have a thermal neutron spectrum.

Low-Confinement Mode (L-Mode): A mode of tokamak plasma operation with energy confinement time about a factor of 2 less than H-mode.

MARFE: Multifaceted asymmetric radiation from edge.

MHD: Magnetohydrodynamic.

Microinstabilities: Small-scale instabilities that produce fluctuations in the plasma density, temperature and electrostatic potential.

Mirror: A class of magnetic confinement devices in which the plasma is confined in magnetic wells by the requirements of conservation of magnetic moment and kinetic energy.

Moderating Ratio: The product of the average logarithmic energy loss in an elastic scattering event and the ratio of the elastic scattering and absorption probabilities.

Multiplication Constant: The average number of fission neutrons produced in the next generation by the ν neutrons produced in a fission event.

Neoclassical Transport: Transport of particles, momentum, and energy in a toroidal plasma due to collisional processes.

Non-Diffusive Transport: Transport of particles, momentum, and energy in a plasma due to Lorentz and other forces not related to gradients of the thermodynamic variables.

Non-Leakage Probability: The ratio of the rate at which neutrons are absorbed in some material in a reactor to the rate at

which neutrons are either absorbed or leak from the reactor.

Nuclear Engineering: The applied nuclear physics discipline concerned with the controlled production, recovery, and conversion to electricity of nuclear energy released in the fission or fusion of atomic nuclei.

Once-Through-Cycle (OTC): A nuclear reactor fuel cycle in which the fuel is removed from the reactor after one residence time, stored temporarily to allow decay of short-life fission products, and then stored in a permanent disposal facility (HLWR).

Plasma Energy Gain Q_p: Ratio of the fusion power production rate to the auxiliary heating and current-drive power injected into the plasma.

Pinch: (i) A magnetic confinement concept based on the attractive force between parallel currents. (ii) The part of the plasma particle flux associated with non-diffusive forces.

Prompt Jump: The very rapid change in the neutron population in a reactor following the introduction of reactivity.

Reactivity: The difference in neutron multiplication (k) between the state of a nuclear reactor and the multiplication ($k = 1$) that would maintain a constant neutron population in time

Reactivity Margin to Prompt Critical: The maximum amount of positive reactivity that could be added to a nuclear reactor without producing a runaway power excursion.

Safety Factor: A plasma parameter proportional to the ratio of toroidal magnetic field to the enclosed toroidal current.

Stellarator: A class of toroidal confinement configurations in which the confining magnetic field is produced entirely by currents in external electromagnets.

Subcritical: A condition in which the multiplication constant $k < 1$ in a nuclear reactor and the neutron population will decay in time unless supplemented by an external neutron source.

Supercritical: A condition in which the multiplication constant $k > 1$ in a nuclear reactor and the neutron population will increase in time.

Tearing Mode: A MHD instability that causes a disruptive loss of plasma confinement.

Temperature Defect: The negative reactivity introduced by raising the reactor temperature from room temperature to operating temperature.

Thermal Instabilities: Plasma power balance instabilities caused by recycling neutrals and impurity ions in the plasma edge and by the strong temperature dependence of the fusion cross section in the plasma core.

Thermal Reactor: A nuclear reactor in which the majority of the neutrons have energy $E < 1$ eV.

Tokamak: A class of toroidal confinement configurations in which the confining magnetic field is produced by currents in external electromagnets and by currents in the plasma.

Transmutation: The change of a nuclide of mass A into one of mass $A + 1$ by neutron capture.

Transuranics: The nuclides above uranium in the periodic table.

Troyon Ballooning Mode Limit: A theoretically based empirical relation for the upper limit on the ratio of the plasma pressure to the magnetic pressure set by MHD ballooning modes.

References

1. Stacey, W.M. (2007) *Nuclear Reactor Physics*, 2nd edn, Wiley-VCH Verlag GmbH, Weinheim.
2. Stacey, W.M. (2012) *Fusion Plasma Physics*, 2nd edn, Wiley-VCH Verlag GmbH, Weinheim.
3. Thompson, T.J. and Beckerley, J.G. (eds) (1964) *The Technology of Nuclear Reactor Safety*, Vol. I, MIT Press, Cambridge.
4. Hummel, H.H. and Okrent, D. (1970) *Reactivity Coefficients in Large Fast Power Reactors*, American Nuclear Society, LaGrange Park.
5. Charles, E. (2011) *Till, Charles E. and Chang, Yoon I., "Plentiful Energy: The Story of the Integral Fast Reactor"*, CreateSpace, Charleston.
6. "World List of Nuclear Power Plants", Nuclear News, 47 (2013).
7. Freidberg, J.P. (1987) *Ideal Magnetohydrodynamics*, Plenum Press, New York.
8. Mikhailovskii, A.B. (1992) *Electromagnetic Instabilities in Inhomogeneous Plasma*, IOP Press, Bristol.
9. Helander, P. and Sigmar, D.J. (2002) *Collisional Transport in Magnetized Plasmas*, Cambridge University Press, Cambridge.
10. Itoh, K., Itoh, S.-I., and Fukiyama, A. (1999) *Transport and Structure Formation in Plasmas*, IOP Press, Bristol.
11. Weiland, J. (2000) *Collective Modes in Inhomogeneous Plasma*, IOP Press, Bristol.
12. Beer, M.A. and Hammett, G.W. (1996) Toroidal gyrofluid equations for simulations of tokamak turbulence. *Phys. Plasmas*, **3**, 4046.
13. Stangeby, P.C. (2002) *The Plasma Boundary of Magnetic Fusion Devices*, IOP Press, Bristol.
14. Snyder, P.B., Wilson, H.R., Ferron, J.R., Lao, L.L., Leonard, A.W., Osborne, T.H., Turnbull, A.D., and Xu, X.Q. (2004) *Nucl. Fusion*, **44**, 320.
15. Stacey, W.M. (2007) A survey of thermal instabilities in tokamak plasmas: theory, comparison with experiment, and predictions for future devices. *Fusion Sci. Technol.*, **52**, 29.
16. Bondeson, A., Parker, R.D., Hugon, M., and Smuelders, P. (1991) MHD modelling of density limit disruptions in tokamaks. *Nucl. Fusion*, **31**, 1695.
17. iter *www.iter.org* (accessed 14 February 2013).
18. Stacey, W.M. (2010) *The Quest for a Fusion Energy Reactor*, Oxford University Press, New York.
19. A. Sakharov, *Memoirs*, (trans. R. Lourie), Vintage Books (1992), p. 143.
20. Bethe, H.A. (1978) "The Fusion Hybrid" Nuclear News 21, 41.
21. Stacey, W. M. (2011) Principles and rationale of the fusion–fission hybrid burner reactor. Proceeding FUNFI-11, Varenna, Italy, 2011 (AIP-2012).

Further Readings

Fission

Bell, G.I. and Glasstone, S. (1970) *Nuclear Reactor Theory*, Van Nostrand Reinhold, New York.

Duderstadt, J.J. and Hamilton, L.J. (1976) *Nuclear Reactor Analysis*, John Wiley & Sons, Inc., New York.

Glasstone, S. and Edlund, M.C. (1952) *The Elements of Nuclear Reactor Physics*, D. Van Nostrand, Princeton.

Henry, A.F. (1975) *Nuclear Reactor Analysis*, MIT Press, Cambridge.

Meghreblian, R.V. and Holmes, D.K. (1960) *Reactor Analysis*, McGraw-Hill, New York.

Radkowsky, A. (ed) (1964) *Naval Reactors Physics Handbook*, U.S. Government Printing Office, Washington, DC.

Templin, L.J. (ed) (1961) *Reactor Physics Constants*, 2nd edn ANL-5800, U.S. Government Printing Office, Washington, DC.

Weinberg, A.M. and Wigner, E.P. (1958) *The Physical Theory of Neutron Chain Reactors*, University Chicago Press, Chicago.

Fusion

Artsimovich, L.A. (1964) *Controlled Thermonuclear Reactions*, Gordon and Breach, New York(1961 in Russian).

Braginksii, S.I. (1965) in *Transport Processes in a Plasma*, Reviews of Plasma Physics, Vol. 1 (ed. M.A. Leontovich), Consultants Bureau, New York, p. 205(1963 in Russian).

Chandrasekhar, S. (1960) *Plasma Physics*, University Chicago Press, Chicago.

Chen, F.F. (1984) *Introduction to Plasma Physics*, 2nd edn, Plenum Press, London.

Freidberg, J.P. (2008) *Plasma Physics and Fusion Energy*, Cambridge University Press, Cambridge.

Glasstone, S. and Lovberg, R.H. (1960) *Controlled Thermonuclear Reactions*, D. Van Nostrand, New York.

Goldston, R.J. and Rutherford, P.H. (1995) *Introduction to Plasma Physics*, IOP Press, Bristol.

Hazeltine, R.D. and Meiss, J.D. (1992) *Plasma Confinement*, Addison-Wesley, New York.

Hutchinson, I.H. (1988) *Principles of Plasma Diagnostics*, Cambridge University Press, Cambridge.

Kadomtsev, B.B. (1992) *Tokamak Plasmas: A Complex Physical System*, IOP Press, Bristol.

Krall, N.A. and Trivelpiece, A.W. (1973) *Principles of Plasma Physics*, McGraw-Hill, New York.

McCracken, G.M. and Stott, P. (2005) *Fusion: The Energy of the Universe*, Elsevier and Academic Press, London.

Miyamoto, K. (1989) *Plasma Physics for Nuclear Fusion*, MIT Press, Cambridge(1976 in Japanese).

Rose, D.J. and Clark, M.J. (1961) *Plasmas and Controlled Fusion*, MIT Press, Cambridge.

Spitzer, L. (1956) *Physics of Fully Ionized Gases*, Interscience Publishers, New York.

Stix, T.H. (1992) *Waves in Plasmas*, AIP Press, New York.

Stacey, W.M. (2010) *FUSION: An Introduction to the Physics and Technology of Magnetic Confinement Fusion*, 2nd edn, Wiley-VCH Verlag GmbH, Weinheim.

Wesson, J. (ed) (2011) *Tokamaks*, 4th edn, Oxford University Press, Oxford.

20
Fundamentals of Controlled Nuclear Fission and Essential Characteristics of Pressurized-Water Reactors

Hartmut Lauer

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

20.1
Introduction

About 440 nuclear power plants are currently onstream producing power worldwide. Their total output fed into the grid amounts to about 370 000 MW. Judging by sheer numbers, nuclear energy contributes about 13% to global power production.

More than 300 units, that is, two-thirds of all nuclear power plant units in the world, are located in the United States, Europe, and Russia. In the densely populated and energy-hungry nations of China and India, nuclear energy so far only contributes 3% to power production. However, Asia is regarded as the emerging continent for nuclear power plant new-builds in the years ahead.

Today, and maybe also in the future, nuclear energy plays a role in industrial nations or transformation countries that, with their relatively high population density and the associated high specific energy demand, energy-intensive industry, and the appropriate infrastructure (including grids), provide the environment in which nuclear power plant units with large capacities above 1000 MW can be operated in a technically and economically viable manner compared with other energy sources. The supply countries for uranium are largely politically stable so that risks for secure and long-term fuel supplies are considerably low. Uranium, which is relatively low-priced and virtually without value for any other technical applications, is used as fuel today to meet a power demand that would otherwise have to be covered largely with fossil fuels. As a result, nuclear energy conserves resources and reduces the demand especially for coal and gas. This is of major economic and ecological importance also on a global basis [1].

The nuclear energy policy of some nations has changed since the reactor accident in the Japanese Fukushima Daiichi Nuclear Power Plant in March 2011, while in other countries, nuclear energy has kept its role as a safe, reliable, and cost-effective source of low-carbon power production.

Most countries with nuclear power plants initiated short-term reviews and measures. Rules and regulations may be tightened in several countries in the long run. New nuclear energy technologies could help allay misgivings in respect of radioactive waste and plant safety.

20.2
Reactor Developments

Nuclear power plants fall into four generations: the first generation of the 1950s comprises the early prototype and

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

demonstration power plants with typically less than 100 MW of electrical capacity. They demonstrated the technical feasibility; economic aspects were of secondary interest at first. Various technology lines of generation II emerged from the prototype reactors of the first generation (generation I), which were then commercially pursued further. Plant costs were reduced and economic efficiencies increased through standardization and larger unit capacities (up to over 1000 MW electric). Only few technologies held their own on a broad basis; the light-water reactors (LWRs), either as boiling-water reactors (BWRs) or as pressurized-water reactors (PWRs), are the dominating technology among the plants in operation. Heavy-water reactors (e.g., CANDU, CANadian Deuterium Uranium) and gas-cooled reactors (GCRs) (e.g., AGR, advanced gas-cooled reactor) were also built in individual countries.

The operating lifetime of many of these plants of generation II can be significantly extended to typically 60 years with careful maintenance and regular retrofits and modernization.

This has gone hand in hand with the development of new, advanced nuclear power plant types of the third generation since the 1990s. They have reached commercial maturity today and provide the basis for the new-builds in the years ahead.

The development of nuclear systems of the fourth generation (generation IV), which could be required in the second half of the twenty-first century or even later, is currently in the focus of international research activities. The development targets of these efforts go to some extent beyond those of generation III (e.g., use for process heat and hydrogen production) (Figure 20.1).

20.3 Principle of Controlled Nuclear Fission, Chain Reaction

It is apparent and can be theoretically explained in nuclear physics that isotopes

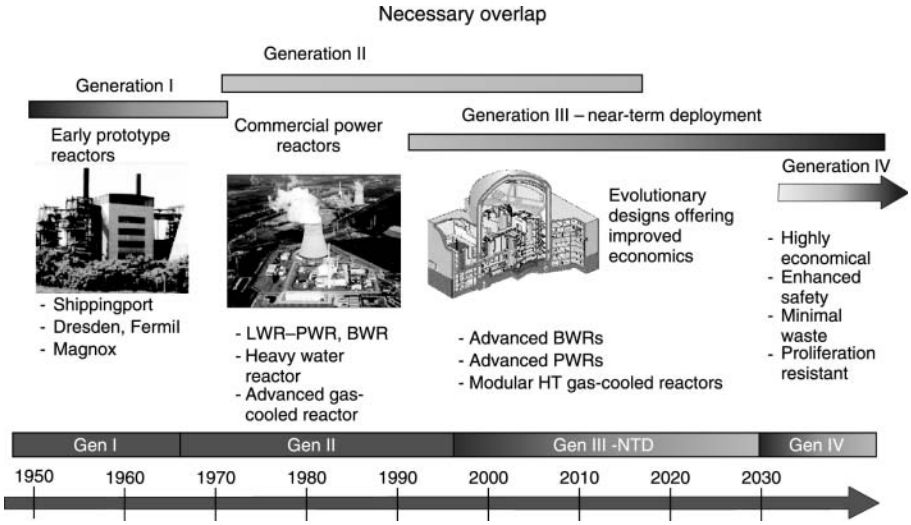


Figure 20.1 Timeline of reactors' developments. (Source: RWE Power.)

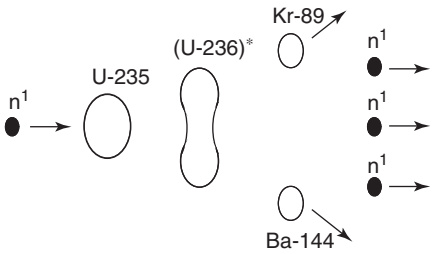


Figure 20.2 Energy from the U-235 atom.

with an uneven mass number have favorable and those with an even mass number have unfavorable fission properties. The isotopes uranium-233, uranium-235, plutonium-239, and plutonium-241 are readily fissile materials as a result of slow (the so-called thermal) neutrons. The nuclear fission mechanism can be described in simplified terms by the so-called droplet model (Figure 20.2) [2–5].

The U-235 target nucleus absorbs a slow (thermal) neutron and creates a highly excited intermediate nucleus of the U-236 isotope. The new nucleus tries to release its excitation energy. The surplus energy is released through either γ-radiation or particle emission, and in most cases, it initiates a fission of the atomic nucleus. In a model, the fission can be described as follows: once the neutron has been captured, the uranium nucleus starts to oscillate, is deformed elliptically, is stretched into a dumbbell shape with a narrow waist, and finally, disintegrates into two debris nuclei of medium weight (Kr-89 and Ba-144 in Figure 20.2) as well as two to three neutrons. Only the last fact enables the chain reaction to happen in which the released neutrons on their part are again available for further fission processes. As, on average, about two to three prompt neutrons arise with every fission, the fission rate increases permanently without interference and renewed fission processes can be triggered and so forth (chain reaction). About 120 trillion (120×10^{18}) nuclei have to be split per second to generate an electrical power of 1300 MW in a nuclear reactor.

The fission process is accompanied by gamma emission. The fission of a uranium-235 nucleus can result in various debris nuclei (fission products). Today, we know approximately 200 different fission products of uranium-235 that relate to 35 different elements. Here, the sum of the atomic numbers of the fission products has to equal the atomic number of uranium, that is, in the case of uranium-235, the atomic numbers of the debris nuclei and the released neutrons always amount to 236. In general, the fission products are not equally large; the mass numbers are between 70 and 160, with the maximum at 95 and the second maximum at 140.

The energy spectrum of the fission neutrons of U-235 extends across a wide range with the main portion being accounted for, however, almost exclusively by fast neutrons ($E > 0.1$ MeV). Over 99% of the arising neutrons are emitted immediately/promptly (10^{-17} s) after fission. The emission is from both heavy and light fragments. The likelihood depends on the level of excitation energy and the neutron surplus in the fragments.

Apart from the neutrons arising directly during fission (the so-called prompt neutrons), the so-called delayed neutrons are also released during the radioactive decay of the fission products, that is, with a delay of up to 1 min. This is because despite

the emitted prompt neutrons, the fission fragments moderated in the fuel still have a neutron/proton ratio in excess of what would be equivalent to a stable core configuration. As a result, the fission products carry out beta decays to get into a stable condition. Beta decays can result in daughter nuclei with a very high level of excitation so that the emission of the neutron becomes possible in energy terms (as process competing with γ-radiation). Overall, this process gives the impression as if the neutron was emitted with the half-life of the conversion of the respective nucleus (for example, Br-87, 54,5s). The mean life of these delayed neutrons (time from fission to absorption in the fuel) is 12.7 s for U-235. The mean life of the prompt neutrons (time from fission to absorption in the fuel), however, is only 2.5×10^{-5} s for the PWR and 5.0×10^{-5} s for the BWR. This results in a mean life of ~0.1 s for U-235 for all neutrons (prompt and moderated). The life of the entire neutron population in the reactor is therefore decisively determined by the moderated neutrons, which enables the reactor to be controlled in the first place. For U-235, the β-share of moderated neutrons of the total number of the newly created fission neutrons amounts to 0.65%.

On account of the neutron surplus, fission products are in a large part radioactive and are typically transformed into stable nuclei while emitting β-rays. In this process, they go through the entire decay series.

If the mass of the nuclear particles of uranium-235 and the primary neutron is compared with the total of the masses of the fission products and the secondary neutrons, a minor loss of mass is identified after nuclear fission. This loss is equivalent to the energy released during fission. This is because the nucleons (protons and neutrons) in an atomic nucleus are bound with certain energy per nucleon, which is the so-called bonding energy. For the heavy atomic nuclei, for example, uranium-235, the bonding energy per nucleon is less than that for the medium atomic nuclei with mass numbers between 80 and 150. When splitting a uranium atomic nucleus into two atomic nuclei of medium mass number, the bonding energy therefore increases, which results in energy being released to the outside.

When splitting a U-235 nucleus, a total energy of approximately 210 MeV is released, which can be roughly divided as follows:

- Kinetic energy of fission products: 175 MeV
- Kinetic energy of fission neutrons: 5 MeV
- Energy of gamma radiation occurring immediately after fission: 7 MeV
- Energy of beta and gamma radiation during the decay of the radioactive fission products: 13 MeV
- Energy of the antineutrinos: 10 MeV.

Approximately 90%, that is, about 190 MeV, of this can be used in the nuclear reactor. The bulk of this fission energy (approximately 80%) occurs as kinetic energy of the fission products. They interact with the ambient matter (fuel) and release their energy after approximately 10^{-3} cm along their path. This energy portion therefore remains completely left in the reactor. The rest of the energy is spread across neutrons, gamma quanta, beta particles, and the antineutrinos created during beta decay. As charged particles, the β particles can only cover short distances in the reactor and therefore release their entire energy in the reactor. The same goes more or less for the gamma quanta because the

reactors are designed in such a way that only a small portion of the gamma quanta can leak out. However, the energy of the antineutrinos can virtually not be used at all in the reactor. Owing to the extremely small interaction with matter, the antineutrinos leave the reactor unimpeded.

20.4 Basic Principle of a Nuclear Reactor

Plants with controlled chain reactions are called *nuclear reactors*, or *reactors* for short. In extremely simplified terms, they consist of five components [6]:

- sufficient mass of fissile material;
- material to moderate the neutrons, an exception is the fast breeder reactor (FBR);
- systems to capture neutrons (control rods);
- heat removal medium;
- barriers for radiation protection and the retention of radioactive materials.

Controlled nuclear fission requires a sufficient number of neutrons to be available. A certain mass (critical mass) is necessary for a chain reaction to happen in a pure fissile material. A nuclear reactor does not provide a critical mass but a critical configuration of fuel, moderator, and reflector. In an LWR of the >1000 MWe capacity class, the nuclear fuel volume totals over 100 Mg with an enrichment of 3–5% of fissile U-235.

20.4.1 Multiplication Factor and Four-Factor Formula

An important measure for the chain reaction is the ratio of the individual fission neutron generations to each other. Fission neutron generation stands for all the free neutrons existing in a reactor at a given point in time. The number of the new free neutrons created by these neutrons by fission is then the next generation. If in any generation, more (fewer) neutrons are created and disappear, the reactor is called *supercritical (subcritical)*. The ratio of the succeeding neutron generations is called the *multiplication factor k*.

$$k = \frac{\text{Number of fissions (or fission neutrons) in one generation}}{\text{Number of fissions (or fission neutrons) in the preceding generation}}$$

For an infinitely extended reactor medium, the multiplication factor (k_∞) is calculated with the so-called four-factor formula:

$$k_\infty = \varepsilon \times p \times f \times \eta$$

The fast fission factor ε takes into account that the fast neutrons created by moderated neutrons during the fission of U-235 also contribute to fission, especially in U-238, with the share ($\varepsilon \approx 1.05$).

$$\varepsilon = \frac{\text{Production rate of all fast neutrons}}{\text{Production rate of fast neutrons from thermal fission processes}}$$

The resonance escape probability p accounts for the likelihood with which the fast neutrons are not captured ($p < 1$) during their moderation process and while passing the absorption resonance area (5 eV < E < 100 eV) of U-238, and so on.

$$p = \text{Fraction of neutrons not absorbed during moderation}$$

The thermal utilization f defines the share of thermal neutrons absorbed ($f < 1$) by the fuel during their diffusion in the core.

$$f = \frac{\text{Absorption rate of thermal neutrons in the fuel}}{\text{Total absorption rate of thermal neutrons}}$$

The neutron yield η stands for the portion of the thermal neutrons absorbed in the fuel, which trigger fission and regenerate fast neutrons. The chain reaction is only possible if $\eta > 1$.

$$\eta = \frac{\text{Number of fast neutrons created by fission processes}}{\text{Number of thermal neutrons absorbed by the fuel}}$$

In a thermal reactor, the variables ε and η are essentially determined by the type of nuclear fuel and the composition of the reactor materials. The quantities p and f are to a certain extent variable through the fuel rod array (fuel rod spacing and diameter and moderator and/or coolant share). Both quantities should be as large as possible. It is important to note, however, that normally an increase in p results in a decrease in f and vice versa.

In a finite reactor, the neutrons can leak out and be lost for the chain reaction. The likelihood that this does not happen is

- Λ, the nonescape probability, is sometimes broken down further into one for fast and one for thermal neutrons. It depends on the shape and size of the core; $\Lambda < 1$ always applies to a finite core volume. Out of k new neutrons created, $k \times \Lambda$ remain in the core and $k(\Lambda - 1)$ that leave the core are lost for the chain reaction.

k_{eff} is the effective multiplication factor of the reactor, according to which the core constitutes a finite volume with boundaries through which neutrons can be lost. It can be expressed as

$$k_{\text{eff}} = k_{\infty}\Lambda$$

The following equation applies

$$k_{\text{eff}} = \frac{\text{Production rate of neutrons}}{\text{Absorption rate} + \text{Leakage rate of neutrons}}$$

$k_{\text{eff}} = 1$	The number of neutrons is constant; the reactor is critical.
$k_{\text{eff}} > 1$	The number of neutrons is growing; the reactor is supercritical.
$k_{\text{eff}} < 1$	The number of neutrons is decreasing; the reactor is subcritical.

Hence, a reactor is critical if as many neutrons are created by nuclear fission as are lost by absorption in the fuel and structure material and leakage. The critical condition is the normal operating condition of a reactor during power operation.

The following holds true for a moderated supercritical reactor operated with uranium-235: $1 < k < 1.0065$ (the β share of moderated neutrons is 0.65%). At $k > 1.0065$, the reactor is promptly supercritical.

The term *reactivity* is also used instead of the multiplication factor. The reactivity is a nondimensional quantity specifying the relative deviation from the critical condition:

$$\rho = \frac{k_{\text{eff}} - 1}{k_{\text{eff}}}$$

$\rho = 0$ applies to the critical reactor; $\rho > 0$, to the supercritical reactor; and $\rho < 0$, to the subcritical reactor. Because the deviation from the critical condition is very small, pcm $= 10^{-5}$ has been introduced as the abbreviation for reactivity (pcm = per centum mil). $\rho = 40$ pcm is then written, for instance, as $\rho = 0.0004$. Besides this, the reactivity ρ can also be expressed in percentage: $\rho = 1000$ pcm $= 1\%$.

What specifically happens with a neutron generation is described in Figure 20.3 in simplified terms, where 2.3 neutrons per fission are assumed.

The core is surrounded by a reflector. It is designed to diffuse part of the neutrons leaking out of the core by way of scattering processes back into the core. This enables the critical size of the reactor to be reduced. Furthermore, a reflector somewhat equalizes the neutron flux spatially in the core in radial and axial directions. This homogenizes the power density and the coolant outlet temperature. Moreover, a sufficiently large reflector design protects the reactor pressure vessel from high neutron flux rates. The moderator and reflector are made of the same material in most reactors.

With a view to the dependency of the neutron yield on the energy of the fission neutrons, reactor systems are most favorable in terms of neutron economics when most of the fission processes happen in the thermal energy range (U-235, U-233) or in the fast energy range (Pu-239). As a result, thermal reactors operate with U-235 or U-233, while fast reactors use Pu-239 as the fuel.

Most reactors built to date are thermal reactors. This is mainly because such a system is the easiest to turn critical. The amount of fissile material required here is relatively small compared with fast systems because the effective cross sections are significantly higher in the thermal energy range than in the fast energy range.

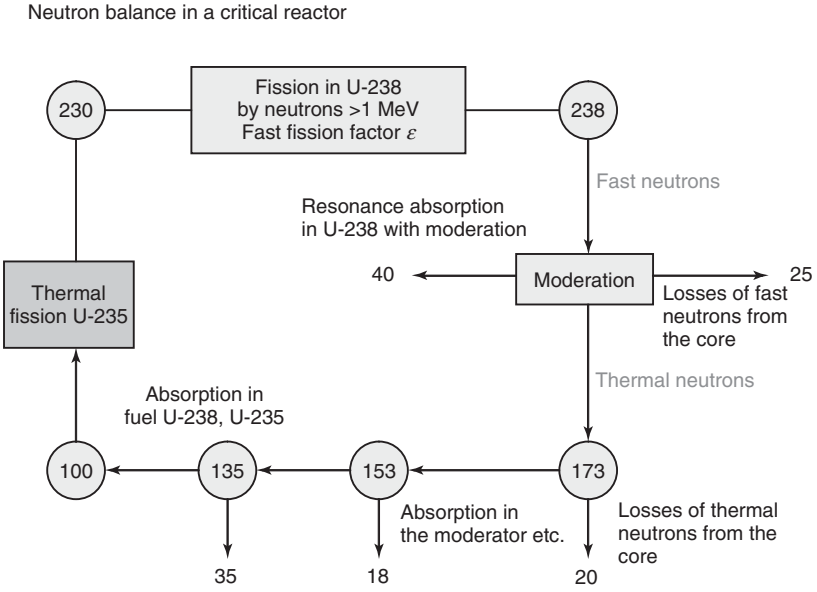


Figure 20.3 Neutron balance in a critical reactor.

20.4.2 Moderator

Neutrons released during the fission of U-235 nuclei have a relatively high speed (approximately 15 000 km s^{-1}). The likelihood for them to cause more fission processes is therefore extremely low. The likelihood of slow neutrons to cause fission processes in U-235 is much higher. Therefore, the speed of fast neutrons created with every nuclear fission has to be lowered (approximately 2 km/s) by appropriate substances. This is done with the so-called moderators (slowing agent). The neutrons are slowed down by hitting against the nuclei of the moderator.

As a rule, a good moderator should have the following neutron physics properties:

- The fast neutrons should, if possible, loose their energy by few impacts on the moderator atomic nuclei so as to be slowed down. As a result, they overcome in large numbers the speed range at which they are increasingly absorbed by U-238. This condition is met if the mass of the moderator atomic nuclei is about half of the neutron's mass.
- The moderator may only have a low susceptibility to capture neutrons because neutrons absorbed by the moderator are lost for further nuclear fission.

It should be pointed out, however, that the capture cross section of the moderators is of relatively less importance in power reactors in actual practice because the absorption losses can be compensated for by a correspondingly higher level of enrichment of the fuel. This is why the lowest priced moderator is typically chosen. In many cases, this is light water with the use of which the savings outweigh the extra costs for enrichment. Heavy water or graphite is only used if special aspects, for example, specific independence of enrichment facilities, justify neutron management austerity. Both moderators enable fuel elements of natural uranium to be used.

20.4.3 Neutron Capture Systems

In a reactor, the chain reaction is controlled by interfering with the neutron balance from outside. This is done by substances that have a large susceptibility to absorb neutrons, for example, boron, indium, silver, cadmium. Extremely neutron-absorbing silver–indium–cadmium alloys, for example, are used in the so-called control elements; they absorb neutrons when immersed into the fission zone of the reactor core to stop the chain reaction.

20.4.4 Heat Removal Medium

The coolant is also the moderator in LWRs. This involves the advantage that the moderation depends on the temperature of the water. When the number of nuclear fission processes and hence the power of the reactor increase, the density of the coolant decreases. Although fewer neutrons are absorbed, even fewer are slowed down. This reduces the number of nuclear fission processes automatically. This self-stabilization is referred to as *inherent stability* of a reactor.

20.4.5 Barriers for Radiation Protection and the Retention of Radioactive Substances

Nuclear power plants have a number of barriers to fulfill two functions. They shield

off the direct radiation, essentially neutrons and gamma radiation, and prevent radioactive substances from leaking out. Alpha and beta radiation are completely shielded off by the coolant. The reactor's pressure vessel is surrounded by a steel concrete shield that almost completely shields off gamma and neutron radiation. The containment and the reactor building constitute additional barriers so that hardly any direct radiation from the reactor leaves the reactor. Some radiation shielding barriers are also safety barriers to prevent radioactive substances from escaping to the outside.

20.5 Reactor Core Design

Reactors with a predominantly *thermal neutron spectrum* are referred to as *thermal reactors*. They include the LWRs (PWR and BWR), which use water as moderator, and the gas-cooled graphite-moderated reactors.

If the share of the moderator in the core is reduced, the neutrons created as fast fission neutrons keep higher energies and their energy spectrum is shifted into the so-called epithermal range. In the absence of a moderator (no lightweight atomic nuclei in the core), the neutrons remain at the high-energy level ("fast"). A fast neutron spectrum is a characteristic of a *fast reactor*. As the effective fission cross section for fast neutrons becomes smaller, this has to be compensated for by a higher density of fission nuclei (higher level of fuel enrichment with fission isotopes) to meet the critical condition.

In fast reactors, the neutron yield per nuclear fission is higher on a statistical average, which is used to produce new nuclear fuel (fast breeder). η factors in excess of 2 enable a fertile material (e.g., U-238, Th-232) to be converted into fissile materials (Pu-239, U-233). A conversion rate (CR) greater than 1 is referred to as *breeding rate*.

$$\mathrm{CR} = \frac{\text{Number of fissile nuclei generated per second}}{\text{Number of fissile nuclei consumed per second}}$$

In most cases, fuel and moderator are arranged heterogeneously. Nonetheless, a *homogeneous reactor* with a homogeneous mix of fuel and moderator is often assumed in theoretical appraisals.

The fuel rods and the moderator are distributed heterogeneously in the core of LWR. The moderator (water) is also the coolant that flows through the spaces in between the rods to remove the nuclear heat generated. The individual fuel rods are cladding tubes welded gastight and filled with the fuel in the form of pellets.

The reactor core consists of an array of parallel fuel elements whose number depends on the power and type of the reactor. Other components are the control rods, neutron sources, flow restrictor assemblies, and parts of the core instrumentation.

Approximately 190 fuel elements are used in PWRs. Every fuel element consists of fuel rods whose number can vary substantially depending on the type of fuel element. Typical values are 236 for a PWR fuel element with 16×16 fuel rod geometry and the cladding tube material being a zirconium alloy.

The fuel elements of BWRs are smaller and typically contain 62 and 80 rods. Four fuel elements are usually assigned to one fuel channel. All in all, approximately 800 fuel elements are used in BWRs.

In PWRs, parts of the fuel elements contain the so-called control elements. Other

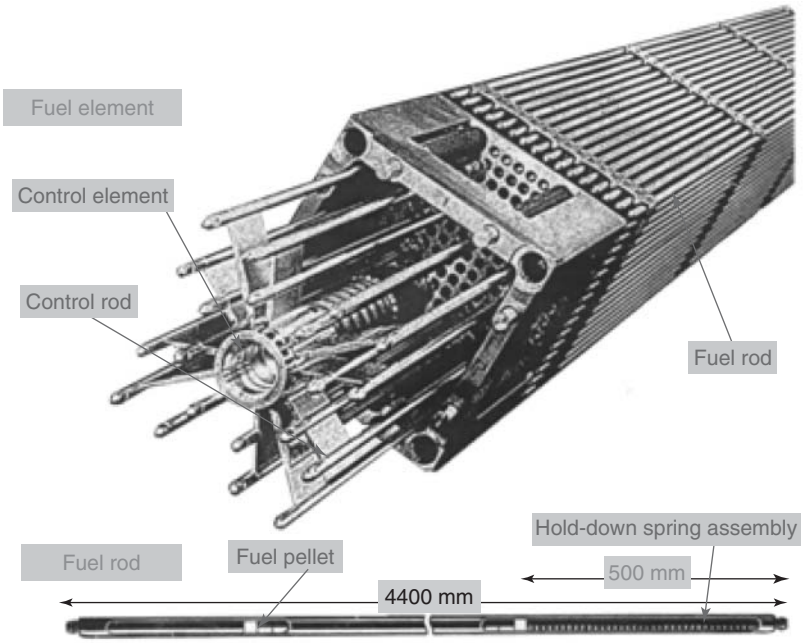


Figure 20.4 PWR fuel element.

fuel elements comprise neutron sources and probes for the core instrumentation. The other free guide thimble tubes are fitted with flow restrictor assemblies to prevent undesirable secondary flows of the main coolant.

A fuel element typically used in PWRs (Figure 20.4) today usually consists of upper/lower tie plates, 8 spacers, 20 control rod guide thimbles, and, depending on the design, 236 or more fuel rods.

The guide thimbles serve the purpose of guiding the control elements and slowing them down hydraulically on the last part of their fall distance in the event of a reactor scram.

The spacers consist of thin-walled plate webs connected at the intersections. They absorb the transverse forces/thermal expansion under the various operating stresses, enable the neutron-induced growth in length, and support the heat transfer from the fuel rods to the coolant by cross-mixing (spacer mixing vanes) (Figure 20.5).

The upper tie plate ensures reliable gripping with fuel element handling tools, for example, the refueling machine gripper.

The lower tie plate holds the fuel element in the desired position. Coolant bores in the lower tie plate distribute the inflowing coolant and impede the ingress of foreign matter.

The fuel rods consist of the fuel pellets, cladding tubes, end plugs, fuel rod springs, support tubes and, if appropriate, insulating pellets. Like the end plugs, the cladding tubes consist of zircaloy, an alloy based on zirconium. They are designed to separate the nuclear fuel from the coolant of the primary circuit as well as to prevent the fission products arising from nuclear fission from mixing with the coolant.

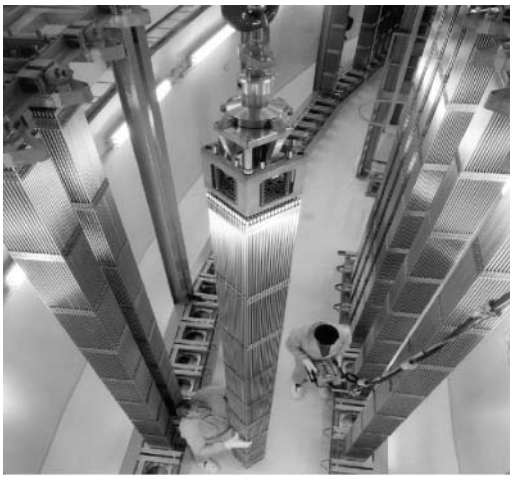


Figure 20.5 PWR fuel elements in dry storage. (Source: RWE Power.)

The fuel rod spring assembly of austenitic spring steel wire presses from above onto the pellets and holds them firmly together in a column. At the same time, this creates space above the nuclear fuel for the noble gases and highly volatile fission products created by nuclear fission. The fission gas space prevents an undue gas pressure increase in the fuel caused by very strong heating.

A fuel element has to meet the following general requirements:

- maximum safety (containment of radioactivity, shape stability)
- operational reliability
- operational flexibility (load changes)
- economic efficiency (burn-up, fuel utilization).

These requirements are met by a design taking account of the stresses imposed on the fuel element across the entire lifetime (power operation to accident).

During operation, the average temperature in the fuel rod is about 1300 °C; temperatures of up to 2300 °C are reached at the hottest point. These temperatures are significantly below the melting point of uranium oxide (2800 °C). Noteworthy is a high temperature gradient on the periphery of the fuel rod where levels of up to 200 $K\,mm^{-1}$ are reached.

During power operation, the fuel undergoes changes in dimension as a result of initial recompaction (closing of fine pores) and subsequent swelling by incorporating fission products into the fuel structure.

The swelling is prevented by setting only about 95% of the theoretical density during manufacture (sintering process). This porosity enables fission products to be absorbed to a certain extent.

Essential design parameters of the fuel lattice are the fuel rod diameter and the water/fuel volume ratio and/or particle density ratio of water and fuel (moderation ratio).

The reactivity of a nuclear reactor depends on

- the volume of the absorbers (e.g., control rods, boron concentration in the coolant) and their changes;
- the thermohydraulics of the core and hence temperatures of the fuel and moderator (feedback processes);

- the concentration of the fission material inventory in the core and its changes (burn-up);
- the concentration of the neutron-absorbing fission products (the so-called neutron poisons) and their changes.

For the PWR, the coolant temperature coefficient, the fuel temperature coefficient, and the control rod effectiveness are relevant over the short term. The coolant temperature coefficient describes, for instance, the correlation of reactivity change and coolant temperature change. Different water temperatures result in changes of density and hence the macroscopic effective cross section. This means that a temperature increase results in poorer moderation, which binds reactivity; as a result, the coolant temperature coefficient is negative.

At higher temperatures, the fuel temperature coefficient leads to an increased neutron absorption in the epithermal range of U-235. This reactivity change is also referred to as the *Doppler effect* (model concept: nuclei oscillate more strongly and hence offer a larger impact area).

Over the long term, burn-up, boron effectiveness, and the concentration of neutron-absorbing fission products (reactor poisoning) are of importance. As an example of poisoning, during reactor operation, fission turns the heavy atomic nuclei of the fuel into a multitude of different lighter nuclei. Many of these fission products have a very large absorption cross section in respect of thermal neutrons and therefore disturb the neutron balance in the reactor. A prime example is Xe-135, which has a very high microscopic absorption cross section compared with thermal neutrons. Xe-135 essentially results from radioactive decay of the direct fission product Te-135.

The higher the constant power, the higher the reactivity loss. This significant loss of reactivity can even be reinforced if the reactor power is reduced by a larger amount following an extended period of constant load operation. The neutron loss by absorption is then higher than that by the equilibrium poisoning. The xenon poisoning drops afterward because the iodine production has also subsided as a result of the power reduction. In the no-load condition, the xenon poisoning has virtually disappeared after about 3 days and the reactor is ''free of xenon'' again.

20.6 Thermohydraulics of the Reactor Core

UO_2 is a good heat conductor in the pellet form. The central temperature of a pellet is determined by the fuel power and the heat transfer from the pellet edge to the cladding tube inside and outside tobreak the coolant. Sufficient cladding tube integrity is only ensured in design-based operation if no molten material arises in the rod center (results in strong fuel swelling and hence damage to the cladding tubes) and if in the long term the cladding tube temperature does not rise above 350 °C (zircaloy structure changing from 500 °C).

A reactor is a system with nominal power. The heat transfer to the coolant can be calculated fairly well in channels with circular cross sections. The cooling channels in reactors, however, do not have a circular cross section. The geometry in most power reactors is that of a bundle or lattice consisting of a large number of parallel circular/cylindrical rods around which the coolant flows in the axial direction. This is why the concept of the ''hydraulic diameter'' and of the cooling channel was defined.

If subcooled water enters into a fictitious cooling channel (e.g., space between four fuel elements) with axially constant heat flux density, the heat is transferred from the fuel elements to the single-phase water flow by convection. As the water temperature keeps increasing, local steam bubbles form at the cladding tube walls, which condensate again, however, because the water temperature remains below the boiling temperature in the core area of the cooling channel. This significantly improves the heat transfer from the cladding tube to the water. This heat transfer is referred to as *subcooled boiling*. In PWRs, heat is transferred by convection or subcooled boiling.

If the water reaches boiling temperature, the vapor bubbles remain unchanged and do not condensate anymore. This heat transfer to the two-phase water/vapor flow is referred to as *nucleate boiling*. The water temperature does not rise anymore because the heat flux entering the coolant is used to evaporate the coolant. Owing to the good heat transfer, the cladding tube temperature does not rise again.

The vapor bubbles coalescing into plugs eventually form a coherent vapor hose. The cladding tube surface is then only covered by a thin water film enclosing the vapor hose as a ring. This almost pure vapor flow with a water ring is referred to as ring flow. The BWR can be operated until the ring flow sets in.

A single-phase vapor flow exists once the last remaining liquid has evaporated. The heat transfer from metal to vapor is much poorer than that onto evaporating water. The cladding tube temperature surges up to a level from which the temperature difference to the coolant is sufficient to evacuate the thermal energy generated by the fuel rod. However, the cladding tubes are not designed for this stress level. This boiling condition is the result of insufficient cooling.

In a real cooling channel, the axial heat flux density is not constant, but is relatively low at the inlet and outlet and reaches a maximum roughly in the middle.

The resulting relevant correlation of heat flux density and temperature difference between the cladding tube surface and the coolant is described by the Nukiyama curve (Figure 20.6). As described by the Nukiyama curve, heat is almost completely transferred by convection when heat flux

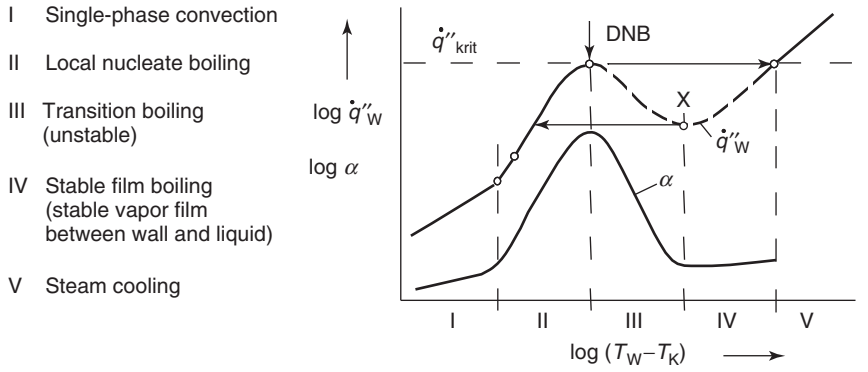


Figure 20.6 Heat flux density, $\dot{q}''_w$, and heat transfer coefficient α as a function of the temperature difference ($T_W - T_C$). W, wall; C, coolant.

densities are low. As a result of good heat transfer, as indicated by the initial steady rise of the heat transfer coefficient, the transferring thermal power increases steadily and is increased once again when nucleate boiling sets in.

At DNB (departure from nucleate boiling), the critical boiling condition has been reached from which an insulating vapor film forms when the thermal power continues to increase (or cooling is reduced). The heat transfer of this vapor film is a 100 times lower; the heat transfer coefficient decreases heavily. As the thermal energy generated in the fuel does not change in this range, the temperature in the fuel and cladding tube shoots up to a level at which the temperature gradient reached is sufficient again to remove heat. The vapor film is unstable at first so that the surface is wetted with water time and again. Stable film boiling does not occur and the heat transfer coefficient does not rise slightly again until this range has been passed completely.

The critical heating surface stress of the cladding tube material is exceeded when there is film boiling; the cladding tube burns through. This is why the actual heating surface stress must have sufficient margin to the critical level, that is, must be lower, for safety reasons.

In reactor operation, decay and radiation of the fission products account for 5–7% of the thermal power. The decay heat power of nuclear fuels is the thermal power generated by radioactive decay of the fission and activation products of the nuclear fuel after a nuclear reactor has been shut down. It is an important physical quantity for the design of systems for which the decay heat power has to be taken into account as the heat source. Cooling of the core is therefore required for weeks after reactor shutdown to prevent undue temperature levels at the fuel rod cladding tubes.

20.7 Design, Function, and Safety Concepts as Demonstrated by a Pressurized-Water Reactor (PWR)

With a share of over 80% of all plants currently in operation, the LWR is the most widespread design of a nuclear reactor for which light water (H_2O) serves as the moderator and coolant. The reactor is typically designed as a BWR or PWR, with the core being accommodated in a pressure vessel, surrounded by coolant.

Instead of building a large pressure vessel, water can also be fed through many thin and pressure-proof tubes past the reactor core. However, because the reactor is surrounded by less water in this case than that in a BWR or PWR, an additional moderator is needed, which is typically graphite. This variant of the BWR is referred to as the *pressure tube reactor*. This includes the RBMK of the Soviet design.

The PWR is the leading design among the LWRs and is therefore explained in greater detail in the following.

The energy released by nuclear fission in the reactor core of a PWR is to be converted into electrical energy (Figure 20.7). This process goes through the intermediate stages of

- thermal energy
- mechanical energy (rotational energy)

The electrical output generated amounts to up to 35% of the thermal power (e.g., 1292 MW).

The fragments and particles created by nuclear fission are flung away from each other with a high level of energy and their kinetic energy is transferred largely by being slowed down across an extremely short distance as thermal energy

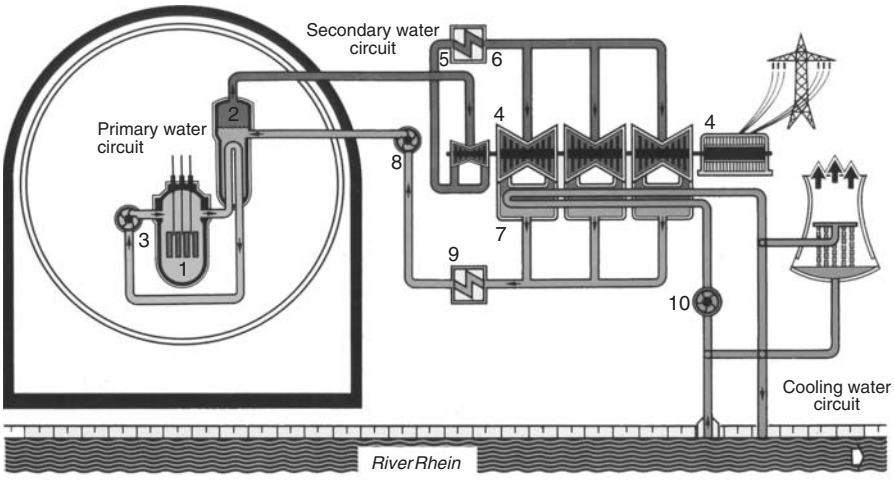


Figure 20.7 Functional diagram of a nuclear power plant with pressurized-water reactor. (Source: RWE Power.)

to the uranium material by which they are surrounded.

The conversion of nuclear energy into thermal energy (in the fuel rod) is ensured by the continuing fission processes. The thermal energy is transferred from the fuel rods through the fuel rod cladding tubes on to the primary coolant. The primary coolant heated up in the reactor core flows to the steam generator where the primary coolant gives off the thermal energy through the steam generator heating tubes to the secondary medium (feedwater).

The function of the secondary circuit of a PWR can be compared with that of large-scale conventional thermal power plants. The steam under a certain working pressure flows to the turbine and drives the turbine rotor. The condenser downstream of the turbine is under negative pressure to increase the useful pressure gradient.

The generator rotor, a rotating electromagnet, is coupled to the turbine rotor. Its corotating magnetic field induces electric current in the lines of the generator stator (conversion of mechanical into electrical energy).

The steam expanded in the turbine is condensed through cooling by means of the cooling water circuit in the turbine condenser. The condensed water (condensate and/or feedwater) is pumped back to the steam generator. The condensation heat of the steam from the turbine absorbed by the cooling water is released as lost energy (cooling losses) to the environment.

The cooling losses are reduced by using the steam in the secondary circuit to preheat the feedwater. To this end, expanded steam is partly extracted from the turbine to heat (preheat) the condensate and feedwater. This reduces the mass flow of the steam to be condensed by cooling, which also decreases the cooling losses to the environment. Another consequence is that less energy has to be provided to the

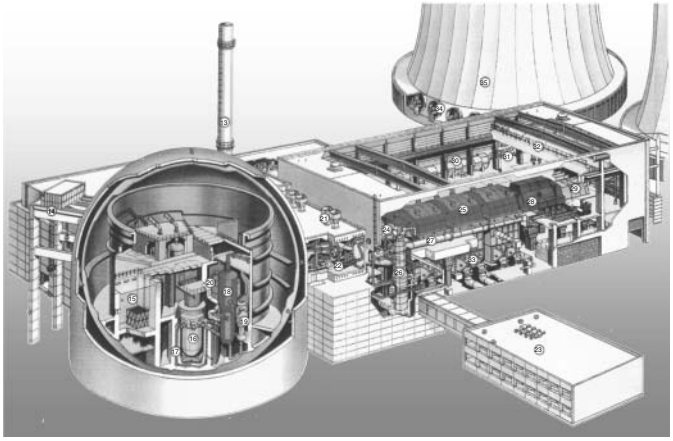


Figure 20.8 Sectional view of the Biblis Nuclear Power Plant Unit B. (13) Vent stack, (14) lift rack, (15) fuel element storage pool, (16) reactor (pressure vessel), (17) biological shield, (18) steam generator, (19) primary coolant pump, (20) pressure controller, (21) exhaust silencer, (22) main steam pipe, (23) ancillary system building, (24) turbine high stage, (25) turbine low stage, (26) reheater, (27) crossover pipe, (28) generator, (29) shaft-mounted exciter, (30) feedwater tank, (31) water tank, (32) lift, (33) condenser, (34) ventilator, (35) cooling tower, and (36) personnel air lock. (Source: RWE Power.)

steam generators because they are fed with preheated feedwater.

In the PWR, the operating pressure of the water in the primary circuit is set at such a level that it does not yet boil at the intended operating temperature. The coolant is passed through the reactor core in the primary circuit at a pressure of 154–160 bar where it absorbs the heat generated by nuclear fission (Figure 20.8).

The mean primary coolant temperature is between 300 and 310 °C depending on the type of reactor. The thermal energy is transferred through several primary coolant circuits (two to four loops) with a corresponding number of steam generators and main coolant pumps to a secondary circuit (feedwater steam circuit). Once the heat has been transferred, the coolant flows back into the reactor core. This involves the advantage that the coolant, which also serves as moderator and is furthermore always slightly contaminated with radioactive materials, always remains within the containment. This is why no radiation protection measures are required in the turbine hall of a PWR as opposed to a BWR (Table 20.1).

20.7.1 Nuclear Auxiliary Systems

Smooth interaction of the main coolant circuit with various auxiliary and ancillary systems is required for undisturbed reactor operation.

A selection of nuclear auxiliary systems with their main functions is listed in the following:

- **Volume control system**: Feeding and extraction of coolant.
- **Coolant cleanup system, coolant gas stripping, coolant treatment and storage system, and fuel element pond cleanup system**: Treatment of coolant and/or fuel element pond water.

Table 20.1 Specifications of the Biblis Nuclear Power Plant Unit B.

Power plant net capacity	1 240 MW
Generator continuous rating	1 300 MW
Power plant efficiency rated to net capacity	33.2%
Uranium weight (U-238 + U-235)	102.7 Mg
Mean U-235 concentration of the first core	2.48%
Mean U-235 concentration of reloadings	Up to 4%
Number of fuel assemblies	193
Number of control rods	61
Reactor coolant flow	72 000 Mg h^{-1}
Mean coolant temperature	303.5 °C
Coolant pressure	154 bar
Main steam flow	7 160 Mg h^{-1}
Main steam pressure	53 bar
Mass flow of reactor coolant pumps	4 × 18 000 Mg h^{-1}
Motor rating of reactor coolant pumps	4 × 8550 kW
Weight of reactor pressure vessel	550 Mg
Internal diameter of reactor pressure vessel	5 000 mm
Wall thickness and plating of reactor pressure vessel	243 + 7 mm
Overall height of reactor pressure vessel	13 250 mm
Cooling water flow of river or cooling tower	218 000 m^3 h^{-1}

- **Core emergency and decay heat removal system**: Decay heat removal from the shutdown reactor, ensuring sufficient core cooling.
- **Nuclear ventilation systems**: Cooling and filtering of air in the controlled area.
- **Sealing water system**: Sealing water supply for the sealing system of the main coolant pumps.
- **Wastewater system**: Treatment of wastewater from coolant cleanup and wastewater in the controlled area.
- **Off-gas system**: Controlled/delayed release of radioactive gases.
- **Chemicals feeding system, boric acid makeup system**: Dosing of neutron-absorbing boric acid and other chemicals.
- **Treatment of radioactive concentrates**: Incorporation of radioactive concentrates, filter residues, consumed resins, and spare parts.

For the protection of the environment, it is required that the release of gaseous, liquid, and solid radioactive products remains below the limits specified by the Radiation Protection Ordinance. The radioactive products created in the closed main coolant circuit are retained there. They decay in there for the most part.

All liquid radioactive products are consigned to intermediate storage sorted by levels of activity. The larger volumes of low-level effluents are cleaned by chemical precipitation and filtering and are discharged in diluted form through the outfall ditch. Liquids with a higher level of activity are concentrated several times by evaporation until they can also be disposed of. Filtering and evaporation residues are solidified and consigned to the solid-waste repository or disposed of for further processing.

Apart from the used fuel elements to be disposed of in special-purpose shipment

containers, depleted ion exchanger resins, evaporator bottoms, and used-up consumables arise as solid radioactive products. The waste products are sorted by levels of activity and are committed to interim and/or ultimate storage packaged in standard shipment drums.

Technical improvements, especially in terms of the quality of fuel elements and process engineering optimization, have led to significantly lower-activity releases than officially permitted. In this context, tritium is an exception whose emission rate is about half the permitted limit.

20.7.2 Safety Principles

The overriding safety principle is the protection of life, health, and material goods against the risks of nuclear energy and the harmful effects of ionizing radiation [6–8]. The radioactive substances, which are mainly concentrated in the reactor core, are enclosed by several barriers to ensure this level of protection. These are the fuel rod cladding tubes, the pressurized enclosure of the reactor coolant circuit, and the containment. The reactivity of the reactor core must be kept within narrow limits, and the fuel elements always have to be cooled sufficiently to ensure the integrity of the barriers. These three protection targets, completed by the fourth protection target of limiting radiation exposure, comprise all the requirements that when met ensure protection. A staged safety concept comprising several levels, referred to as *defence-in-depth*, ensures that the protection targets are met and the integrity of the barriers is maintained.

Originally, this concept consisted of a combination of measures to prevent disruptions and accidents as well as measures designed to safely control them. The high degree of automation of accident management systems, designed to prevent human error and several redundancies of physically separate and also partly diverse safety systems, plays a key role in managing the so-called design-basis accidents. Given its extremely low probability of occurrence, a core meltdown was originally part of the residual risk.

In the 1970s, probabilistic safety analyses came to be used in addition to the deterministic safety assessment. For these analyses, appraisals of the reliability of the operating systems and the safety systems of a nuclear power plant are made using the methods of probabilistic theory.

With a view to minimizing risks further, events were also taken into account as from the 1980s, which had not been regarded as design-basis accidents as part of the original design owing to their low probability of occurrence.

Likewise, measures of on-site emergency protection were put in place in order to reliably identify beyond-design-basis conditions early on, to control them, and to stop them with as little damage as possible. The preventive measures of on-site emergency protection are designed to prevent severe core damage; their main objectives are maintaining or restoring core cooling and making the plant a safe environment. The damage-preventing measures are designed to mitigate severe radiological effects in the plant and its vicinity in the event of core damage; their main purpose is to maintain the remaining activity-retaining barriers and to secure a long-term controlled condition to protect the environment. The on-site emergency measures are based on the flexible use of available safety and operating systems also outside their field of application intended by design and involving the risk of damaging them, as well as on external systems.

Wide-ranging technical and organizational precautions have been taken in many nuclear power plants to carry out the planned measures effectively during an emergency. In the preventive area, for instance, systems were backfitted in reactors for fast pressure relief of the reactor coolant circuit, by which core meltdown scenarios under high pressure are avoided. Damage-mitigating measures were implemented, such as the installation of catalysts to decompose the hydrogen released during core damage so that hydrogen explosions with a risk for the containment would be avoided in the event of a core meltdown, which was still assumed after all. Another measure is the installation of a pressure relief system through which the filtered gases can be released from the containment to prevent failure of the containment caused by excessive pressure and to largely retain the radioactive substances.

The terrorist attacks of 11 September 2001 have led to an intensive debate on how we can protect ourselves against warlike terrorism. Against this background, the discussion is focused, in particular, on the protection of nuclear power plants against a targeted aircraft crash. Technical and organizational precautions were already taken at the design stage to provide protection against acts of sabotage and other unauthorized interference. These have been improved on an ongoing basis over the entire period of operation. In so far, nuclear power plants rank among the best protected industrial facilities. Through the interaction of precautions taken by governments and operators, the risk of a terrorist attack on nuclear power plants with serious effects can be reduced significantly.

In sum, it is safe to say that in many countries, safety and security have been permanently updated to match the development of the state of the art and that numerous improvements, especially also in the beyond-design-basis area, have been implemented in the past few years.

20.7.3 Safety Barrier Concept

The safety design of a nuclear power plant is based on the safe enclosure of the radioactive substances in the reactor and in irradiated fuel elements by means of several barriers (Figure 20.9). The required number and the design of the barriers depend on the risk potential of the enclosed radioactive materials and substances and the potential of them being released.

These targets are to be achieved by

- meeting high-level quality requirements in terms of design, erection, operation, monitoring, and maintenance of the plant to prevent incidents, events, and accidents;
- providing several separate and independent systems and prepared measures based on a defence-in-depth concept to protect the integrity of the barriers and, if the integrity is damaged, to contain the associated effects, but if the systems and measures fail on a given safety level of this staged concept, the effects of this failure are controlled and/or contained by systems and measures at the next higher level;
- virtually excluding events, which could lead to an overriding failure of the staged safety systems, by reliable systems and prepared measures;
- mitigating release of radioactive substances, which can nonetheless be caused by an event or accident, systems, and prepared measures.

The following barriers are in place to effectively retain the radioactive substances in the reactor:

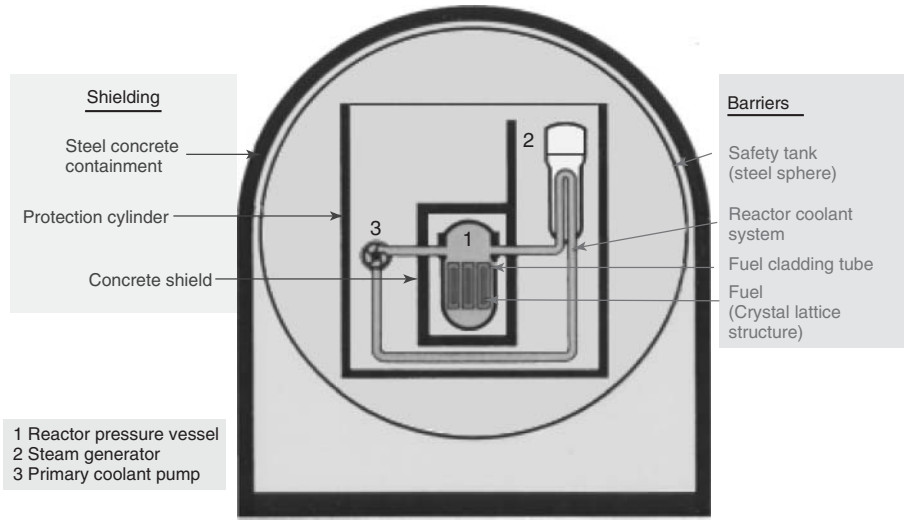


Figure 20.9 Safety barriers to retain radioactive substances and for direct radiation shielding. (Source: RWE Power).

- the fuel matrix and the fuel rod cladding tubes;
- the pressure boundary of the reactor coolant consisting of the reactor pressure vessel, the primary-coolant-conveying parts of the steam generators, the pressurizer, the main coolant pump housing, and connecting pipe work branching off with larger nominal diameters up to and including the first shutoff valve;
- the containment (includes the component of the safety tank as well as safety functions for its isolation, if required).

Facilities and systems-engineering functions to support the retention function of the barriers have to be taken into account with their effects as part of the barrier function. Such facilities and/or functions are, in particular,

- incorporation of radionuclides in solids (e.g., ceramics);
- negative pressure staggering and associated retention systems;
- separation of activity-containing and non-activity-containing systems;
- submersion in water during nonpower operation with open primary system and during the storage of fuel elements in the storage pond;
- protective effect of the reactor building and other buildings against airborne and/or liquid releases.

If individual barriers are opened or unavailable for operating processes, safe retention of the radioactive substances has to be ensured by means of the remaining barriers and other retention functions and, should this not be enough, by additional backup measures.

The barriers should be sufficiently independent of each other so that in the event of incidents with design-based effectiveness of the safety functions, one barrier does not fail because of the failure of another.

The following principles are applied, among others, to ensure a high level of reliability of measures and safety systems to manage and control incidents and events:

- *Conservative design*: This so-called design reserves are provided for at many points of the entire reactor plant. Moreover, components with a proven operational or experimental track record are used as far as possible. The plant is inherently safe.
- *Quality management*: The safety-engineering functions of all plant parts, systems, and components as well as their quality requirements aligned with the respective safety-engineering relevance have to be clearly defined, and the quality checks have to be documented. This goes not only for the construction phase but also for the operation of the plant (e.g., periodic testing and inspections).
- *Redundancy*: Important safety devices are generally installed in multiple sets. This guarantees that when repair or maintenance is required or in the case of a single failure, one or more backup systems immediately and automatically step in and thus stoppages are always controlled. For example, there are four sets of the systems that are used to cool the reactor, while only two are required in a stoppage.
- *Diversity*: Different systems are installed to meet the same safety function. Examples are pump drives driven by steam, diesel, and electricity and actuators of valves, electric and pneumatic.
- *Spatial separation*: The spatial separation of redundant and diverse facilities will ensure that multiple systems cannot fail simultaneously because of a single cause.
- *Fail-safe principle*: In the event of failure, all devices respond in the direction of safety. For example, if the power supply fails, valves and flaps rise to the safety position.
- *Multistage control and monitoring system*: In addition, each nuclear power plant is equipped with a multistage control and monitoring system. During the operation of the plant, this continually controls all important measurable data, compares each reading with the target state, and corrects recognized abnormal operating conditions. If any previously defined limits are reached, the control and monitoring system automatically triggers active security measures. These include the following.
  - Fast shutdown of the reactor by interruption of the power supply: In a shutdown, control rods are inserted between the fuel rods and the chain reaction is stopped. There is also a system for fast shutdown by adding boric acid, which stops the chain reaction. In BWRs, the control rods are inserted into the reactor from below. In normal operation, this is done by electric motors. An independent hydraulic system is used for fast shutdown.
  - The tight sealing of the various parts of the buildings: This prevents leakage of the radioactive substances from the primary circuit to the outside.
  - The residual heat removal system: In shutdown or in case of a loss of coolant in the primary circuit, the residual heat removal system is

automatically activated. This not only removes the residual heat from the deactivated reactor but also simultaneously replaces the coolant loss.
- The emergency power supply: This takes over the power supply for all relevant safety systems when the normal power supply to the plant is interrupted or has failed. If the external supply of power is not possible, power supply is ensured by several emergency diesel generators. In addition, there are battery systems that can also be used as a stopgap in the short term.
- The emergency core cooling systems: If there is a failure in the primary reactor cooling system, various emergency cooling systems are available to feed cooling water into the reactor pressure vessel and ensure cooling of the fuel assemblies.

The reactor protection system is activated automatically so that interference by the operating staff is not required within 30 min following identification of an event. The safety system is self-controlled and has priority over the actions of operating systems. The functional capacity of safety equipment and security measures is systematically tested through a predetermined program of periodic inspections.

20.8 Radioactive Waste Management

All activities for the treatment, processing, and disposal of the depleted fuel elements and radioactive waste are commonly referred to as *waste management* [1, 6, 7].

When it comes to nuclear waste, a distinction is made between low-level and medium-level as well as high-level waste.

- **High-level waste**: This accounts for approximately 10% of the calculated waste volume and comprises approximately 99% of the entire radioactivity. This includes mainly irradiated fuel elements from nuclear power plants, research and experimental reactors, and waste from the reprocessing of fuel elements. The radioactive decay of the contained radionuclides releases a significant amount of heat, which is why this waste is also referred to as *heat-generating waste*.
- **Low-** to **medium-level waste**: These account for 90% of the forecast waste volume, but contain only 1% of the radioactivity of all the waste. This is essentially the waste from the operation and dismantling of nuclear power plants, from reprocessing and from research and the medical sector. The radioactive decay of this waste does not produce any significant amount of heat, which is why this waste is also referred to as *radioactive waste with negligible heat generation*. A small portion of the medium-level waste is accounted for by heat-generating waste (e.g., fuel cans, structure parts from reprocessing).

The radioactive inventory and the heat generated by radioactive decay are particularly relevant to the safety-analysis appraisals for ultimate storage. Most countries have opted for the separate ultimate disposal of the waste types in deep geological formations. This involves safety-related

advantages because the two waste types have different properties. With this in mind, the safety requirements of the two ultimate repositories can best be aligned with the respective waste requirements.

A nuclear power plant with an electrical capacity of 1300 MW annually produces about 50 m^3 of conditioned radioactive operating waste with negligible heat generation as well as waste from the disposal of the annually unloaded depleted fuel elements, the volume of which depends on the disposal path: fuel element reprocessing results in 10 m^3 of radioactive waste with negligible heat development and 3 m^3 of heat-generating waste – the vitrified high-level fission products. Direct ultimate disposal of the depleted fuel elements results in 45 m^3 of heat-generating waste.

20.9 Decommissioning and Dismantling of Nuclear Power Plants

Like any other industrial facility, nuclear power plants also have to be decommissioned at the end of their operating life.

Decommissioning and dismantling of nuclear power plants comprise:

- the postoperational phase serving to prepare the power plant for dismantling; it is covered by the valid operating license, and this phase ends with the implementation of the final and definitive decommissioning permit;
- the dismantling of all facilities and removal of all radioactivity of the building structures, right down to releasing the site from the scope of nuclear supervision.

The following decommissioning options are available [7, 9]:

- ''Immediate dismantling'' of the plant or of plant components right after the postoperational phase.
- ''Safe enclosure'' is the option that postpones the final removal of controls for a longer period, usually in the order of 30 years. The plant is placed in a safe storage configuration until the dismantling activities occur.

Decommissioning variants may also be mixed forms of these two basic options.

Dismantling ends when the site is removed from regulatory control. Afterward, the buildings can be torn down conventionally. Very often the aspiration is the so-called green field, which means the complete removal of all buildings and the reclamation of the site. However, a large number of other alternatives are also possible; for instance, the industrial use of a released site, partial dismantling of the controlled area, and another use for the remaining buildings. In countries that intend to continue operating a nuclear industry in the future, the continued nuclear use may also be an important option.

There are pros and cons to both variants (immediate dismantling and safe enclosure). A trend toward immediate dismantling has emerged in the international field. The following aspects, in particular, are regarded as being relevant to the decision in favor of immediate dismantling:

- planning and implementation of decommissioning (expertise and specialist know-how still available in the plant can be used);
- financing (guaranteed financing has to be provided for a shorter period);
- social concerns (job security, minimization of local economic problems);

- new techniques developed for decontamination and (remote controlled) dismantling; the decay of radionuclides during safe enclosure has lost importance from a radiation perspective.

When dismantling the nuclear part (controlled area) of a large nuclear power plant with a PWR, to go back to the ''green field,'' more than 95% of the total mass of controlled area (approximately 156 500 Mg) can almost completely be put to conventional use again once potential surface decontamination has been performed. About 1–2% can be recycled following decontamination and/or once molten down (i.e., steel structures). This means that only about 2–3% of the total mass of the controlled area has to be disposed of as radioactive waste in an ultimate repository.

20.10 Outlook on New Reactor Technologies of the Third and Fourth Generations

Advanced nuclear power plant types of the third generation have been developed since the early 1990s. These have reached commercial maturity today and form the basis for the new-builds in the years ahead. The development of nuclear systems of the fourth generation (Gen-IV), which may be needed in the second half of the twenty-first century or even later, is currently the focus of international research activities [7].

The power plant types of the third generation constitute in many cases the advanced development of the reactor types of the second generation. It was this evolutionary approach with moderate changes while keeping the design features of the second generation that enabled the general technical progress of recent decades to be combined with the practical operating experience and over 40 years of international reactor safety research.

The continued improvement of safety, resource conservation, and economic efficiency remain at the center of the developmental work.

Nuclear power plants of the third generation are expected to confine the effects of accidents, even including core meltdown, to their immediate vicinity. The so-called passive safety systems are an innovative development of the past few decades. They are based on the physical laws of Nature, such as gravity. Unlike active safety systems, passive systems need no pumps or motor-actuated valves and also function without external energy supply. In the event of a severe malfunction in the power plant, they fulfill their function without any intervention. Many of the reactor systems of the third generation come with the so-called passive safety systems.

The reactor systems of the third generation have the following in common:

- standardized plant design to reduce consenting periods, capital expenditure, and the construction period;
- optimized design simplifying plant operation and maintenance;
- optimized man–machine interfaces;
- shorter outage periods and hence higher availability factors;
- longer operating lifetime, typically 60 years;
- increased efficiency and improved nuclear fuel utilization;
- further reduction of nuclear core meltdown probability and ''virtual elimination'' of accidents, which result in major radioactive releases early on;
- further increase in plant safety, also as protection against external events, for example, improved protection against

aircraft crash, also of large passenger aircraft.

Vendors from Europe, North America, East Asia, and Russia offer nuclear power plants of the third generation in the world market today. Most of these plants are more or less innovative advanced developments of the LWRs of the second generation.

In 2001, nine nations – Argentina, Brazil, Canada, France, Japan, the Republic of Korea, South Africa, England, and the United States – joined forces in the Generation IV International Forum (GIF) to prepare the fourth generation of nuclear systems for the second half of the twenty-first century by way of joint research. Switzerland, Euratom, China, and Russia also joined in the following years.

The first survey produced over 130 responses with conceivable systems that were rated based on the following criteria: the systems should be safe and reliable and should permit economically viable electricity generation, conserve resources, and produce a minimum amount of waste. Particular importance was given to concepts enabling uranium to be used as fuel in a sustainable manner. The required fuel cycle should also be proliferation-proof.

Six reactor concepts and/or technology lines resulted from the assessment, which the forum intends to explore together: one for LWRs with supercritical steam conditions, one for ultra-high-temperature reactors, three for fast reactors, and one for molten salt reactors.

As the market demand in the next few decades is not foreseeable today, Gen-IV-reactor concepts should also develop new markets apart from electricity generation, for example, hydrogen production, process heat for chemical plants, seawater desalination, and actinide transmutation.

Glossary

Actinide: The actinide series encompasses the elements with atomic numbers from 89 to 103. Actinides are radioactive and typically have long half-lives. They are therefore significant in wastes arising from nuclear fission, for example, spent fuel.

Activity: The number of disintegrations per unit time inside a radioactive source. The SI unit of "activity" is becquerel, Bq.

AGR: See gas-cooled reactor

Becquerel: The SI unit of intrinsic radioactivity in a material. One Bq indicates one disintegration per second and is thus the activity of a quantity of radioactive material that averages one decay per second. (In practice, GBq or TBq are the common units.)

Boiling-Water Reactor (BWR): The boiling-water reactor is a type of light-water nuclear reactor (LWR) where water is allowed to boil in the core thus generating steam directly in the reactor vessel.

Burn-Up: A measure of thermal energy released by nuclear fuel relative to its mass, measured typically in megawatt days per ton of fuel ($MWd\,t^{-1}$).

CANDU: The CANadian Deuterium Uranium reactor that is moderated and cooled by heavy water. This is the most common heavy-water reactors.

Chain Reaction: A chain reaction occurs when one nuclear reaction causes an average of one or more nuclear reactions, thus leading to a self-propagating series of these reactions.

Cladding: The metal tubes containing oxide fuel pellets (cf. zircaloy).

Control Rods: Devices to absorb neutrons so that the chain reaction in a reactor core may be slowed or stopped by inserting

them further, or accelerated by withdrawing them.

Coolant: The liquid or gas used to transfer heat from the reactor core to the steam generators or directly to the turbines.

Core: The central part of a nuclear reactor containing the fuel elements and any moderator.

Critical Mass: The smallest mass of fissile material that will support a self-sustaining chain reaction under specified conditions.

Criticality: Condition of being able to sustain a nuclear chain reaction.

Decommissioning: Dismantling of a nuclear power plant and decontamination of the site to a state no longer requiring protection from radiation for the general public and making the site available for unrestricted use.

Defence in Depth: In nuclear safety, this denotes the practice of having multiple, redundant, and independent layers of safety for the single, critical point of failure: the reactor core. The aim is to reduce the risk of harm to individuals or the public at large in operational states and, for some barriers, in accident conditions.

Delayed Neutrons: Neutrons emitted after a nuclear fission event, by one of the fission products (or actually, a fission product daughter after beta decay), anytime from a few milliseconds to a few minutes after the fission event. These enable control of fission in a nuclear reactor.

Dose: The energy absorbed by tissue from ionizing radiation. It is equal to the energy deposited per unit mass of medium, which may be measured as joules per kilogram and represented by the equivalent SI unit gray (Gy). Further calculation can be performed to find the equivalent dose for whole-body external exposure, the effective dose for partial-body external exposure, or the committed dose for internal exposures. These adjusted doses, measured in units of sievert (Sv), are much more representative of the stochastic risks to human health.

Enriched Uranium: Uranium in which the proportion of U-235 (to U-238) has been increased above the natural 0.7%.

Fast Breeder Reactor (FBR): A fast neutron reactor configured to produce more fissile material than it consumes, using fertile material such as depleted uranium in a blanket around the core.

Fast Neutron: neutron released during fission, traveling at very high velocity (15 000 km s^{-1} and more) and having high energy.

Fast Neutron Reactor: A fast neutron reactor or simply a fast reactor is a category of nuclear reactor in which the fission chain reaction is sustained by fast neutrons. Such a reactor needs no neutron moderator but must use fuel that is relatively rich in fissile material when compared to that required for a thermal reactor.

Fission Products: Nuclear fission products are the atomic fragments left after a large atomic nucleus fission event. Typically, a large nucleus like that of uranium fissions by splitting into two smaller nuclei. The two smaller nuclei are the ''fission products.''

Fuel Assembly: Structured collection of fuel rods or elements, the unit of fuel in a reactor.

Gas-Cooled Reactor (GCR): In the gas-cooled reactor, the moderator used is graphite. Inert gas, for example, helium or carbon dioxide, is used as the coolant. The fuel consists of natural uranium metal clad with an alloy of magnesium known as *Magnox* (thus the name for the reactor type). The newer advanced gas-cooled

reactors (AGRs) use a slightly enriched uranium dioxide clad with stainless steel. Carbon dioxide is the coolant gas used.

Heavy-Water Reactor: A reactor that uses heavy water as its moderator, for example, CANDU, which is a pressurized-heavy-water reactor.

High-Level Waste: Extremely radioactive fission products and transuranic elements in used nuclear fuel (heat-generating waste).

Intermediate-Level Waste: Radioactive waste that requires shielding to protect people nearby, but not cooling (generally waste with negligible heat development).

Light Water: Ordinary water (H_2O) that is distinct from heavy water.

Light-Water Reactor (LWR): A common nuclear reactor cooled and usually moderated by ordinary water. It is a generic designation including BWR and PWR types.

Low-Level Waste: Radioactive waste that can be handled safely without shielding (waste with negligible heat development).

Moderator: A material such as light or heavy water or graphite used in a reactor to slow down fast neutrons by collision with lighter nuclei so as to expedite further fission.

Natural Uranium: Uranium with an isotopic composition as found in Nature, containing 99.3% U-238, 0.7% U-235, and a trace of U-234. It can be used as fuel in heavy-water-moderated or graphite-moderated reactors.

Plutonium: A transuranic element formed in a nuclear reactor by neutron capture. It has several isotopes, some of which are fissile and some of which undergo spontaneous fission, releasing neutrons. About one-third of the energy in a light-water reactor comes from the fission of Pu-239, and this is the main isotope of value recovered from reprocessing used fuel.

Pressurized-Water Reactor (PWR): The most common type of light-water reactor (LWR), it uses water at very high pressure in a primary circuit and steam is formed in a secondary circuit.

Radionuclide: A radioactive isotope of an element.

Reactor Pressure Vessel: The main steel vessel of a nuclear reactor containing the reactor fuel, moderator, and coolant under pressure.

Sievert (Sv): Unit indicating the biological damage caused by radiation dose.

Spent Fuel: Used fuel assemblies removed from a reactor after several years of use and treated as waste.

Temperature Coefficient: The change in reactivity due to changed temperature in the fuel. A negative temperature coefficient means that the reactivity is diminished as the temperature rises, so as to reduce the rate of fission and hence reduce the temperature – a natural negative feedback.

Thermal Reactor: A reactor in which the fission chain reaction is sustained primarily by slow neutrons, hence requiring a moderator.

Transmutation: Changing atoms of one element into those of another by neutron bombardment, causing neutron capture and/or fission. In an ordinary reactor neutron, capture is the main event, whereas in a fast reactor, fission is more common, and therefore, it is best for dealing with actinides. Fission product transmutation is by neutron capture.

Transuranic Element: A very heavy element formed artificially by neutron capture. Neptunium, plutonium, americium, and curium are the best known.

Uranium (U): A mildly radioactive element with two isotopes that are fissile (U-235 and U-233) and two that are fertile (U-238 and U-234). Uranium is the basic fuel of nuclear energy.

Void Coefficient: The change in reactivity due to the boiling of coolant or moderator in the actual core. A negative void coefficient means that the reactivity is diminished, and the balance of the chain reaction is affected so as to reduce the rate of fission and hence reduce the temperature - a natural negative feedback.

Zircaloy: Zirconium alloy used as a tube to contain uranium oxide fuel pellets in a fuel rod (part of a reactor fuel assembly).

References

1. RWE Corporate (2007) power: perspektiven 2007: Die Kernenergie und ihr Beitrag zu einer nachhaltigen Energieversorgung, Herausgeber RWE Power AG, *www.rwe.com* (accessed 16 February 2013).
2. Mandel, H. (1966–1968) Einführung in die Reaktortheorie, Teil 1 und Teil 2; Vorlesung an der RWTH Aachen.
3. Smidt, D. (1976) *Reaktortechnik*, Grundlagen, Band **1**, Anwendungen, Wissenschaft und Technik, Band **2**, G. Braun, Karlsruhe.
4. Emendörfer, D. and Höcker, K.H. (1993) *Theorie der Kernreaktoren Bibliographisches Institut and F. A. Brockhaus*, BI-Wissenschaftsverlag, pp. 1982–1993.
5. Loth, R. (2001) Vorlesungsskript Energietechnik III (Reaktortechnik), Technische Universität Darmstadt.
6. Volkmer, M. (2007) *Basiswissen Kernenergie*, Informationskreis KernEnergie, ISBN: 3-926956-44-5, *www.kernenergie.de*.
7. World Nuclear Association (2012) Public Information Service, London, United Kingdom, *http://world-nuclear.org* (accessed 16 February 2013).
8. International Atomic Energy Agency (2006) Fundamental Safety Principles, IAEA Safety Standards Series No. SF-1, *www.iaea.org* (accessed 16 February 2013).
9. The Federal Office for Radiation Protection (2009) Decommissioning Guideline, Germany, *www.bfs.de* (accessed 16 February 2013).

Further Readings

In this accessible overview, the author provides an authoritative account of the key facts about nuclear energy: Ferguson, C.D. (2011) *Nuclear Energy: What Everyone Needs to Know*, Oxford University Press, ISBN: 978-0199759460.

The World Nuclear University Primer on Nuclear Energy in the 21st Century is an authoritative resource for educators, students, policy-makers and interested lay-people alike. With balanced and accessible text, it provides an introduction to nuclear science for the non-specialist, a valuable account of many aspects of nuclear technology, including industry applications, answers to public concerns about nuclear power including safety, proliferation, radiation and waste and up-to-date data and references: Hore-Lacy, I. (2012) *Nuclear Energy in the 21st Century: World Nuclear University Primer*, 3rd edn, World Nuclear University Press, ISBN: 978-0955078453.

This comprehensive text provides an introduction to basic nuclear physics, including nuclear decays and reactions and nuclear structure, while covering the essential areas of basic research and practical applications: Krane, K.S. (1987) *Introductory Nuclear Physics*, John Wiley & Sons, Inc., ISBN: 978-0471805533.

Lamarsh, J.R. and Baratta, A.J. (2011) *Introduction to Nuclear Engineering*, 4th edn, Prentice Hall, ISBN: 978-0132764575
The text is designed for junior and senior level Nuclear Engineering students:.

''Nuclear Energy'' is a popular text on basic nuclear physics, systems, and applications of nuclear energy: Murray, R.L. (2009) *Nuclear Energy: An Introduction to the Concepts, Systems, and Applications of Nuclear Processes*, 6th edn, Butterworth-Heinemann, ISBN: 978-0123705471.

An ideal introduction to the fundamentals of nuclear science and engineering, this book presents the basic nuclear science needed to understand and quantify an extensive range of nuclear phenomena: Shultis, J.K. and Faw, R.E. (2007) *Fundamentals of Nuclear Science and Engineering*, 2nd edn, CRC Press, ISBN: 978-1420051353.

21
Generation IV Nuclear Reactors

Thomas Schulenberg

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

21.1
Introduction

Generation IV systems under development in the framework of the Generation IV International Forum (GIF) aim at improving the future contribution and benefits of nuclear energy utilization. Nuclear power technology has evolved through roughly three generations of system designs: the first generation of prototypes and first-of-a-kind units implemented during the period 1950–1970; the second generation of industrial power plants built from 1970 to the turn of the century, most of which are still in operation today; and the third generation, usually called *Generation III or III+*, of evolutionary advanced reactors built after 2000, which incorporate technical progress based on lessons learnt through more than 12 000 reactor-years of operation.

21.1.1
Goals of the Generation IV International Forum

The trajectory from current nuclear systems to Generation IV systems is described in ''A Technology Roadmap for Generation IV Nuclear Energy Systems'' [1]. The development of Generation IV systems is based on the use of advanced technologies and designs to improve reactor and fuel cycle performance as compared with current nuclear systems and further increases the attractiveness of nuclear energy. Additionally, Generation IV systems should offer a potential for new applications compatible with expanded use of nuclear energy, in particular, in the fields of process heat supply, water desalination, and hydrogen production.

Eight goals have been defined for Generation IV systems in four broad areas: sustainability, economics, safety and reliability, and proliferation resistance and physical protection. These highly ambitious goals are shared by a large number of countries as they aim at responding to economic, environmental, and social requirements of the twenty-first century. They set a framework and identify targets for focusing research and development (R&D) efforts on the field of nuclear energy:

- **Sustainability-1**: Generation IV nuclear energy systems shall provide sustainable energy generation that meets clean air objectives and provides long-term availability of systems and effective fuel utilization for worldwide energy production.
- **Sustainability-2**: Generation IV nuclear energy systems shall minimize and

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

manage their nuclear waste and notably reduce the long-term stewardship burden, thereby improving protection for public health and the environment.

- **Economics-1**: Generation IV nuclear energy systems shall have a clear life-cycle cost advantage over other energy sources.
- **Economics-2**: Generation IV nuclear energy systems shall have a level of financial risk comparable to other energy projects.
- **Safety and reliability-1**: Generation IV nuclear energy systems operations shall excel in safety and reliability.
- **Safety and reliability-2**: Generation IV nuclear systems shall have a very low likelihood and degree of reactor core damage.
- **Safety and reliability-3**: Generation IV nuclear energy systems shall eliminate the need for offsite emergency response.
- **Proliferation resistance and physical protection**: Generation IV nuclear energy systems shall increase the assurance that they are very unattractive and the least desirable route for diversion or theft of weapons-usable materials, and provide increased physical protection against acts of terrorism.

In the framework of the ongoing GIF activities, the goals are guiding the joint R&D efforts undertaken by the members. The challenges in meeting the GIF goals are intended to stimulate innovative R&D covering all aspects of reactor and fuel cycle facility design and implementation.

In the light of the ambitious nature of the goals, international collaboration is considered essential to progress efficiently toward the development of Generation IV systems, keeping open multiple options and avoiding premature downselection that may result from the lack of adequate human, financial, and equipment resources at the national level.

21.1.2 Generation IV Systems

The goals adopted by GIF have served as a basis for identifying and selecting six nuclear energy systems, which are particularly promising. The six selected systems cover a variety of reactor and fuel cycle technologies, including thermal and fast neutron spectra, closed and open fuel cycles, and include a wide range of sizes from very small to very large reactors. Depending on their respective technology preparedness levels, the Generation IV systems were envisaged to become commercially available in the period between 2015 and 2030 or beyond. Table 21.1 summarizes the main characteristics of the six Generation IV systems, in alphabetic order, which are described in more detail in Section 21.2.

These six different systems address the whole variety of future market requests, which are conceivable today for nuclear energy consumption. They go far beyond the conventional technologies of Generation III reactors but, as outlined later, they are often based on innovative concepts that were studied and even built in the past. Thus, they are usually not entirely new but rely as well, to some extent, on proven technologies. Today, one can hardly predict when the market needs for such systems will arise, and the envisaged R&D schedules should not be taken as a forecast for the future demands of such systems.

At least three of these systems are based on a fast neutron spectrum, for which the following section provides some introductory remarks.

Table 21.1 Overview of generation IV systems.

System	Neutron spectrum	Coolant	Temperature (°C)	Fuel cycle	Size (MW_e)	Main uses
GFR (gas-cooled fast reactor)	Fast	Helium	850	Closed	200–1200	Electricity and hydrogen production
LFR (lead-cooled fast reactor)	Fast	Pb or PbBi	550–800	Closed	50–150, 300–600, 1200	Electricity and hydrogen production
MSR (molten salt reactor)	Epithermal	Fluoride salts	700–800	Closed, *in situ*	1000	Electricity and hydrogen production
SFR (sodium-cooled fast reactor)	Fast	Sodium	550	Closed	300–1500	Electricity production
SCWR (supercritical-water-cooled reactor)	Thermal or fast	Water	500–600	Open or closed	1000–1500	Electricity and hydrogen production
VHTR (very high temperature gas reactor)	Thermal	Helium	1000	Open	250	Electricity and hydrogen production

21.1.3 Why Fast Reactors?

The present light water reactors have proven to be a very competitive, reliable, and safe technology for electricity production with almost no CO_2 emissions and small environmental impact. However, they use less than 1% of the uranium available in Nature in an open fuel cycle. According to the OECD/NEA (Organisation for Economic Co-operation and Development/Nuclear Energy Agency) Redbook 2011 [2], the identified and presently economically extractable worldwide resources of uranium will be sufficient for more than 100 years, taking into account especially the energy projections of the Asian region. In order to get long-term sustainability with nuclear energy from fission, new technological solutions improving the usage of this natural resource by at least 50 times had been developed since the 1950s. The new technology is based on the combination of fast neutron systems with multirecycling of the fuel in advanced, closed fuel cycles. Using fast reactors also means build up of significant reprocessing capability.

The physical mechanism behind this idea is the conversion of the normally

useless U-238, having a low fission cross section, to Pu-239, which is fissionable even with thermal neutrons. In a reactor with a thermal neutron spectrum, U-238 is absorbing neutrons converting it to U-239, which in turn decays rapidly to Np-239 and Pu-239.

When Pu-239 is fissioned with thermal or fast neutrons, it produces three or more fast prompt neutrons with 1–5 MeV to continue the chain reaction. The ratio of emitted fast neutrons to the number of absorbed neutrons is called the *regeneration factor*, η. It depends on the energy of the neutron causing fission as shown in Figure 21.1.

If one intends to produce more fissionable material than consumed, which is called *breeding* of nuclear fuel, the regeneration factor must be at least 2, so that one neutron can be taken for fission and one neutron can be absorbed and can thus produce new fissionable material. Taking into account the absorption in structural core material or other nonfuel material, as well as neutron leakage from the core, the regeneration factor must be even greater than 2. Figure 21.1 shows that Pu-239, above about 10^5 eV neutron energy, is ideal for this purpose. It implies, however, that the reactor must be operated with a fast neutron spectrum to get a positive breeding gain.

When neutrons are absorbed in Pu-239, there is production of Pu-240, Pu-241, Pu-242, and even heavier isotopes, which belong chemically to the group of minor actinides. Most of these isotopes can only be fissioned with fast neutrons as shown in Figure 21.2 in a comparison of a reactor with fast neutron spectrum with a conventional light water reactor having a thermal neutron spectrum. This means that multirecycling in a light water reactor as MOX fuel (U-Pu mixed oxide fuel) is accumulating Pu isotopes such as Pu-240 or Pu-242, also creating minor actinides such as Am and Cm. Some of

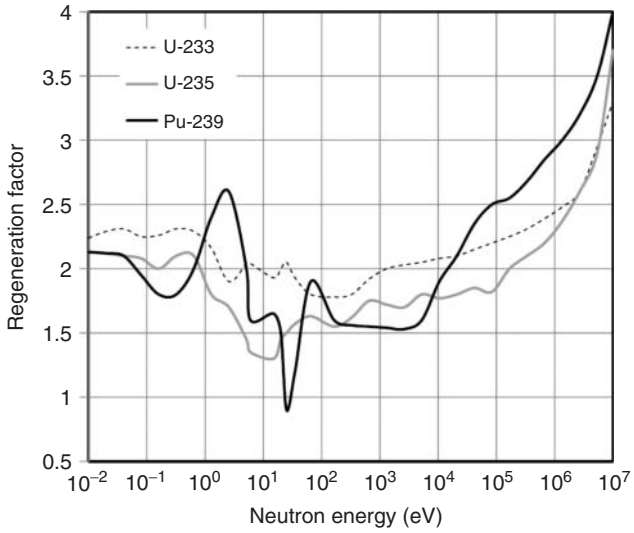


Figure 21.1 Regeneration factor η (number of neutrons emitted per neutron absorbed) as a function of the neutron energy causing fission.

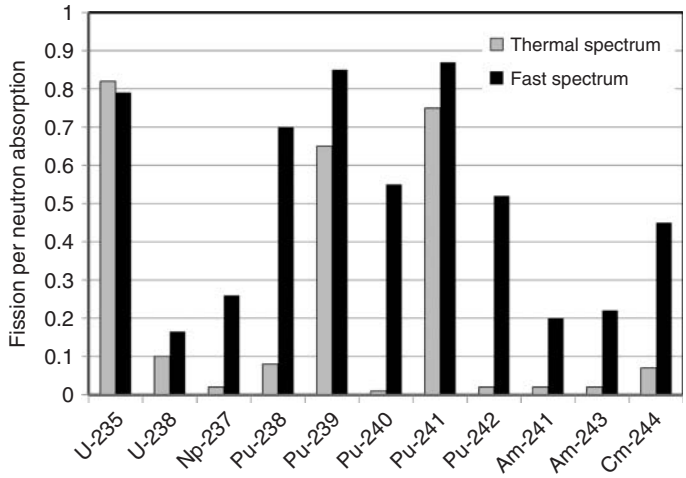


Figure 21.2 Fraction of minor actinides that can be fissioned in thermal and fast reactors.

these isotopes, however, will need up to 1 million years to decay to Pb along the α-decay chains in the final repository, whereas the fission products will remain to be radioactive only for a few hundred years or even far less. Therefore, the radioactive inventory of the spent fuel repository can significantly be reduced, if the recycled Pu is used in fast reactors only, which can fission (or ''burn'') all minor actinides to fission products. As a conclusion, the concept of sustainable nuclear energy requires fast reactors and a closed fuel cycle with reprocessing of the spent fuel. It consumes U-238 from spent fuel and depleted uranium from enrichment facilities, which enables to utilize nuclear energy at least for the next 5000 years. On the other hand, this concept reduces the radioactive inventory of the final repository by more than a factor of 100, as only fission products and about 0.1% losses of minor actinides in the reprocessing plant will enter the final repository. This concept provides a strong political argument in favor of nuclear power. However, there are still a number of challenges left with fast reactors.

Most fast reactors, which were built in the past, were using liquid Na as coolant. Different from water, which slows down the neutron energy by elastic scattering with the hydrogen nucleus to less than 0.1 eV, the heavier Na nucleus is taking far less energy from the neutron with each collision. The neutron spectrum gets even harder, that is, it is shifting to higher energies, if Na starts boiling or is lost during an accident. This, however, causes a safety issue, as a harder neutron spectrum increases the regeneration factor (see Figure 21.1) and thus the reactivity of the core; in other words, the reactivity of a Na-cooled reactor has a positive void coefficient, at least locally. Even though fast reactors with a local positive void coefficient had been licensed in the past, it is doubtful if they can ever be licensed again in a western country.

Therefore, the following methods and ideas have been brought forward to overcome this issue.

1. Development of advanced core concepts with core zones of different axial heights, large upper coolant plena,

and internal axial blankets with a sufficiently high neutron leakage such that an increase in neutron energy increases the neutron leakage from the core, which overcompensates for the increase in the regeneration factor and thus the positive void coefficient.
2. Use of a heavy liquid metal such as Pb or PbBi with very high boiling temperature such that boiling of the coolant can reasonably be excluded. However, the entrainment of gas from other sources must still be excluded.
3. Use of a pressurized gas as a coolant such that the core is almost voided and a loss of coolant will hardly increase the reactivity, but the Doppler effect decreases the reactivity with fuel temperature instead.
4. Use of a softer neutron spectrum by adding some moderator to the core. Such an epithermal core can easily achieve a negative void coefficient throughout the entire burn-up cycle but reduces the breeding gain (produced/consumed Pu-239) to less than 1 because of a smaller regeneration factor.
5. Breeding of U-233 from Th-232 in a thermal reactor. Thorium is available in abundance and the production of minor actinides such as Pu, Am, or Cm is physically excluded in such a reactor. This concept requires, however, an initial inventory of fissile material, as U-233 does not exist in natural uranium.

Another general issue of fast reactors is that Pu-239 could also be misused for nuclear weapon production. The main strategy to overcome this proliferation concern is to remove the breeding blankets containing depleted U or poisoning the blankets with minor actinides. The key idea of "breeding" is however jeopardized by this procedure.

This way, the fuel is kept in the core to a burn-up of more than 100 MWd per kilogram of heavy metal (MWd kg_{HM}^{-1}). While almost pure Pu-239 is produced within the first 3 months, a higher burn-up will produce significant amounts of Pu-238, as indicated in Figure 21.3. This isotope provides physical protection to plutonium as it decays with a half time of 87.7 years, causing significant heat production such that a nuclear weapon cannot be built without the risk of self-ignition. Therefore, unlike earlier reactor

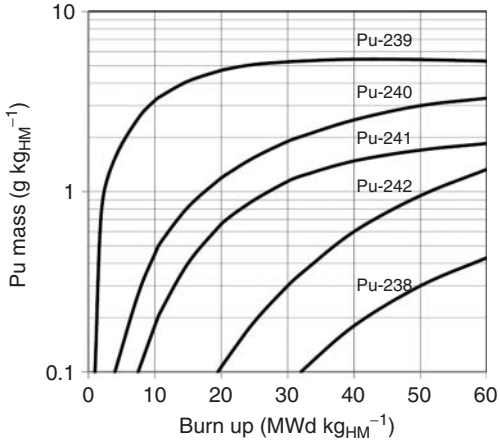


Figure 21.3 Production of plutonium isotopes during burn-up in light water reactors.

concepts, latest core concepts are designed without removable blanket layers or they are even encapsulated to be operated for many years to become proliferation resistant.

Besides the above-mentioned political reasons to build a fast reactor, there should also be economic reasons in the future. Liquid-metal-cooled reactors are characterized by a very high power density of more than $200\,MW_{th}\,per\,m^3$ of the reactor core, which makes the reactor concept very compact. If the reactor does not need to be opened for many years because of a high burn-up of more than even $200\,MWd\,kg_{HM}^{-1}$, which can be achieved in principle with a fast reactor, the availability of the nuclear power plant can be improved and the service costs for refueling can be decreased. As long as the plant erection costs of fast reactors are more expensive than those of light water reactors, they will remain to be less attractive, provided sufficient nuclear fuel will be available in natural uranium. Therefore, a major objective of the R&D programs of the GIF is to simplify the design, reduce the number of required components, and ease inspection and repair.

Recycling of spent fuel, on the other hand, requires reprocessing technologies that separate not only U and Pu but also minor actinides such as Am and Cm from fission products. (Np is not contributing a significant radioactive inventory to the final repository so that it would not need to be separated from fission products.) Chemical separation techniques for Am and Cm, called *partitioning* in this context, are under development. Moreover, a novel fuel has to be manufactured that includes these minor actinides to be irradiated and fissioned in a fast reactor. These additional isotopes, however, for example, Cm-244 with its decay time of only 18.1 years, result in significant dose rates in the reprocessing and fuel manufacturing facilities. Remote handling will be required for the entire fuel cycle, which can increase the costs of the fuel significantly. It might be advisable, therefore, to cool down the fuel in an interim storage for a longer period before reprocessing it.

Section 21.2 discusses these issues in more detail using the examples of Na-, Pb- or PbBi-, and He-cooled fast reactors.

21.2 Outline of the Generation IV Reactor Concepts

21.2.1 Supercritical-Water-Cooled Reactor (SCWR)

The supercritical-water-cooled reactor (SCWR) is a high-temperature, high-pressure, water-cooled reactor that operates above the thermodynamic critical point of water (374 °C, 22.1 MPa). Two design options – pressure vessel and pressure tube – have been proposed for the SCWR.

A potential plant could have a power of around $1000–1500\,MW_e$, an operating pressure of 25 MPa, and a reactor outlet temperature of 500–625 °C. Owing to the low density of supercritical water, additional moderator must be added to the coolant to slow down the neutron energy. Compared with a boiling water reactor, the SCWR is considerably simplified because the coolant does not change phase (boil) in the reactor. However, safety features similar to those of advanced boiling water reactors may be incorporated in the design.

The main advantage of SCWR is its improved economics because of the higher thermodynamic efficiency (up to about 50 vs 34% for present light water reactors) and

the potential for plant simplification. While the focus of SCWR has been on a reactor with a thermal neutron spectrum, a fast version has also been studied for a closed fuel cycle.

Thus, the SCWR is a long-term vision for light water reactors, but it requires significant development in a number of technical areas. As a general advantage, however, the SCWR benefits from worldwide experience with light water reactors as well as from an established technology for supercritical water power plants in the coal power industry.

Moreover, different from all other Generation IV concepts, the SCWR can be developed evolutionarily from current water-cooled reactors. Looking at the thermal efficiencies of fossil fired power plants in the past 50 years, which increased from 36% in 1960 to almost 50% at present, light water reactors have a huge potential for further improvement. The main driver for these improvements in fossil fired power plants had been the increase of live steam temperature and pressure, which are supercritical since around 1990.

Therefore, significant cost savings and high thermal efficiencies will need high core outlet temperatures of 500 °C or more, such that the core produces superheated steam. Steam generators, droplet separators, and primary pumps can then be omitted. This allows a significantly smaller containment and smaller turbines, condensers, reheaters, and feedwater pumps than those of today's light water reactors.

The higher core outlet temperature, on the other hand, requires significant changes in the core design concept. Like with coal-fired boilers, the increased steam temperature can only match with today's cladding material limits at around 620–630 °C if the coolant is mixed twice while heated up from 280 °C feedwater temperature. Schulenberg and Starflinger [3] propose to use the coolant first as moderator water in gaps between fuel assembly boxes and in water channels inside these assemblies. After homogeneous mixing in a plenum underneath the core, the preheated water is used as coolant in assemblies running upward, being heated up to around 390 °C. As this heat generated passes the pseudocritical point, this core section is called the *evaporator*. Above the core, the coolant needs to be mixed uniformly to eliminate hot streaks. The first superheater with downward flow surrounds the evaporator. The second outer mixing chamber underneath the core homogenizes the coolant again before it is finally heated up in the second superheater at the core's periphery. A European consortium adopted this idea to work out a reactor and plant design concept called the high-performance light water reactor (HPLWR). The consortium's latest reactor concept is sketched in Figure 21.4.

In Japan, an SCWR concept with 2740 MW thermal power and 1217 MW net electric power has been worked out by the University of Tokyo, as summarized by Oka *et al.* [4]. The core design concept is similar to the HPLWR but with two heat-up steps only and hotter peak cladding temperatures. As in the HPLWR, control rods are inserted from the top to run inside the water channels. This core has been analyzed for peak cladding temperatures during a burn-up cycle. The hottest fuel rod reached more than 730 °C, not including tolerances and allowance for operation yet.

If the water channels in the core are omitted, the neutron spectrum becomes fast so that the SCWR could be used for breeding Pu and for burning minor actinides from spent fuel. Care has to be taken, however, that the void reactivity coefficient of the coolant is negative in

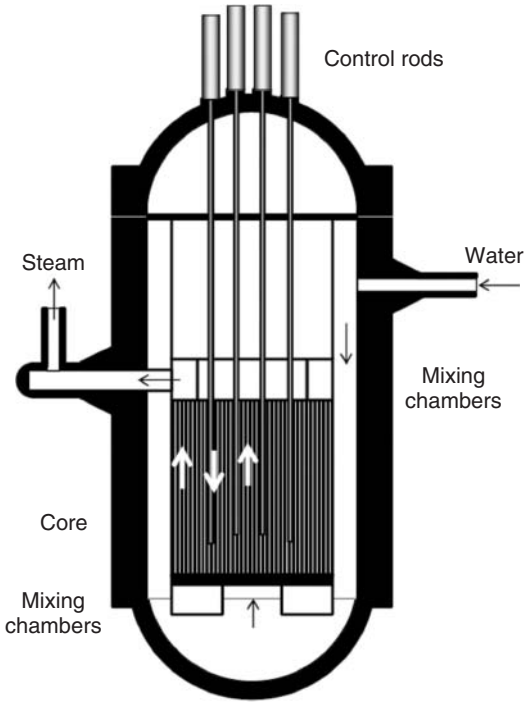


Figure 21.4 Concept of a pressure-vessel-type supercritical-water-cooled reactor (SCWR) cooled and moderated with supercritical water [3].

each fuel assembly, as the coolant changes its density significantly in the core. Oka *et al.* [4] have reported about a reactor and core design with fast neutron spectrum, in which a heterogeneous arrangement of seed assemblies (i.e., with Pu fuel) with upward and downward flows, and blanket assemblies with downward flow and with ZrH layers, has been optimized such that the void coefficient is always negative. The ZrH layers as well as the downward flow of colder feedwater cause some moderation of the neutrons, which shifts the neutron spectrum to lower energies. The resultant epithermal spectrum not only makes the regeneration factor (Figure 21.1) less sensitive to a spectrum shift, causing a negative void coefficient, but also reduces the breeding gain below zero. Such a fast SCWR, therefore, can rather be used for burning spent fuel than for breeding efficiently new fuel. Figure 21.5 illustrates this assembly design concept for a fast neutron spectrum.

Multiple heat-up steps with intermediate coolant mixing are the easiest to be optimized in a pressure tube reactor. Such reactors mix the coolant in headers anyway, so that additional heat-up steps would not increase the complexity of the design. Yetisir *et al.* [5] proposed a new pressure tube reactor with supercritical water based on the CANDU (CANadian Deuterium Uranium) reactor design of AECL, Canada. Different from a pressure vessel reactor, the coolant pressure inside the tubes may differ significantly, so that even a steam reheat at lower pressure could be feasible inside the reactor.

The improvements from the current CANDU concept to the Generation IV concept with supercritical water, however, imply a lot of significant changes. The Generation III CANDU reactor has a closed

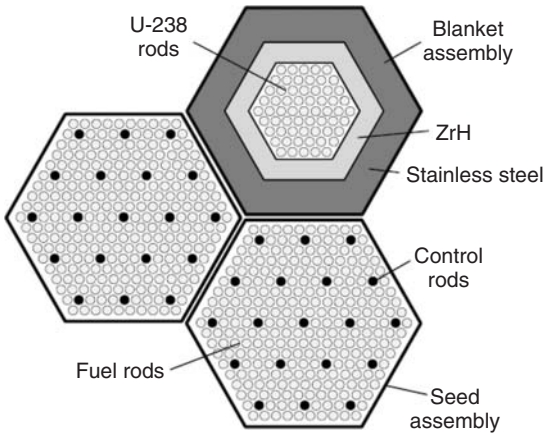


Figure 21.5 Concepts of assembly design for pressure-vessel-type SCWR with fast neutron spectrum [4].

primary loop, whereas the supercritical pressure version is open at least to the high-pressure turbine, like a boiling water reactor. Moreover, the pressure tubes should rather be vertical as the conventional, horizontal arrangement would cause thermal stratifications and risk bending of the tubes. Online refueling, a general feature of CANDU reactors, becomes difficult as the wall thickness of the tubes at supercritical pressure does not leave enough space between the tubes for suitable end flanges.

Hotter components such as the claddings of fuel rods and assembly boxes need to be resistant against corrosion, creep, and radiation damage. Tests with zircaloy at supercritical pressure and 500 °C temperature showed severe oxidation just after a few hundred hours, so that stainless steel claddings with more than 20% Cr will need to be developed instead and qualified for this purpose. Another challenging area of technology development is the prediction of heat transfer when the supercritical fluid changes from liquidlike to steamlike properties. Last but not least, it should be remembered that an SCWR has never been built in the past and R&D must start first with small-scale component tests inside nuclear reactors.

21.2.2 Very High Temperature Reactor (VHTR)

The very high temperature reactor (VHTR) is the next generation in the development of high-temperature reactors (HTRs) and is primarily dedicated to the cogeneration of electricity, hydrogen, and process heat for industry. Hydrogen can be extracted from water using thermochemical, electrochemical, or hybrid processes. The reactor is cooled by helium gas and moderated by graphite with a core outlet temperature greater than 900 °C (with an ultimate goal of 1000 °C) to support the efficient production of hydrogen by thermochemical processes. The high outlet temperature makes it attractive also for the chemical, oil, and iron industries. The VHTR has the potential for high burn-up (150–200 MWd kg_{HM}^{-1}), passive safety, low operation and maintenance costs, and modular construction.

Two baseline options are available for the VHTR core, namely, the pebble-bed-type and the prismatic-block-type reactor, both with a thermal neutron spectrum. The fuel cycle will initially be open (i.e., without recycling) with low-enriched uranium fuel and very high fuel burn-up. The system,

however, has the flexibility to apply closed fuel cycles and offers burning of minor actinide fuel later on. The electric power conversion may employ a steam cycle via a steam generator, a direct cycle with a helium gas turbine, or an indirect Brayton cycle with a gas mixture.

The basic technology for the VHTR has been established in former high-temperature gas reactors such as the US Peach Bottom and Fort St. Vrain plants as well as the German AVR and THTR-300 prototypes.

As a general feature, the VHTR does not need an emergency cooling system. If the helium flow through the core is stopped or if helium were lost from the pressure vessel, the ceramic fuel will heat up and the Doppler effect will interrupt the fission chain reaction physically even if the control rods could not be inserted. The fuel will not melt during this accident. This unique feature makes the reactor concept inherently safe against core melt accidents. The thermal power of the HTR, however, is limited then to around 250 MW to keep its inherent safety features against core melt accidents.

A commercial reactor following the two German prototypes was the HTR module with 400 MW thermal power, generating 77 kg s^{-1} live steam at 190 bar and 530 °C, which could be used to produce up to 124 MW electric power. It was designed in 1989 by Siemens Interatom, but was never built. The concept is sketched in Figure 21.6. It served as a basis for the HTR-10 module at the Tsinghua University in China with 10 MW thermal power and 750 °C helium temperature, where helium at a pressure of 3 MPa is heated up from 250 to 700 °C, producing steam in a steam

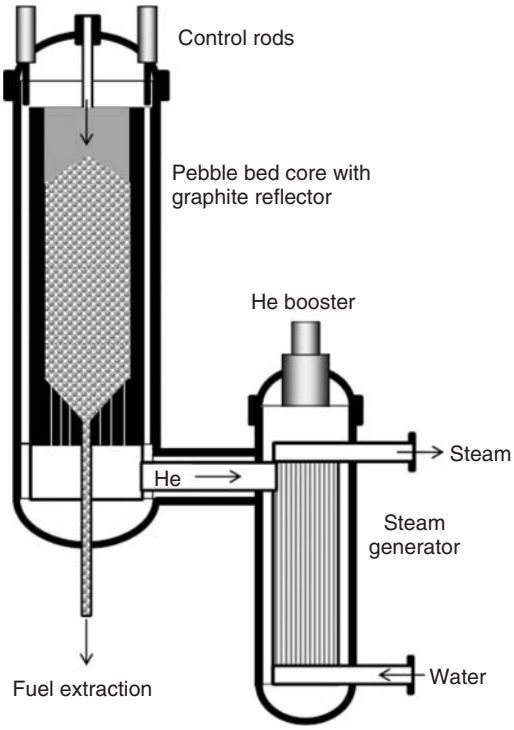


Figure 21.6 Sketch of a pebble bed type very high temperature reactor (VHTR) with steam generator.

generator connected with the reactor. Different from the THTR-300, control rods were inserted into the outer reflector. Construction began in 2000 and the first criticality was achieved in 2003. China intends to scale up this prototype to the HTR-PM (high-temperature reactor pebble bed module) reactor with 250 MW thermal power. Zhang *et al.* [6] have reported more details on the Chinese HTR concepts.

The first HTR with hexagonal prismatic graphite fuel blocks was the 115 MW_{th} reactor built in 1960 at Peach Bottom (USA). It was operated from 1964 to 1974. The Fort St. Vrain reactor with 770 MW thermal power was built a few years later in 1968. It was operated from 1976 to 1989. Like with the THTR-300, a prestressed concrete vessel was chosen. The graphite blocks have the advantage that control rods may run into the core, and the core power and temperature distribution can be predicted more precisely than with the random arrangement of pebbles, but online refueling like with the pebble bed reactor is impossible.

A similar core design can be found in the Japanese HTTR (high-temperature test reactor) with 30 MW thermal power, which heats up a helium gas flow of 12.4 $kg\,s^{-1}$ at 4 MPa from 395 °C to 850 °C. Test runs were also performed with 950 °C core outlet temperature.

A more advanced concept of a VHTR with prismatic graphite blocks is the gas turbine modular helium reactor (GT-MHR) designed by General Atomics (USA) for an electric power of around 300 MW_e and a thermal power of 600 MW_{th}, sketched in Figure 21.7. The reactor and the entire Brayton cycle are integrated into two pressure vessels, closely connected with a coaxial pipe (called *cross vessel*) to exchange helium to

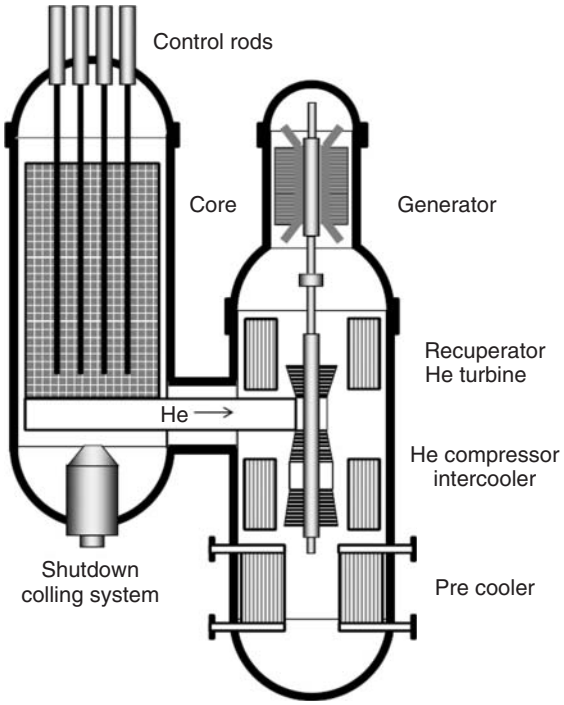


Figure 21.7 Sketch of a very high temperature reactor (VHTR) with prismatic graphite blocks and the closed recuperated Brayton cycle [7].

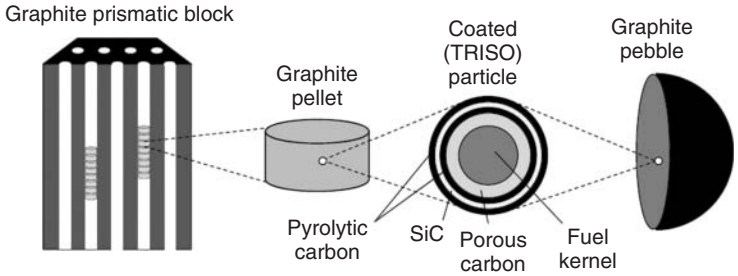


Figure 21.8 Fuel concept with TRISO particles for pebble-bed- or prismatic-block-type reactors.

and from the reactor with the turbines. Baxi *et al.* [7] have reported more details.

Both the design concepts, the pebble bed reactor and the prismatic blocks, use TRISO (tristructural-isotropic) particles to confine the fuel and fission products, as sketched in Figure 21.8. These particles – built from a UO_2 fuel kernel surrounded by a layer of porous carbon to capture the fission gas, as well as three layers of pyrolytic carbon and silicon carbide to keep the fission gas inside – have a diameter less than 1 mm. They are embedded in a graphite matrix from which either spherical fuel elements of 60 mm diameter (the ''pebbles'') or fuel pellets are formed, which are then filled into larger graphite blocks.

Several unique components are needed for the VHTR, including the reactor pressure vessel, intermediate heat exchangers, and the Brayton cycle turbomachinery. The pressure vessels are unique because their size and thickness are larger than those of modern boiling water reactor vessels. Their development includes welding and fabrication methods, as well as the means to assure the thermal emissivity of the outer vessel walls.

For core outlet temperatures up to about 900 °C, it is expected that existing materials can be used; however, temperatures above this, including safe operation during off-design conditions, require the development and qualification of new materials. Current research is focused on (i) graphite for the reactor core and internals; (ii) high-temperature metallic materials for internals, piping, valves, high-temperature heat exchangers, and gas turbine components; and (iii) ceramic composites for control rod, cladding, and other core internals, as well as for high-temperature heat exchangers and gas turbine components. Moreover, the VHTR should enable the production of hydrogen from water, for which processes are being developed. Further research will be needed to couple a nuclear reactor with the hydrogen production process. This involves the thorough analysis of safe and reliable control and operation, including the hazards or risks that each system might pose to the other.

21.2.3 Sodium-Cooled Fast Reactor (SFR)

The sodium-cooled fast reactor (SFR) uses liquid sodium as the reactor coolant, allowing high power density in the core with low coolant volume fraction. While the oxygen-free environment prevents corrosion, sodium reacts chemically with air and water and requires a sealed coolant system.

Plant size options under consideration range from small, 50–300 MW_e, modular reactors to larger plants up to 1500 MW_e. The outlet temperature is 500–550 °C for

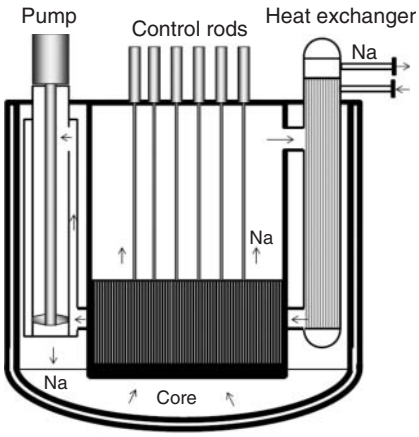


Figure 21.9 General sketch of a pool-type sodium-cooled fast reactor (SFR).

both options, which affords to use materials that had been developed and proven in prior fast reactor programs.

The SFR closed fuel cycle enables recycling of fissile fuel and facilitates management of high-level waste, in particular, plutonium and minor actinides. However, the latter option requires that recycled fuels will be developed and qualified for use. Important safety features of the SFR include a long thermal response time to reactivity changes in the core, a reasonable margin to coolant boiling, a primary system that operates near atmospheric pressure, and an intermediate sodium system between the radioactive sodium in the primary system and the power conversion system. Water/steam and supercritical carbon dioxide were considered as working fluids for the power conversion system to achieve high performance in terms of thermal efficiency, safety, and reliability.

Much of the basic technology for the SFR has been established in former fast reactor programs. Experience goes back to the 1950s to 1960s, when the first experimental reactors were built in the United States, Russia, and France. Germany, Japan, and the United Kingdom followed with prototypes in the 1970s. Details are given, for example, in the IAEA Fast Reactor Database [8]. The next step to commercialization had been demonstration or prototype reactors, of which the French Superphénix with almost 3000 MW thermal power is still the largest unit ever built. Some of these reactors are still running as a fast neutron source for irradiation test.

A general sketch of a pool-type reactor is shown in Figure 21.9. The largest SFR of this type is the Superphénix reactor. The reactor core containing 16–19.7% Pu in the fissile zone had a height of only 1 m and a diameter of 3.7 m, which is surrounded by a large blanket in axial and radial directions. The power density of this active zone exceeds by far that of light water reactors. With a breeding gain of 0.18, the reactor produced 18% more Pu in the blanket than it consumed in the fissile zone. The reactor had a positive void reactivity coefficient, while the overall temperature coefficients were clearly negative. The primary sodium system of such reactor is built as a pool in which the secondary heat exchangers and primary pumps are inserted. This arrangement provides enough thermal inertia of the primary system to start a residual heat removal system in case of a failure of all primary pumps.

The new Indian PFBR and the Russian BN 800 concepts are rather similar. Differences can be found in the emergency decay heat removal system, which is a passive sodium loop connecting a dip cooler inside the pool with an air heat exchanger in case of the PFBR. The residual heat is finally removed by natural air convection in the stack. On the other hand, the decay heat removal system of the Russian BN 800 is connected with the secondary sodium loop.

An example of a pool-type SFR considered for Generation IV is the Korean SFR concept KALIMER with 600 MW electric power. It can reach sodium temperatures of 366–510 °C in the primary pool, 310–495 °C in the secondary sodium loops, and a life steam temperature of 460 °C in the steam cycle, thus resulting in a net efficiency of 38%. The decay heat removal system is designed similar to that in the PFBR, with dip coolers inside the sodium pool and natural convection sodium lines connected to an auxiliary heat exchanger in the air stack. The primary system is built with two pumps and four intermediate heat exchangers. Hahn *et al.* [9] have reported about more details.

An example of a large-sized loop-type reactor for Generation IV is the Japanese JSFR 1500, sketched in Figure 21.10. It consists of two primary coolant loops with a maximum temperature of 550 °C and a sodium inventory of only 1330 t. The secondary loop has a total sodium inventory of 852 t, reaching 520 °C at maximum. The steam generators are built with double-walled straight tubes to minimize the risk of sodium contact with water, and 12% Cr steels are used for both loops. Ichimiya [10] has reported briefly about the latest design status.

An innovative approach for smaller unit sizes to save costs is to use the high burn-up that can be achieved with the SFR for longer refueling intervals. As an ultimate goal, a run time of 30 years without refueling

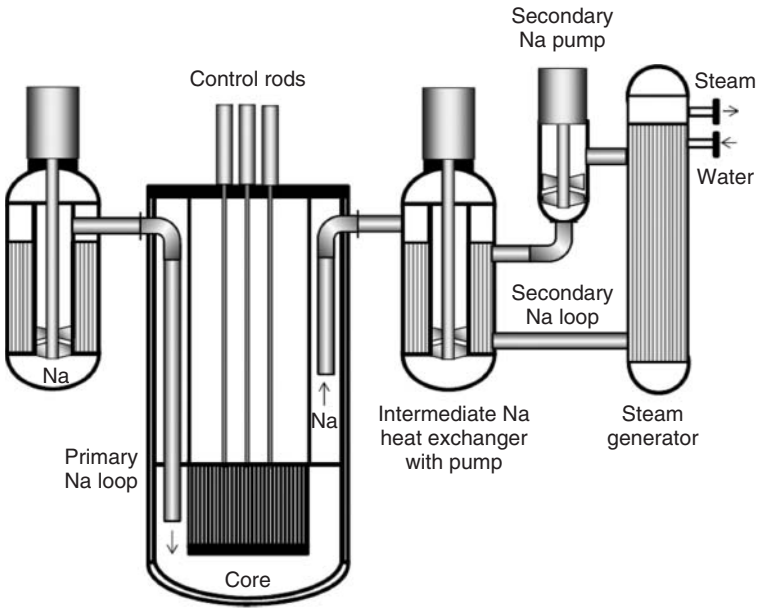


Figure 21.10 Sketch of a loop-type sodium-cooled fast reactor (SFR).

seems feasible. This avoids not only outage and maintenance costs but also the need for on-site fuel handling and storage. Instead of removing and shuffling single assemblies, the entire core or even the entire reactor is exchanged after 30 years, combined with an upgrade of the turbines or other major components in parallel. A small-scale design for such an application is the 4S Nuclear Battery of Toshiba, an encapsulated reactor with 30 MW thermal power producing about 10 MW net electric power. The concept includes an underground reactor, a secondary sodium loop like that in the large reactor versions, and a steam cycle with 10.5 MPa life steam pressure and 453 °C temperature. Passive air cooling of the reactor vessel is sufficient to remove the decay heat, which simplifies the emergency cooling system significantly. Tsuboil *et al.* [11] have reported about more details.

In summary, the SFR is the most advanced fast reactor concept. It has already been built in large, commercial scale and has been operated for many years. Corrosion of fuel claddings or structural materials has never been an issue, not even after 30 years of operation of the EBR II in the United States. There are still safety issues, however, that need special attention. The violent reaction of sodium with air and water requires an intermediate sodium loop. The handling of spent fuel must avoid any contact with air or water, which makes is more complicated. Moreover, the risk of sodium boiling during severe accidents and the increase of reactivity by voiding or any rearrangement of the fuel or structure is an issue of investigation in all programs. However, successful licensing procedures from the past demonstrate that all safety issues could be handled successfully.

Most experience exists with oxide fuels. MOX fuel (U-Pu mixed oxide fuel) has been used in light water reactors for more than 40 years and also in most of the past SFRs. The fuel fabrication and reprocessing techniques (PUREX, plutonium and uranium recovery by extraction) have been well established for many years. An alternative fuel development is aiming at a harder neutron spectrum to improve the incineration of minor actinides and the breeding ratio, improve the thermal conductivity of the fuel and thus to decrease the fuel centerline temperature, and increase the melting temperature of the fuel to allow higher power of fuel rods. Nitride fuel is characterized by a higher thermal conductivity than MOX fuel, allowing a higher linear heat rate of the fuel rods of 700–900 W cm^{-1}. Higher thermal conductivity but a lower melting temperature will be achieved with carbide fuel, which is not compatible, however, with the PUREX process.

21.2.4 Lead-Cooled Fast Reactor (LFR)

The lead-cooled fast reactor (LFR) is an alternative option for a fast neutron spectrum and a closed fuel cycle for efficient conversion of U-238 to fissile fuel. Like the SFR, it can also be used as a burner of minor actinides from spent fuel or as a breeder. An important feature of the LFR is the enhanced safety that results from the choice of a relatively inert coolant, provided issues regarding weight and corrosion of lead can be overcome. It has the potential to meet the electricity needs of smaller, remote sites as well as of larger grid-connected power stations.

The designs that are currently proposed by the GIF as future options are pool-type reactors, as sketched in Figure 21.9 for the SFR. Examples are the modular fast reactor SUPERSTAR, described by Bortot

et al. [12], and the European Lead-cooled System (ELSY), presented by Alemberti *et al.* [13]. The history of lead-cooled reactors, however, started in Russia.

The history of LFR goes back to the use of liquid lead bismuth for fast reactors to power the Russian Alpha Class submarines. Here, like in later land-based applications, the eutectic mixture of 44.5% lead with 55.5% bismuth has been preferred as the coolant to lower the melting temperature to 125 °C. This reduces material problems and the risk of coolant freezing, with the disadvantage, however, that bismuth is rare and expensive and will be activated to build up Po-210 under neutron irradiation. Later commercial applications are aiming therefore at using pure lead with 327 °C melting temperature, such that the operating temperature range could rise to around 400–600 °C, improving the thermal cycle efficiency significantly. Both coolants have in common that the boiling temperature is higher than 1670 °C so that coolant boiling can practically be excluded and a positive void reactivity coefficient would play a minor role. The weight of the coolant requires special seismic provisions.

A small, pool-type reactor using lead bismuth as the coolant is the SVBR designed by SSC IPPE, Russia, for electricity production [14]. The prototype version SVBR 75 reaches 265 MW thermal power and 80 MW gross electric power with core inlet and outlet temperatures of 286/435 °C, respectively. A later commercial version SVBR 100 could run at elevated temperatures and thus reach and increase gross electric power to more than 100 MW. The breeding gain using MOX fuel is slightly positive, which allows running the reactor for 2200 days without refueling or shuffling the assemblies. Instead, the entire core is designed to be removed in total as a cartridge after use.

The steam cycle is made of two natural convection drum boilers, each with six steam generator modules inside the reactor pool. An intermediate loop like that in the SFR is not required.

The safety system is made of active and passive residual heat removal systems. In case of a pump failure, a water tank around the pool vessel can remove the heat without any power supply. In case of steam leakage into the pool, the steam above the liquid lead bismuth condenses in a gas system condenser connected with the gas purification system. In case of a steam generator tube rupture, a membrane opens to spargers in the water pool around the vessel to depressurize the system. The arrangement of steam generators in the vessel is such that steam ingress into the core can physically be excluded, which is important because of a positive void reactivity coefficient that cannot be strictly avoided. Superheating of the core is protected with fusible locks at some control rods. In case of massive corrosion of the vessel, an additional safeguard casing keeps the coolant in place.

A larger reactor concept with 300 MW gross electric power has been designed by Nikiet, Russia. The BREST-300 prototype reactor uses pure lead as coolant. With core inlet and outlet temperatures of 420/540 °C, respectively, the net plant efficiency is expected to be significantly higher than that of SVBR 75/100. Vessel preheating is required to avoid freezing of the pool. It also helps to remove the residual heat in case of a pump failure. More information is found in the IAEA Fast Reactor Data Base 2006 update.

As liquid lead or lead bismuth does not react violently with air or water, the LFR can be simplified significantly compared with the SFR, which reduces the plant erection costs. The intermediate coolant loop can be omitted and the spent fuel can be exposed

simply to air or water. On the basis of the first cost estimates in Russia and Japan, it appears that the LFR will even be cheaper per installed kilowatt electrical power than a pressurized water reactor. Therefore, there is an obvious market interest in this fast reactor technology.

The greatest technical challenge to build such an LFR is the prevention of corrosion of fuel claddings, the reactor vessel, and its internal structures. If a small oxygen concentration is maintained in the liquid metal, it can protect the steel surface by building oxide coatings. The required concentration range, however, is small and an excess oxygen concentration will cause formation of lead oxide crystals that can plug the coolant flow path. Control of the oxygen concentration in this narrow range in the entire, large reactor pool will not be easy. As long as these material issues are not solved, the reliability of the LFR during long-term operation will remain to be an issue.

Like with the SFR, oxide fuel with Pu can be used for the first generation of LFR, and the use of nitride fuel could improve both the breeding gain and the core power density.

21.2.5 Gas-Cooled Fast Reactor (GFR)

The gas-cooled fast reactor (GFR) is a high-temperature helium-cooled fast reactor with a closed fuel cycle. It combines the features of fast-spectrum systems with those of high-temperature systems. The fast spectrum enables more sustainable use of uranium resources and waste minimization through fuel recycling and burning of long-living actinides, and the high temperature affords high thermal cycle efficiency and industrial use of the generated heat, for example, for hydrogen production. The largest reactor studied until now is a $2400\,MW_{th}/1100\,MW_e$ helium-cooled system operating with an outlet temperature of 850 °C using three indirect power conversion systems with a combined cycle power plant. A direct Brayton cycle gas turbine can be considered as another alternative.

The GFR has a long-term vision for highly sustainable reactors that require significant development in a number of technical areas. Unlike the SFR, the GFR does not have the benefits from operational experience worldwide and will require more time to develop. Like the VHTR, however, the GFR does use helium coolant and refractory materials to access high temperatures, thus allowing it to provide process heat.

The GFR is dedicated to actinide management and to the production of electricity and heat. A small positive void reactivity coefficient also exists for this fast reactor in case of loss of the pressurized coolant, but a strong negative reactivity feedback is still due to the Doppler effect. However, the inherent safety features of the HTR with its thermal spectrum and large moderator heat sink are lost and so residual heat removal systems will be required.

Therefore, unlike the VHTR that uses its considerable thermal mass to limit the rise in core temperature during transients, the GFR requires the development of a number of unique subsystems to provide in-depth protection for its considerably higher power density core and low thermal inertia of the coolant. As an example, the core exceeds 1100 °C within 30 s if a helium fan trips but the reactor fails to scram. A fuel pin failure can be expected after 20 s. This requires a robust decay heat removal system with added provisions for natural circulation heat removal, such as a core with a low

pressure drop. Under accidental conditions, radiation effects on structures play a significant role at high-temperature levels. An external core catcher thus seems to be mandatory.

A gastight envelope acting as additional guard containment must be provided to maintain backup pressure in case of large gas leak from the primary system. It can be a metallic vessel that is initially filled with nitrogen slightly at over the atmospheric pressure to reduce the air ingress potential. Dedicated loops for decay heat removal in case of emergency are directly connected to the primary circuit, using cross-duct piping from the pressure vessel, and are equipped with heat exchangers and fans.

Present reactor design concepts that are available in Europe have been summarized by Stainsby *et al.* [15] (Figure 21.11). The joint European project GCFR designed a concept of an experimental reactor with 75 MW thermal power, called *ALLEGRO*. A pressure vessel of about 15 m height with control rods driven from the bottom is connected with a U-tube heat exchanger and a helium fan. The large space above the core is needed for the fuel-handling structures, as the pressure vessel remains closed during refueling to keep the helium inventory inside and thus to cool the core. Three decay heat removal systems, driven by pony motors, with three accumulators, are needed for emergency cooling. They are connected to water heat exchangers that deliver the residual heat in a natural convection driven loop to water pools. The core is operated at 7 MPa and heats up helium from 260 to 560 °C.

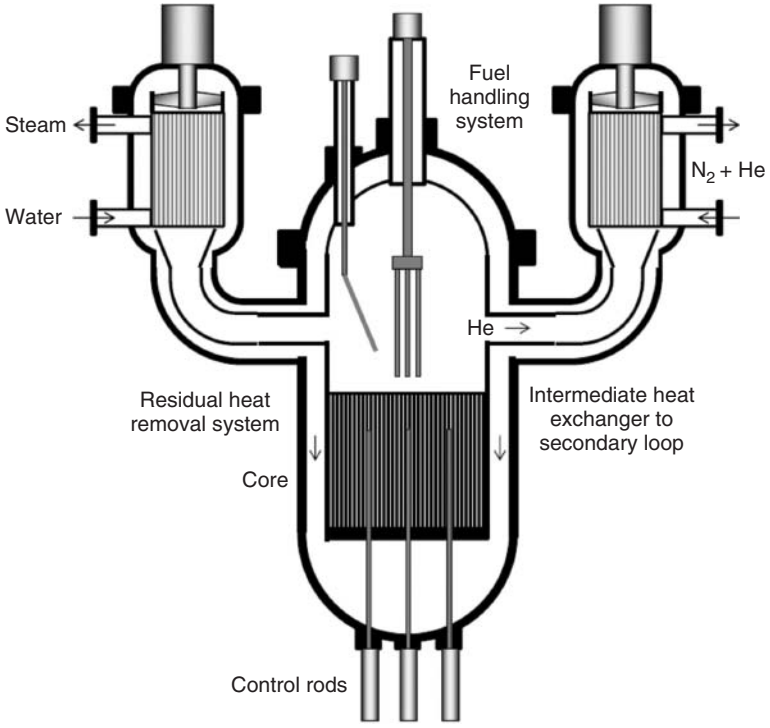


Figure 21.11 Sketch of a gas-cooled fast reactor (GFR) concept [15].

The large-scale reactor with 2400 MW_{th} power is larger than the ALLEGRO design, but includes basically the same design features. Figure 21.11 shows the fuel-handling equipment in the upper half of the reactor and two of the three decay heat removal systems . Three intermediate heat exchangers, of which one is shown in Figure 21.11, heat up a secondary $He-N_2$ loop, which drives a gas turbine. A heat recovery boiler behind the gas turbine drives a bottom steam cycle to reach a net efficiency of 45% at a core outlet temperature of 850 °C.

The greatest challenge facing the GFR is the development of robust high-temperature ceramic fuels and core structural materials, capable of withstanding the thermal, mechanical, and radiation environment in the reactor core. Spent fuel treatment for the GFR can be accomplished with aqueous processes similar to those of the SFR but qualified for the unique GFR fuel form. At least two fuel forms have the potential to meet the GFR requirements: a ceramic plate-type fuel element and a ceramic pin-type fuel element. In general, however, the fuel needed for such reactors is new. Experience from HTR-coated particles can hardly be used, and the new fuel needs to be tested in research reactors before it could be applied to this novel reactor design. This extends the development time significantly.

21.2.6 Molten Salt Reactor (MSR)

In a molten salt reactor (MSR), the fuel is dissolved in a fluoride salt coolant. The technology was partly developed in the 1950s and 1960s in the United States (ORNL) leading to operating test reactors. Compared with solid-fueled reactors, MSR systems have lower fissile inventories, the absence of radiation damage that can limit fuel burn-up, the possibility of continuous fission-product removal, the avoidance of the expense of fabricating fuel elements, and the possibility of adding fuel continuously as needed. Noteworthy is also the homogeneous isotopic composition of fuel in the reactor. These and other characteristics may enable MSRs to have potentially unique capabilities and competitive economics for actinide burning and for extending fuel resources. One important aspect in the development of the MSR was the idea to use the thorium (Th) fuel cycle. The MSR capable of continuous removal of fission products opened up the option of breeding U-233 with a thermal neutron spectrum. The Th fuel cycle has its advantages that it produces much less minor actinides than the U-cycle during irradiation. The Th cycle has however to be started with enriched uranium, as natural U-233 fission material does not exist.

Advantages of the MSR in combination with the thorium fuel cycle are the following.

- Fluoride inorganic salts used as a carrier for the fuel and as a coolant are very stable chemical compounds under reactor operating conditions.
- High solubility for actinides, very low vapor pressure, and good heat transfer properties.
- No reaction with air or water, and inert to some commonly used structural materials.
- Soluble fission products can be removed online in a chemical reprocessing plant, while nonsoluble fission products and the noble metals can be extracted from the salt by helium bubbling.
- Freeze plug and drain tank as safety measures: If the salt heats up to levels above design values or if the power

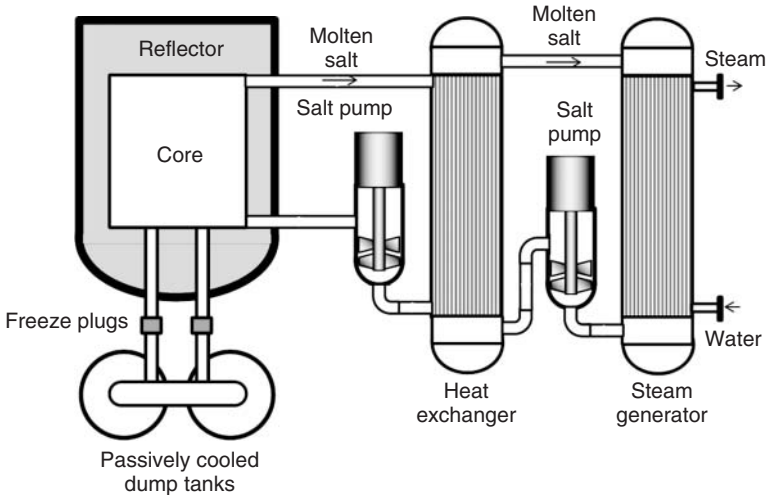


Figure 21.12 Concept of the molten salt reactor (MSR).

supply fails, the plugs will melt and the salt will be drained into a drain tank.
- A fast increase in the fuel temperature leads to salt expansion, providing instantaneous negative reactivity feedback.
- Optical transparency of salt.
- The primary and secondary circuits are operated under ambient pressure.
- The thorium fuel cycle produces much less long-lived nuclear waste.

New aspects include the use of the Brayton power cycle (rather than a steam cycle) that eliminates many of the historical challenges in building MSRs as well as the conceptual development of fast-spectrum cores for the MSR that have large negative temperature and void reactivity coefficients – a unique safety characteristic not found in any solid-fuel fast reactor. In addition, the development of higher temperature salts as coolants would open the MSR to new nuclear and nonnuclear applications. These salts are being considered for intermediate heat transport loops within all types of HTR systems (helium and salt cooled) and for hydrogen production, oil refineries, and shale oil processing facilities, among other applications. For most of these applications, the heat would have to be transported up to a kilometer or more.

Besides the thermal neutron spectrum, an innovative thorium-fueled molten salt fast reactor (MSFR) with an epithermal neutron spectrum has been proposed by Nuttin *et al.* [16] and Renault *et al.* [17] with a thermal power of 3000 MW_{th}, as sketched in Figure 21.12. The mean feature of the MSFR concept is the removal of the graphite moderator from the core. The core is a single cylinder (2.3 m diameter, 2.3 m height) where nuclear reactions take place within the flowing fuel salt. The fuel salt is a binary salt, composed of LiF enriched with ^{7}Li with a melting temperature of 550 °C. The operating temperature range is from 700 to 800 °C. In total, 18 m^3 of salt is used, of which 9 m^3 is in the core. The initial inventory is 5.06 t U-233 and 38.3 t Th-232. Each year, 1.112 t of thorium is consumed and 93 kg of U-233 is produced, resulting in a breeding ratio of 1.085.

The level of safety reached by the concept is excellent because the feedback coefficients of the MSFR are negative, even if larger amounts of minor actinide fuel is used for waste incineration.

As a conclusion, the MSR could be the choice and option in case there is increased interest in the Th fuel cycle. Thorium is at least three times more abundant than uranium. As today's reactor technology is uranium oriented, this is also a disadvantage for the promotion of the MSR because investments in the Th fuel cycle are needed. In this regard, the MSR seems to be a concept farther in the future compared, for example, to the SFR. However, MSR is a concept where already practical experience has been gained and reactors have been licensed and operated. So, it is definitely more than pure paperwork. Nevertheless, in the reactor community, the MSR is a very innovative and challenging approach to breed fissile material from thorium as well as to burn minor actinides from light water reactor spent fuel. Basic technologies needed for this reactor are still to be developed and optimized. The online reprocessing is not fully settled, and the structural materials are still an issue. Reactor design studies for larger units have been rather generic until now. The design of a realistic reactor or even a nuclear power plant will certainly need some more decades of preparatory work.

21.3 Outlook

The first 10 years of joint research in the GIF, from 2002 to 2012, have produced a large number of innovative design concepts and technologies that would be useful for future reactor systems. The decision on building a prototype for any of these reactors, however, is still pending. This is due to the renewed focus of worldwide R&D activities on nuclear safety and emergency management, as a consequence of the nuclear accident in Fukushima in 2011, requiring a significant part of the limited human and financial resources. Besides this issue, there is no urgent market need yet for fast or highly efficient reactors as long as plenty of cheap natural uranium is available in the world.

Therefore, the GIF road map for the second decade is primarily concentrating on technology development such as material research programs and component tests, which have the longest lead times. Some examples are listed in the following.

- The SCWR research program will concentrate on further material development and tests of fuel claddings and will perform an in-pile test of a small-scale fuel assembly at supercritical water conditions in a research reactor.
- The VHTR research program will focus on fuel and graphite qualification, materials for heat exchangers and pressure vessels, and coupling to process heat applications including hydrogen production.
- The SFR will continue with innovative fuel development, both metal and oxide, and continue with conceptual designs of both large base load and small modular SFRs.
- The LFR has development needs in the areas of fuels, materials performance, and corrosion control.

Besides further technology development, the GIF projects will continue to provide a field for high-level scientific education and training in nuclear technologies, an advantage that should not be

underestimated for future recruitment of nuclear experts.

Glossary

Blanket: Fuel assemblies surrounding the core, containing U-238 or Th-232 for breeding.

The Brayton Cycle: Thermodynamic cycle of a gas turbine.

Breeding: Conversion of U-238 or Th-232 by neutron absorption to fissile material.

Burning: Fission of spent fuel to convert minor actinides to fission products.

Burn-up: The thermal energy gained from nuclear fission in a given mass of fuel.

Cladding: Metal or ceramic tube housing the fuel.

Closed Fuel Cycle: Recycling of spent fuel.

Conversion Ratio: The ratio of produced fissile material divided by the spent fissile material. A conversion ratio >1 means breeding.

Core: The central part of a nuclear reactor where nuclear fission occurs.

The Doppler Effect: The reduction of reactivity in a reactor core when the fuel heats up, caused by an increase in neutron absorption later.

Excess Reactivity: Additional reactivity in a fresh reactor core to compensate for later burn-up.

Fast Reactor: Nuclear reactor in which fission is dominantly caused by fast neutrons from fission.

Fuel: Uranium, plutonium, or other material for nuclear fission.

Fuel Assembly: Reactor core component containing fuel.

Minor Actinides: The heaviest group of chemical elements.

Moderator: Material such as water or carbon that slows down the neutron energy.

Open Fuel Cycle: Direct disposal of spent fuel without recycling.

Reactivity: Degree of change of the neutron multiplication in a reactor core. The reactivity is zero if the reactor runs at steady state.

Reprocessing: Recycling of spent fuel by separating U and Pu from fission products.

Thermal Reactor: Nuclear reactor in which fission is dominantly caused by slow neutrons in thermal equilibrium with the coolant.

Void Effect: The reduction or increase of reactivity in a reactor core if gas or steam is appearing in the coolant.

References

1. GIF (2002) A Technology Roadmap for Generation IV Nuclear Energy Systems, issues by the U.S. DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum, Washington, DC.
2. OECD Nuclear Energy Agency and IAEA (2012) *Uranium 2011: Resources, Production and Demand (the Redbook)*, NEA No. 6891, OECD Publishing.
3. Schulenberg, T. and Starflinger, J. (eds) (2012) *High Performance Light Water Reactor, Design and Analyses*, KIT Scientific Publishing, Karlsruhe, ISBN: 978-3-86644-817-9.
4. Oka, Y., Koshizuka, S., Ishiwatari, Y., and Yamaji, A. (2010) *Super Light Water Reactors and Super Fast Reactors*, Springer, ISBN: 978-1-4419-6034-4.
5. Yetisir, M., Diamond, W., Leung, L.K.H., Martin, D., and Duffey, R. (2011) Conceptual mechanical design for a pressure-tube type

supercritical water-cooled reactor. Proceedings of the ISSCWR-5, Vancouver, Canada, March 13–16, 2011.

6. Zhang, Z., Wu, Z., Wang, D., Xu, Y., Sun, Y., Li, F., and Dong, Y. (2009) Current status and technical description of Chinese 2x250 MWth HTR-PM demonstration plant. *Nucl. Eng. Des.*, **239**, 1212–1219.
7. Baxi, C.B., Shenoy, A., Kostin, V.I., Kodochigov, N.G., and Vasyaev, A.V. (2008) Evaluation of alternate power conversion unit designs for the GT-MHR. *Nucl. Eng. Des.*, **238**, 2995–3001.
8. IAEA (2006) *Fast Reactor Data Base 2006 Update*, IAEA TECDOC 1531, ISBN: 92-0-114206-4.
9. Hahn, D., Kim, Y.I., Kim, S.O., Lee, J.H., Lee, Y.B., and Jeong, H.Y. (2007) Conceptual design of the sodium-cooled fast reactor KALIMER-600. *Nucl. Eng. Technol.*, **39** (6), 205–218.
10. Ichimiya, M. (2011) The status of generation IV sodium-cooled fast reactor technology development and its future project. *Energy Procedia*, **7**, 79–87.
11. Tsuboil, Y., Kasuga, S., Sakashita, Y., Matsumiya, H., Hasegawa, K., Ueda, N., Grenci, T., and Handa, N. (2009) Development of the 4S and related technologies, plant system overview and current status. Proceedings of ICAPP '09, Tokyo, Japan, May 10–14, 2009, Paper 9214.
12. Bortot, S., Moisseytsev, A., Sienicki, J.J., and Artioli, C. (2011) Core design investigation for a SUPERSTAR small modular lead-cooled fast reactor demonstrator. *Nucl. Eng. Des.*, **241**, 3021–3031.
13. Alemberti, A., Carlsson, J., Malambu, E., Orden, A., Struwe, D., Agostini, P., and Monti, S. (2011) European lead fast reactor ELSY. *Nucl. Eng. Des.*, **241**, 3470–3480.
14. Zrodnikov, A.V., Toshinsky, G.I., Komlev, O.G., Dragunov, Y.G., Stepanov, V.S., Klimov, N.N., Kopytov, I.I., and Krushelnitsky, V.N. (2008) Innovative nuclear technology based on modular multi-purpose lead-bismuth cooled fast reactors. *Prog. Nucl. Energy*, **50**, 170–178.
15. Stainsby, R., Peers, K., Mitchell, C., Poette, C., Mikityuk, K., and Somers, J. (2011) Gas cooled fast reactor research in Europe. *Nucl. Eng. Des.*, **241**, 3481–3489.
16. Nuttin, A., Heuer, D., Billebaud, A., Brissot, R., LeBrun, C., Liatard, E., Loiseeaux, J.M., Mathieu, L., Meplan, O., Merle-Lucotte, E., Nifenecker, H., and Perdu, F. (2005) Potential of thorium molten salt reactors: detailed calculations and concept evaluation with a view to large scale energy production. *Prog. Nucl. Energy*, **46**, 77–99.
17. Renault, C., Delpech, S., Merle-Lucotte, E., Konings, R., Hron, M., and Ignatiev, V. (2009) The Molten Salt Reactor (MRSR), R&D status and perspectives in Europe. Proceedings of FISA 2009, Prague, Czech Republic, June 22–25, 2009.

22
Transmutation of High-Level Nuclear Waste by Means of Accelerator Driven System (ADS)*

Hamid Aït Abderrahim, Didier De Bruyn, Gert Van den Eynde, and Sidney Michiels

* This article was previously published in WIREs Energy and Environment, Wiley 2013. doi:10.1002/wene.82

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

22.1
Introduction

22.1.1
Energy in the World

One of the major challenges that our society faces is the increasing demand for energy, in general, and electricity, in particular. According to the World Energy Council and the International Energy Agency of the Organisation for Economic Co-Operation and Development (OECD), global energy requirements will increase by 50% by 2035 [1]. Several factors play a role in this:

- the expected growth of the world population to 8.1 billion people by 2035;
- the availability of energy (electricity) to the part of the global population that currently has no access to electricity;
- higher consumption in countries that are still in full development today;
- an expected increase in the average life expectancy, which is associated with higher average energy consumption.

Of all energy sources used for electricity production that are being tapped worldwide today, fossil fuels account for nearly 70%. Oil represents 5.1%; coal, 40.5%; and natural gas, 21.4%. Fossil fuels represent the most widely used source of energy worldwide, despite having a less favorable performance with regard to CO_2 emissions and greenhouse effects. At a global level, hydropower accounts for 16.2% of the overall energy production, and nuclear energy supplies 13.4% of the worldwide energy requirement. Other energy sources, including renewable energy and biomass, currently supply 3.3% of the global energy needs. This is graphically represented in Figure 22.1.

While the increase of the global energy requirement is a continuing trend, the available sources of energy are also being increasingly tapped. Fossil fuels will continue to maintain their prominent role: they are still in sufficient supply and are easily accessible. They are often the first source that countries rely on during periods of rising energy requirement. Nowadays, we are confronted with excessive CO_2 emissions. In order to achieve the global targets with regard to these emissions, it will become necessary to limit the use of fossil fuels. At the same time, renewable energy sources cannot satisfy the demand. For this reason, different countries recognize that nuclear energy needs to be part of the ''energy basket'' of the future.

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

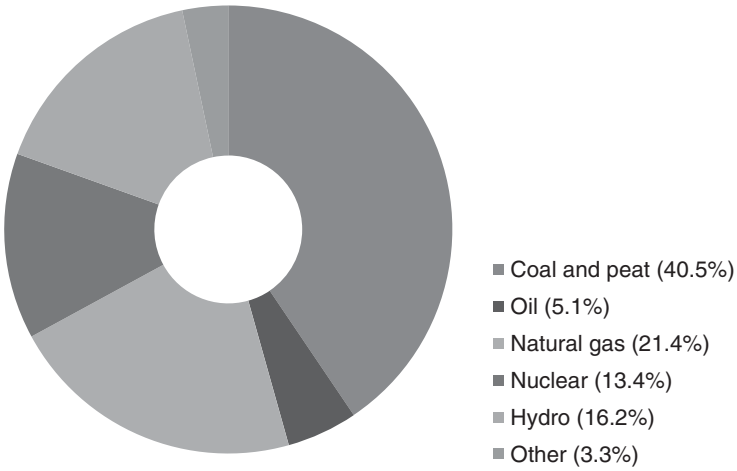


Figure 22.1 The globally used sources of electrical energy (2011) [2].

22.1.2 Nuclear Energy

22.1.2.1 Basic Principles: Fission vs Spallation

In nuclear fission, a uranium atom is split. The uranium atom explodes as a result of the deliberately induced impact of a neutron with the uranium nucleus. The nucleus splits into two new, smaller nuclei and, on average, two or three neutrons, and about 200 MeV of energy is released. The released neutrons can in turn split another uranium atom. This creates a chain reaction.

In a nuclear reactor, the released neutrons from a fission reaction can undergo various reactions: they can induce a new fission leading to the birth of new neutrons; they can be absorbed in the structures of the reactor or even in the fuel without inducing fission, leading to the production of heavier elements than uranium (Np, Pu, Am, Cm, etc.); or they can leak out of the reactor core and then are considered as lost for the chain reaction. Thus, to maintain the chain reaction, the critical mass and the critical dimensions of the reactor need to be considered and these correspond to quantities of fissile materials arranged in a given geometry in such a way that the amount of neutrons produced by fission events is balanced by the amount of neutrons lost either by sterile absorptions or by leaking out of the reactor core. When these conditions are met, one speaks about ''critical'' reactor.

An accelerator driven system (ADS) will not have this chain reaction. Indeed, its core is subcritical. In this context, the term *subcritical* means that, on average, for each generation of neutrons, less than one secondary neutron is capable of initiating a subsequent nuclear fission. This situation can be achieved by limiting the amount of fissile material in the core; in this way, the chain reaction is not self-sustaining. Despite this subcriticality, one can operate such a reactor in continuous mode, but an external neutron source is needed. In an ADS, these supplementary neutrons are produced by spallation reactions, during which high-energy protons, coming from a particle accelerator, impinge on a spallation target made of a heavy metal such as lead. In the course of the spallation process, the

Figure 22.2 (a) Spallation versus (b) fission.

liquid metal nuclei emit a large number of neutrons whose energy spectrum is made of two parts: a conventional fission spectrum and a high-energy tail up to the energy of the incident proton. At 600 MeV, each impinging proton on a lead target would produce about 10 primary neutrons (Figure 22.2).

22.1.2.2 **Radioactive Waste**

The generation of nuclear energy from uranium produces, besides energy, high-level nuclear waste. The high-level radioactive waste (HLW) quantity depends on the considered fuel cycle (FC). To illustrate this let us consider the various fractions produced by using 1 t of uranium as fuel in a pressurized water reactor (PWR)-type nuclear reactor after 4.5 years (corresponding to an average burnup of 50.000 MWd t_{HM}^{-1}), as shown in Figure 22.3.

Depending on the operational FC chosen, the resulting amount of HLW will be

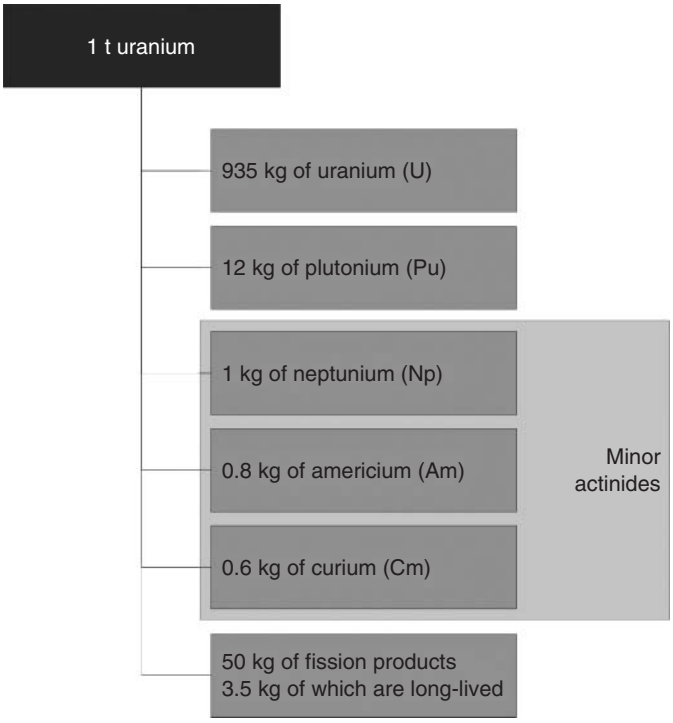


Figure 22.3 Fractions of waste produced by 1 t of uranium in a PWR after a burnup of 50.000 MWd t_{HM}^{-1}.

1000 kg for the "once-through" FC, ~56 kg representing the minor actinides (MAs) and fission products (FPs) if reprocessing is applied and the retrieved uranium and plutonium are recycled, and even less if "advanced reprocessing" to separate the MAs and the LLFPs (long-lived fission products) is applied. These HLWs can then be treated by transmutation in dedicated waste burners such as ADS.

22.1.2.3 Better Use of Raw Materials and Resources

Nowadays, nuclear energy is generated by splitting uranium nuclei in thermal reactors. The use of thermal neutrons has a negative effect on the efficient use of uranium. Indeed, thermal neutrons can fission only ^{235}U, which is present at 0.7% in natural uranium. The major part of the ore, ^{238}U, remains unusable for fission in reactors of the current generation with a thermal neutron spectrum.

If nothing changes from a technological point of view and in the case of a rapid development of nuclear energy, known uranium resources will become scarce before the end of the century.

Reactors with a fast neutron spectrum allow using the remaining 99.3% of mineral uranium as fuel by transforming ^{238}U to ^{239}Pu. As a result, the present uranium resources can be used more efficiently, leading to uranium resources for more than several thousands of years.

22.1.2.4 Subcriticality, a Must or a Luxury for Transmutation of Minor Actinides (MAs)

When considering transmutation of MAs, the aiming is to destroy these nuclides through the fission process for turning them into FPs. Figure 22.4 illustrates the

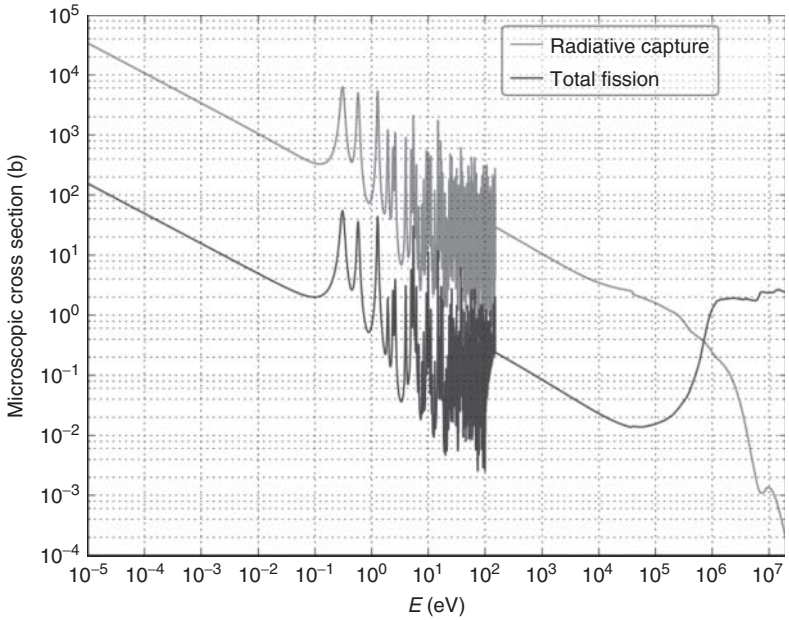


Figure 22.4 Neutron spectrum of ^{241}Am. Figure produced using JANIS 3 (Java-based Nuclear Information Software, *http://www.oecd-nea.org/janis*. (Source: Am241 from JEFF 3.1.2 library.)

capture and fission cross sections of ^{241}Am as a function of energy. We can see clearly that below 1 MeV, the capture cross section is two to three decades higher than the fission cross section. The latter becomes predominant beyond a high-energy threshold, the level just below 1 MeV. Thus, it is obvious that MA transmutation needs to be conducted in fast neutron spectrum reactors. This is the first condition for a successful transmutation of MAs.

In a nuclear reactor, the fission process produces neutrons. Each fission event delivers ν neutrons, out of which a given fraction designated β is delivered with a delay (ranging from 0.1 to a few tens of seconds) and is called the *delayed neutron fraction*. The delayed neutron fraction plays an essential role in the kinetic behavior of the neutron population ($\Delta n/\Delta t$) in the reactor core and hence on the reactor power variation:

$$\frac{\Delta n}{\Delta t} = \frac{[k_{\text{eff}}(1-\beta)-1]}{l} n + \lambda C + S$$

where

k_{eff}	=	the criticality level of the reactor ($k_{\text{eff}}=1$ for a critical reactor, <1 for a subcritical one);
β	=	the delayed neutron fraction;
l	=	the average life time of a neutron in the reactor, which is $\sim 10^{-7}$ s for a fast spectrum system;
n	=	the neutron population in the reactor;
λ	=	the decay constant of the delayed neutrons (in general, one works with six delayed neutron groups);
C	=	the concentration of the precursors emitting the delayed neutrons;
S	=	the fixed neutron source (such as the spallation source within the ADS).

The delayed neutron fraction appearing in the above equation is the reactor core volume weighted average value $\langle\beta\rangle$:

$$\langle\beta\rangle = \frac{\Sigma_i \int_{\nu_{\text{core}}} \nu_i \beta_i N_i \sigma_{\text{fi}}(E)\phi(E)\mathrm{d}E\mathrm{d}V}{\Sigma_i \int_{\nu_{\text{core}}} \nu_i \beta_i N_i \sigma_{\text{fi}}(E)\phi(E)\mathrm{d}E\mathrm{d}V}$$

where

ν_i	=	the number of neutrons/fission of the ith fissile nuclide present in the core;
β_i	=	the delayed neutron fraction of the ith fissile nuclide present in the core;
N_i	=	the concentration of the ith fissile nuclide present in the core;
$\sigma_{\text{fi}}(E)$	=	the microscopic fission cross section of the ith fissile nuclide present in the core;
$\varphi(E)$	=	the scalar neutron flux present in the core.

The values of β_i are nuclide dependent as illustrated in the table below:

Nuclide	β_i (pcm)
^{235}U	650
^{238}U	1480
^{237}Np	334
^{238}Pu	120
^{239}Pu	210
^{240}Pu	270
^{241}Pu	490
^{242}Pu	573
^{241}Am	113
^{243}Am	208
^{242}Cm	33
^{244}Cm	100

To keep a fast spectrum critical reactor controllable, there should be an average b delayed neutron fraction, around 400–500 pcm. In a critical reactor, the fraction of MAs will thus be limited to a small percentage (2–5%). If one wants to have a limited number of MAs burning, with the reactor core containing up to 40–50% of MAs, the only solution is to have a subcritical reactor. Thus, for a concentrated and an efficient burning of MAs, the subcriticality is not a luxury but a must.

22.2 Partitioning and Transmutation

The concept of partitioning and transmutation (P&T) has three main goals:

- reduction of the radiological hazard associated with spent fuel by reducing the inventory of MAs;
- reduction of the time interval required to reach the radiotoxicity of natural uranium;
- reduction of the heat load of the HLW packages to be stored in geological disposal, leading to their efficient use.

Partitioning is the separation of the radiotoxic elements from the spent fuel. Regarding P&T, various FCs have been considered for the optimal use of the natural resources and the HLW production optimization, but the consideration should include the economic impact on the ultimate cost of the produced electric kilowatt hour and the technical and technological readiness of the various steps to be taken into account from the separation point of view, as well as the advance fuel fabrication and the burning capabilities of the dedicated transmuters.

Presently, the possible FCs can be limited to the four following options (Figure 22.5):

- The *"once-through" nuclear fuel cycle* (a), where there is no reprocessing of the spent fuel.
- Within a simple recycle program, a part of the plutonium is reused, together

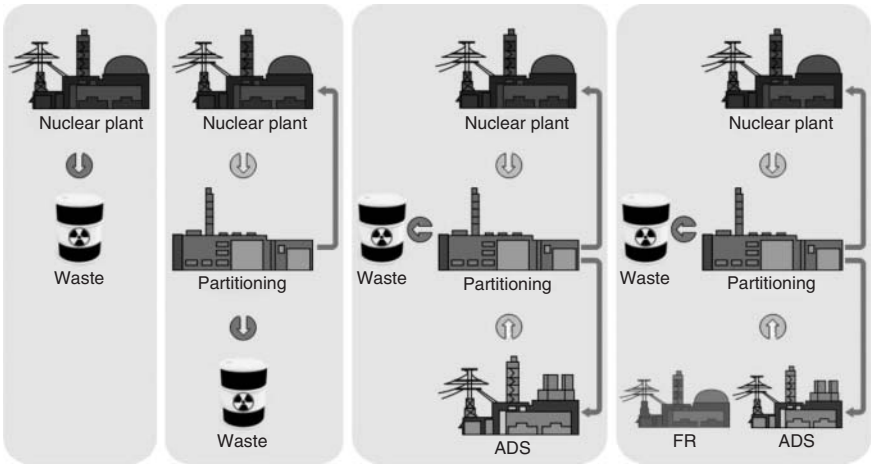


Figure 22.5 Schematic overview of partitioning.

with fresh uranium, to produce mixed oxide (MOX) fuel (b).
- In the next step, the unprocessed actinides can be "burned" in an ADS (c).
- In the *double strata concept* (d), fuels with a high content of MAs can be "burned" in ADS or in dedicated fast reactors (FRs).

Transmutation of high-level radioactive elements with a long half-life present in the nuclear waste reduces the radiological impact of the actinides (such as americium, curium, and neptunium) and FPs. The timescale needed for the radiotoxicity of the waste to drop to the level of natural uranium will be reduced from a "geological" value (500 000 to 1 million years) to a value that is comparable to that of human activities (several hundreds of years) (Figure 22.6) [2–4]. During transmutation, the nuclei of the actinides are split into shorter-lived FPs.

In order to efficiently transmute the elements with long half-life, a radiation source with high intensity and energy efficiency is necessary. Therefore, only nuclear fission reactors and spallation targets can be utilized. These sources deliver neutrons in the energy range from thermal to about 20 MeV, which is necessary to induce transmutation reactions.

If the aim is to transmute large amounts of MAs in one installation, it is necessary to have some dedicated fast systems, such as ADS. The key reaction for the transmutation of actinides is the fission reaction, which transforms long-lived high-radiotoxic actinides into mostly short-lived less-toxic FPs.

22.3 ADS Technology

The concept of ADS combines a particle accelerator, a spallation target, and a nuclear reactor with a subcritical core (see Figure 22.7). In most proposals, a proton accelerator can be found. The protons are injected by the accelerator onto a spallation target, where they produce neutrons for driving the subcritical core. Except for the subcritical state, the core of an ADS is similar to that of a critical reactor.

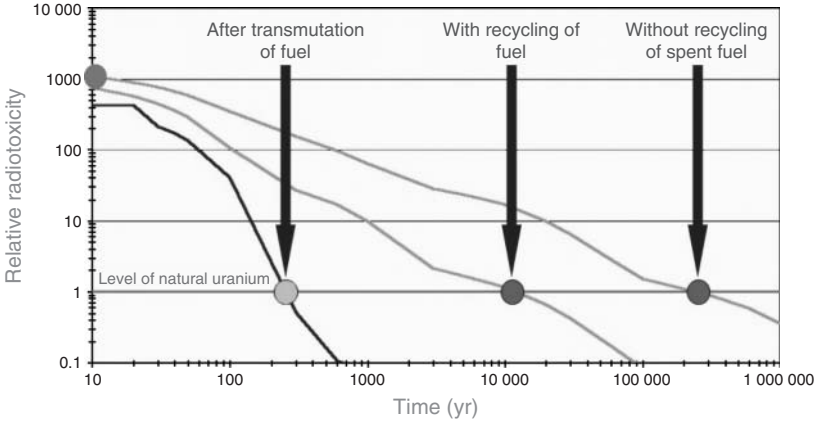


Figure 22.6 Radiotoxicity of radioactive waste. (Source: Reference: Figure 171 in Ref. 5.)

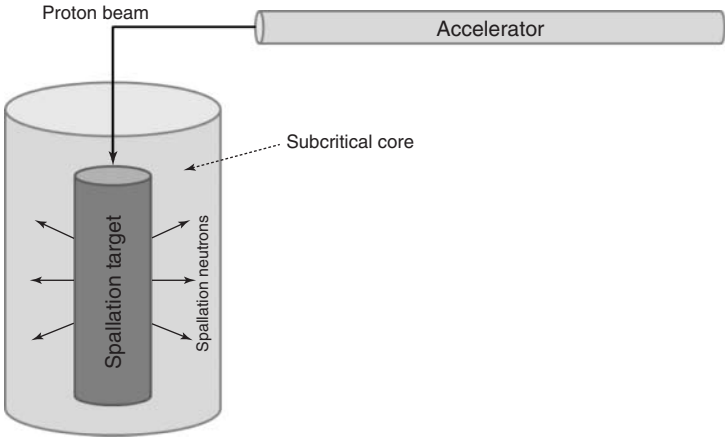


Figure 22.7 Concept of an accelerator driven system.

The different components of an ADS are discussed in detail in the following sections.

22.3.1 Accelerator

The accelerator is the driver of the ADS. The accelerator provides the high-energy protons that are used in the spallation target to create neutrons that feed in their turn the subcritical core. The accelerator can both be a linear accelerator (linac) or a circular accelerator (cyclotron). The accelerators used in an ADS must be high-power accelerators and must have an extremely high reliability and availability. These accelerators are under continuous development, and the construction of these machines seems feasible.

Particle accelerators delivering a beam on a target can be organized according to the time structure of their beam:

- **Direct current (DC) beam delivery**: This beam with no time structure is typically obtained from a DC accelerator.
- **Continuous wave (CW) beam delivery**: In order to be accelerated by radiofrequency (RF) fields, the beam has to be subdivided into small packets called *bunches* and the RF has to be synchronized with the bunch progression through the accelerator. If the resulting bunch train is delivered continuously to the target, the beam delivery is called *continuous wave*. For providing such a beam, the accelerator must operate in a steady state and must be characterized by constant guiding fields and constant RF.
- **Pulsed beam delivery**: If the bunch train is interrupted periodically (but with a frequency that is much smaller than the RF) a beam on/off cycle is obtained, leading to pulsed operation. Within the pulse, the bunch structure of the beam stays.

The choice between a linac or a cyclotron depends on many factors. In both cases, it is important to specify the beam specifications for ADS applications, and more specifically, the little amount of beam interrupts. For most ADSs, the accelerator delivers the particles in continuous mode, although pulsed operation could also be used (e.g., for testing purposes).

Linear accelerators have the advantage that they operate with RF-cavity frequencies that are typically 10–20 times higher than those of cyclotrons. The possibility to place the focusing elements (quadrupoles) at frequent intervals as well as the longitudinal focusing set a higher limit to the charge per bunch that can be accelerated without significant beam losses. Together, these factors result in the fact that a linac could accelerate a current one to two orders of magnitude higher than a cyclotron, without problems at extraction. The major drawback of linacs is the length, which depends on the final beam energy and the accelerating gradient. The length of the linac is an important factor in the cost of the facility, while most of the accelerator needs to be shielded against radiation.

The base of a cyclotron is the so-called cyclotron resonance: in a constant magnetic field, perpendicular to the beam orbit, the particle revolution frequency is fixed and independent of the particle energy. The major problem with cyclotrons is the beam extraction system. The deflecting system, which guides the beam out of the accelerator, is permitted to touch only a negligible fraction of the beam. The current limit of the cyclotron is given by a design requirement to produce a clean beam at the outer radius of the machine.

The accelerator must have a high reliability and efficiency. The number of beam trips should be limited because they can cause temperature and mechanical stress transients comparable to fast control rod insertions (SCRAM) in critical reactors.

22.3.2 Spallation Target

The spallation target is a neutron source providing primary neutrons that are multiplied by the surrounding subcritical core. The primary neutrons are produced by the spallation reaction of heavy-metal target nuclei, bombarded by the high-energy particles generated by the accelerator.

The spallation target in one of the most stressed components of an ADS. The spallation target is designed to generate the maximum amount of neutrons while it must be able to evacuate the heat generated by the spallation process. In addition, the mixed proton–neutron irradiation field, which typically can be found in the spallation target of an ADS, imposes specific conditions on the design and operability of the target and influences all the thermomechanical options of these targets. The target of an ADS must be cooled, mostly by gas, heavy water, or liquid metal.

There are two options for the spallation target: a solid target or a liquid metal target. Within the operational low-power facilities, there is a lot of experience with solid targets, especially when combined with an accelerator that delivers a pulsed beam. These solid targets are difficult to cool and experience radiation damage. The target should be cooled further, even after the accelerator is shut off. The path for solid targets for use in an ADS is abandoned in favor of the liquid metal targets.

For high-power facilities, such as an ADS, the use of a liquid metal spallation target is preferred. Lead and lead-bismuth eutectic (LBE) are the two primary candidates for liquid metal target for the production of spallation targets in an ADS. The cooling of a liquid metal spallation target is inherent to the design of the system; there is a constant flow of liquid metal coolant. At the same time, it is possible to achieve high power densities without having significant irradiation damage to the target. The drawback of a liquid metal target is

that the coolant, especially lead or lead bismuth, can cause erosion or corrosion of the structural materials that are in contact with it.

The use of a liquid metal spallation target introduces some design difficulties. First of all, we have the choice of materials for the target. The materials must withstand the thermomechanical loads on the targets and must be compatible with the liquid metal. Second, an interface between the target and the beam guide must be foreseen. This may take the form of a solid beam window or a windowless design. A beam window is exposed to a significant flux of high-energy protons and neutrons at high temperatures and to a highly corrosive environment. Therefore, it is very likely that the beam window must be replaced periodically because of the expected material damage. The use of a windowless target where the window is integrated with the beam guide tube will result in approximately 1.5–2 times lower radiation dose to this window because neutron irradiation will be negligible in comparison with proton irradiation.

22.3.3 Subcritical Core

The subcritical core of an ADS is constructed in such a way that it does not have enough fissile material to reach criticality. Source neutrons are needed as an input to drive the system. They are obtained from an energetic beam delivered by the accelerator impinging on a target, through spallation reactions. The subcritical core of an ADS acts as an amplifier; on average, ~15 neutrons are produced for each proton.

The energy spectrum of the subcritical core depends on the application of the core. For the transmutation of MAs, a fast neutron spectrum has various advantages, including a larger excess of neutrons, reduced amount of MA production in the core, and better energy yield for eventual energy production.

The evolution of power with time and temperature feedback effects will have an effect on the multiplication factor of the core, as in critical systems. However, for a subcritical core, the behavior is dominated by the external source and its variation in time. The closer to a critical system, the more important the feedback elements of the core will become. There are no final criteria to define an optimal level of subcriticality.

In case of critical cores, reactivity of the core and the power level of the reactor monitored through control rods. With an ADS, it should be possible to control the power level of the subcritical reactor by an external source. When a core is working in the ''source dominated'' mode, the shutdown of the source has an instantaneous effect of power reduction. The inverse effect, a sudden increase of the source, has the consequence of an instantaneous increase in the reactor power.

22.3.4 Advantages of an ADS

The main advantages of an ADS compared to critical reactors are twofold: an ADS allows greater flexibility with respect to the fuel composition and it has a potentially enhanced safety.

ADSs are able to burn fuels that are problematic from the standpoint of critical reactor operation, namely, fuels that would degrade neutronic safety characteristics of the critical core to unacceptable levels because of the small delayed neutron fractions and short neutron lifetimes. Among other systems, an ADS is capable of coping with ^{233}U and MA fuel. Additionally, an

ADS allows the use of nonfissile fuels, such as thorium, without requiring the incorporation of U or Pu into the fresh fuel.

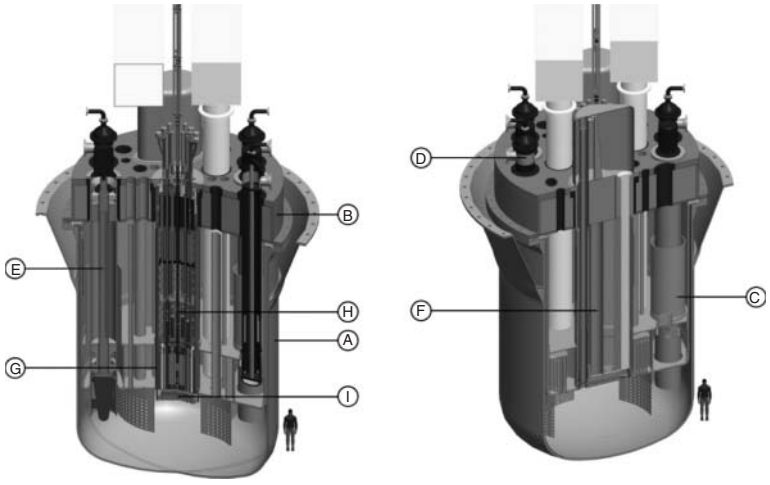
The enhanced safety of an ADS is due to the fact that once the accelerator is turned off, the system shuts down. When the margin toward criticality is sufficiently large, reactivity-induced transients will never result in a supercritical accident.

As the accelerator is the driver of the system, the power of the ADS is controlled by the beam current of the accelerator. This feature can be utilized to compensate for fuel burnup.

In the context of transmutation, another advantage of the ADS is its noncritical core. This results in increased core design and fuel management flexibility. The subcriticality also opens opportunities for new reactor concepts, including concepts that are otherwise ruled out by insufficient neutron economy. An ADS allows transmuters to be designed as purely transuranic (TRU) of MA burners. This minimizes the fraction of specialized transmuters in the reactor park.

22.3.5 Example: the Development of MYRRHA in Belgium

Multi-purpose hYbrid Research Reactor for High-tech Applications (MYRRHA) is a pool-type ADS cooled by LBE. While MYRRHA is a pool-type design, the reactor vessel houses all the primary systems and the primary coolant. The vessel is closed by the reactor cover. The diaphragm, located within the reactor vessel, divides the LBE coolant in a hot, low-pressure upper pool and a cooler, high-pressure lower pool. It serves as the supporting structure for the major internal components and foresees feedthroughs to the cold plenum. Figure 22.8 shows a section of the MYRRHA-FASTEF reactor



A. Reactor vessel
B. Reactor cover
C. Diaphragm
D. Primary heat exchanger
E. Primary pump
F. In-vessel fuel-handling machine
G. Core
H. Above core structure
I. Core restraint system

Figure 22.8 Overview of the MYRRHA reactor.

with its main internal components. The MYRRHA reactor can be switched to operate in subcritical ADS mode or as a critical reactor.

22.3.5.1 The MYRRHA Accelerator

The accelerator is the driver of MYRRHA while it provides the high-energy protons that are used in the spallation target to create neutrons, which in turn feed the subcritical core. The accelerator must be able to provide a proton beam with an energy of 600 MeV and a maximum current of 4 mA, which will be delivered to the core in the CW mode. The beam is delivered to the core from above through a beam window.

Accelerator availability is a crucial issue for the operation of the ADS. A high availability is expressed by a long mean time between failure (MTBF), which is commonly obtained by a combination of overdesign and redundancy. Besides these two strategies, fault tolerance must be implemented to obtain the required MTBF. Fault tolerance will allow the accelerator to recover the beam within a beam trip duration tolerance after the failure of a single component. In case of MYRRHA, the beam trip duration tolerance is 3 s. Within an operational period of MYRRHA, the number of allowed beam trips exceeding 3 s must remain under 10; however, shorter beam trips are allowed without limitations. The combination of redundancy and fault tolerance should allow obtaining an MTBF value in excess of 250 h.

At present, proton accelerators with megawatt-level beam power in CW mode only exist in two basic concepts: sector-focused cyclotrons and linear accelerators (linacs). Cyclotrons are an attractive option with respect to construction costs, but they do not have any modularity, and hence, a fault tolerance scheme cannot be implemented. Also, an upgrade of its beam energy is not a realistic option. A linear accelerator, especially if made superconducting, has the potential for implementing a fault tolerance scheme and offers high modularity, resulting in the possibility to recover the beam within a short time and increasing the beam energy.

22.3.5.2 MYRRHA Spallation Target and Subcritical Core

The reactor core consists of MOX fuel pins, typical for FRs. A total of 151 hexagonal positions can be taken by fuel assemblies, control and scram rods (in the critical configuration), the spallation target (in subcritical configuration), and in-pile test sections (IPSs). When MYRRHA is operating in ADS mode, the central position is always taken by the spallation target. Totally, 37 positions are available for the IPS: these positions are accessible from the top and have the possibility to house on top extra out-of-pile equipment to create other irradiation conditions. Figure 22.9 shows a core configuration for MYRRHA.

The requested high fast flux intensity has been obtained by optimizing the core configuration geometry (fuel rod diameter and pitch) and maximizing the power density. We will be using, for the first core loadings, 15-15Ti as the cladding material instead of T91, which will be qualified progressively during MYRRHA operation. Thanks to the use of LBE as coolant, it permits to reduce the core inlet operating temperature to 270 °C. Consequently, the risk of corrosion is decreased and the core ΔT might be increased. This, together with the adoption of reliable and passive shut down systems, will permit to meet the high fast flux intensity target.

The spallation target assembly, located in the central of the core, brings the proton

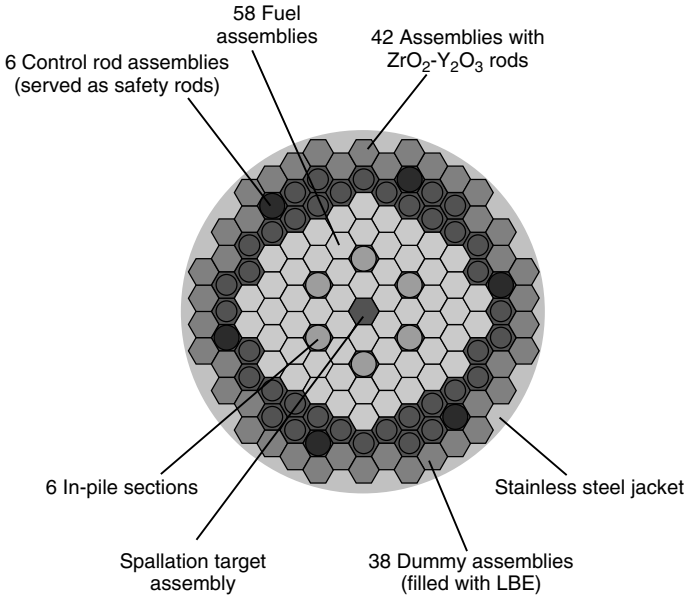


Figure 22.9 Cross-sectional view of the MYRRHA core.

beam via the beam tube into the central core region. The assembly evacuates the spallation heat deposit, guarantees the barrier between the LBE and the reactor hall, and assures optimal conditions for the spallation reaction. The assembly is conceived as an IPS and is easily removable or replaceable.

22.4 Conclusions

In order to reduce the amount of long-lived radioactive waste legacy and for the countries willing to continue the use of nuclear energy in their energy mix with limited amount of natural resources, the FC of the nuclear energy systems needs to be closed for plutonium and MAs. P&T is an attractive solution to reach the goal of sustainable nuclear energy systems. P&T could address the HLW issue and initiate the strategy toward a more resource-efficient nuclear energy system in the future. P&T will not replace the need for appropriate geological disposal of high-level waste, but it will reduce significantly the long-term radiotoxicity of the waste and thus will reduce the impact of HLW on the safety of geological disposal, therefore increasing highly its acceptability by the public.

All transmutation strategies with closed FCs could achieve high reductions in the inventory of actinides and the radiotoxicity of the nuclear waste by using fast neutron spectrum systems. The latest system being fast critical reactor or ADSs.

Nevertheless, the subcritical ADS offers interesting degrees of freedom with respect to the core design and fuel management as well as flexibility to evaluative policy toward the use of nuclear energy.

With the extensive development of accelerator technology, especially the linac technology, it appears today that proton beam powers of 10–100 MW with the needed reliability for operating ADSs are achievable in the coming decade.

List of Abbreviations

ADS: Accelerator driven system

CW: Continuous wave

DC: Direct current

FC: Fuel cycle

FP: Fission product

FR: Fast reactor

HLW: High-level radioactive waste

IPS: In-pile test section

LBE: Lead-bismuth eutectic

LLFPs: Long-lived fission products

MA: Minor actinide

MOX: Mixed oxide

MTBF: Mean time between failure

MYRRHA: Multi-purpose hYbrid Research Reactor for High-tech Applications

OECD: Organisation for Economic Co-Operation and Development

P&T: Partitioning and transmutation

RF: Radiofrequency

TRU: Transuranic

References

1. U.S. Energy Information Administration (EIA) (2011) International Energy Outlook 2011. Report No. DOE/EIA-0484(2011), U.S. Energy Information Administration (EIA), Washington, DC, September 2011, *www.eia.gov/oiaf/ieo/index.html* (accessed 16 February 2013).
2. OECD (2012) *OECD Factbook 2011–2012: Economic, Environmental and Social Statistics*, OECD Publishing. doi: 10.1787/factbook-2011-en
3. OECD/NEA (2006) *Physics and Safety of Transmutation Systems, A Status Report*, OECD Publishing, ISBN: 92-64-01082-3.
4. PATEROS (2006–2008) Partitioning and Transmutation European Roadmap for Sustainable Nuclear Energy, FI6W, Contract Number 036418.
5. M. Lecomte, *Le Traitement-Recyclage du Combustible Nucléaire usé* CEA, Editions Le Moniteur, 2008, ISBN: 978-2-281-11376-1.

Further Readings

Aït Abderrahim, H., Baeten, P., De Bruyn, D., Heyse, J., Schuurmans, P., and Wagemans, J. (2010) MYRRHA, a multipurpose hYbrid research reactor for high-end applications. *Nucl. Phys.*, **20**, 24–28.

Greneche, D., Quiniou, B., Boucher, L., Delpech, M., Gonzalez, E., Alvarez, F., Cuñado, M.A., Serrano, G., Luis Cormenzana, J., Kuckshinrichs, W., Odoj, R., von Lensa, W., Wallenius, J., Westlén, D., Zimmerman, C., and Marivoet, J. (2008), *RED-IMPACT – Impact of Partitioning, Transmutation and Waste Reduction Technologies on the Final Nuclear Waste Disposal*, Jülich Forschunszentrumn.

Mansani, L., Artioli, C., Schikorr, M., Rimpault, G., Angulo, C., and De Bruyn, D. (2012) The European lead-cooled EFIT plant: an industrial-scale accelerator-driven system for minor actinide transmutation. *Nucl. Technol.*, **180**, 241–263.

Martinez-Val, J. (2009) PATEROS P&T Roadmap Proposal for Advanced Fuel Cycles Leading to a Sustainable Nuclear Energy – Syntheses Report.

23
Fusion Energy by Magnetic Confinement

Friedrich Wagner

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

23.1 Introduction

The primary energy source of the universe is based on fusion processes. Fusion has been providing the cosmos with energy for 13.8 billion years. Developed on Earth, fusion would represent a clean and basically inexhaustible energy source. Worldwide, science and research are engaged in the development of this technology. In this chapter, we describe the fundamental principles of fusion and the status and prospects of fusion R&D on the basis of magnetic confinement. An easily readable introduction on all aspects of fusion for a not necessarily technical readership is Ref. [1]. There are many excellent textbooks on high-temperature plasma physics and fusion technology; the author recommends that in Ref. [2]. A more detailed version of this chapter is given in Ref. [3].

23.2 The Role of Fusion in the Universe

The big bang left leptons along with protons, deuterons, a little helium, and spurious traces of lithium. The rest of the 92 elements up to uranium are the spin-offs of the self-organization of the universe enforced by gravity – the formation of galaxies with stars consisting of hydrogen acting as the breeder material for new elements. In the core of the sun the temperature is 15 million K ($\sim$1.5 keV)[1] and the pressure is 10^{16} Pa. Under these extreme conditions, matter is in the fourth state – the plasma state. Like in the other cases – solid, liquid, gas – the new state emerges out of the previous one by a disintegration process through elevated temperatures. In the transition from gas to plasma, the constituents – molecules and atoms – are dissociated and ionized and free electrons and ions are produced – for example, protons in case of hydrogen as working gas. Although we are initially describing a process of disintegration, at the end, new elements are formed during the life cycle of stars by fusion of light elements to heavier ones.

The initial fusion process is the fusion of hydrogen to helium. The most basic process is the pp (proton–proton) I cycle.[2] In a complex step process, four protons

1) In plasma physics, temperatures are expressed in electronvolts (eV) and kiloelectronvolts (keV); 1 eV corresponds to 11 600 °C.

2) There are several chain processes starting from hydrogen and ending in helium. Under solar conditions, the pp I cycle is the most dominant one producing 86% of the He. The CNO cycle is dominant in stars hotter than the sun.

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

(^{1}H) yield helium (^{4}He). The individual steps are

$$^1\mathrm{H} + {}^1\mathrm{H} \rightarrow \mathrm{D} + \mathrm{e}^+ + \nu_\mathrm{e} + 0.42\ \mathrm{MeV} \quad (23.1)$$

$$\mathrm{D} + {}^1\mathrm{H} \rightarrow {}^3\mathrm{He} + 5.49\ \mathrm{MeV} \quad (23.2)$$

$$^3\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^4\mathrm{He} + 2\ {}^1\mathrm{H} + 12.86\ \mathrm{MeV} \quad (23.3)$$

The transformation of a proton into a neutron in the initial step (Eq. (23.1)) is governed by the weak interaction and is, therefore, an extremely slow process. This is essential for the lifetime of the sun and our planetary system. Besides a more massive element, He, a positron, e^+, and a neutrino, ν_e, are additionally created, which share the released energy. The total energy gained in this process is 26.72 MeV. This is the sum of the exothermal processes 23.1 and 23.2 taken twice and the process (23.3). In process (23.1), the annihilation γ-energy of e^+ with an abundant electron (1.02 MeV) is included in the energy balance (because the γ-radiation adds to the fusion pressure to be balanced by gravity). Less than 1% of the masses involved are transferred into energy.

Fusion processes yielding higher mass elements are exothermal, releasing vast amounts of energy. The energy process behind fusion is the transformation of mass into energy according to Einstein's most famous relation $E = mc^2$, which is one of the consequences of the special theory of relativity published in 1905.

The fusion process between positively charged nuclei has to overcome the Coulomb repulsion. The potential energy at a distance of the proton diameter is about 0.6 MeV. But Nature eases fusion because the particle, the proton, adopts in this interaction the characteristics of a wave, which allows it to leak into the energetically prohibited zone and to finally tunnel through the Coulomb barrier into the deep bound states of the newly formed nucleus. The potential energy at the distance of the de Broglie wavelength is about two orders of magnitude lower than the one within the range of the nuclear forces.

The fusion yield depends on the density of the particles, their relative velocity, and the nuclear interaction in the form of the fusion cross section σ_{fus}. Fusion is only one form of interaction in a plasma, even the one with the smallest cross section. The competing interaction processes are Coulomb collisions between charged particles. As the Coulomb collision rate is large compared to the fusion rate, the proton assembly (as well as one of the electrons) can be described by the Maxwellian distribution. The governing kinetic parameter is temperature rather than the individual particle energy in the assembly interactions – scattering and fusion. This is the reason why we are discussing controlled thermonuclear fusion.

23.3 Processes for Technical Fusion

Fusion scientists are trying to realize a technology, which basically transfers mass into energy. (Of course, this is again only a transformation process; energy is conserved when mass is transferred, e.g., into radiation.) A gram of matter transferred into energy, for example, via processes (23.1–23.3), would yield 10 GWh. This energy would serve the members of a typical OECD (Organisation for Economic Co-operation and Development) family throughout their life. However, technical fusion cannot resort to the process acting in

the interior of the sun. A reaction has to be selected that circumvents the weak interaction and maintains the number of protons and neutrons. There are many such processes with different interesting features, such as the avoidance of free neutrons and thus the avoidance of neutron-induced radioactive fusion waste. The selection of the proper reaction is guided, however, by searching for the highest reaction rate providing the highest fusion yield. This condition is fulfilled by fusing deuterium $D = {}^2H$, the heavy isotope of hydrogen, with tritium $T = {}^3H$, the superheavy isotope. The reaction reads:

$$D + T \rightarrow {}^4He(3.52\ MeV) + n(14.06\ MeV) \qquad (23.4)$$

The fusion energy per process $E_{fus} = E_\alpha + E_n = 17.58\,MeV$. The DT fusion cross section is plotted against energy in Figure 23.1 and compared with the alternative DD fusion process. DT fusion has the higher cross section and the maximum is at lower energy: larger fusion yield is possible at less demanding conditions. The basic reaction (23.4) highlights the major characteristics of fusion energy: there is neither release of radioactive process products nor CO_2. Deuterium can be gained from seawater. The DT reaction is, however, not neutron-free. The neutron is actually the dominant energy carrier. It activates structural material and fusion energy leaves radioactive waste, which, however, quickly decays and is characterized by a comparatively short-lived radiotoxicity (see Section 23.11.2).

The reaction products, α-particle and neutron, have additional tasks. The α-particle provides the necessary inner heating of the plasma to maintain the conditions for fusion; the fusion neutron allows producing tritium via a breeding reaction with lithium.

Tritium is a radioactive substance with a short half-life of 12.3 years. Therefore, it is naturally not available in any relevant concentration rather a technology has to be developed for its production. From Eq. (23.4), the annual tritium use of a fusion power station with $3\,GW_{th}$ power can be calculated to be – in an idealized case – 167 kg. The breeding technology of

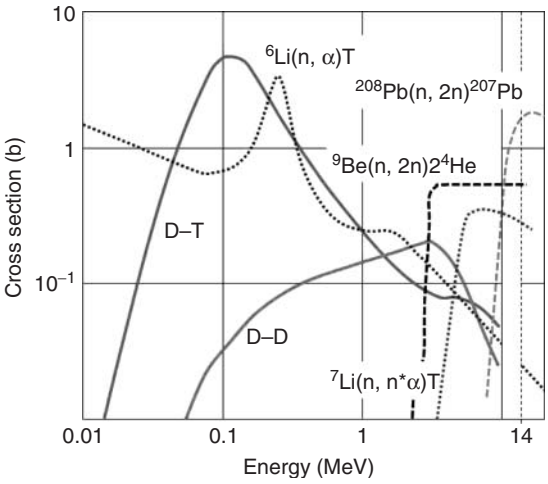


Figure 23.1 Fusion-relevant cross sections are plotted against energy. In case of the fusion reactions, D–T and D–D, the energy is the kinetic energy of the incident particle; in case of the Li-breeding processes and the (n,2n) processes, it is the energy of the incident neutron.

tritium is based on the nuclear interaction of a neutron with lithium (Li), which yields tritium. There are two (stable) Li isotopes of relevance: ^{6}Li and ^{7}Li. Natural lithium contains 7.5% ^{6}Li and 92.5% ^{7}Li. The breeding reactions are

$$n + {}^6Li \rightarrow T + {}^4He + 4.8\,MeV \qquad (23.5)$$

This process, involving a thermal neutron, is exothermic and represents additionally a substantial contribution to the overall fusion energy yield.

$$n + {}^7Li \rightarrow T + {}^4He + n^* - 2.5\,MeV \qquad (23.6)$$

This process with fast neutrons is endothermic corresponding to a mass difference $\Delta m \sim -0.0027$ u, whereas $Q = \Delta m^*931.5$ MeV; the resulting n* has a reduced energy, which is distributed according to the kinetic conditions of its release.

The neutron necessary to initiate the breeding process emerges from the central DT fusion itself, starting with an energy of 14.1 MeV. The breeding of tritium is a process that happens within the reactor chamber. This is an important safety aspect. In particular, there is no need of external transportation of tritium apart from the initial fueling of a fusion reactor to start the fusion and breeding processes.

The processes (23.5) and (23.6) have highly different cross sections – both in magnitude and energy dependence (see Figure 23.1). The (n,α) cross section of ^{6}Li has the classical form with a $1/v$ dependence at low velocities, v, and a sharp maximum corresponding to a resonance with an energy state of the compound nucleus ^{7}Li. The cross section of the ^{7}Li reaction has the typical shape of a threshold reaction with a fast rise of σ beyond a critical energy.

Tritium is produced in a breeding blanket, which surrounds the fusion plasma and absorbs the neutrons from the fusion process. The fusion blanket is a complex compound structure because it has to fulfill three major tasks–: (i) produce tritium and enable collecting it, (ii) transfer the kinetic energy of the neutrons into a cooling medium, and (iii) screen the superconducting coils used for magnetic confinement (see Section 23.6.) from the central neutron flux to avoid material damage.

23.4 The Conditions for Controlled Nuclear Fusion

The plasma parameters relevant to fusion and the fusion conditions are obtained from an energy balance. It is easy to formulate the conditions for controlled nuclear fusion but it is difficult to establish them. John Lawson published the Lawson criterion in 1957 [4]. Different formulations exist depending on the terms and processes considered – whether the α-particle heating is equated to the power losses (ignition condition), whether the total fusion power is related to the external heating power P_{ext} (break-even condition), whether part of the fusion power is recycled to provide electricity for the plant operation, whether the ash concentration (He, α-particles) is consistently considered, or whether radiation losses are included. The outcome of such balances yields average values and can only be considered as rough guidance. The design of a fusion reactor such as the International Thermonuclear Experimental Reactor (ITER) (see Section 23.9.1) following the definition of its objectives requires detailed analysis as, for example, summarized in Ref. [5].

The power balance for the ignition condition including radiation is

$$p_\alpha = \frac{1}{4} n_{DT}^2 \langle \sigma_{fus} v \rangle E_\alpha = \text{constant } n_e n_z Z^2 \sqrt{T} + \frac{3}{2} \frac{(n_{DT} T)}{\tau_E} \quad (23.7)$$

The left-side term represents the α-particle power density deposited inside the plasma maintaining the self-sustained burn. The first term on the right side represents radiation,[3] and the second one represents the plasma transport losses parameterized by the energy confinement time τ_E. n_{DT} is the deuterium–tritium density; n_e, the electron density; and n_Z, the impurity density (e.g., He or those caused by wall erosion). T is the temperature. The factor 1/4 results from the particle balance: $n_{DT} = n_e = n_D + n_T$ and an equal mix between D and T.

We encounter the term *energy confinement time* τ_E for the first time, although we have already used the phrase ''magnetic confinement''. The energy confinement time is the energy replacement time. In case of a house, τ_E would represent the cooldown time when the stove is turned off in winter. In case of good thermal insulation, τ_E is long. In case of a plasma, τ_E is determined by plasma transport. The plasma core is heated externally with P_{ext} or – after ignition – internally with the α-particle heating power P_α. Through heat conduction and convection, this power is transported to the plasma edge and radiated off or deposited onto specific technical installations – limiters or divertor targets. The transport rates are characterized by τ_E, which is a measure of the thermal isolation of the plasma. The global power balance can be written under steady-state conditions as $\dot{W} = 0 = P_{ext} + P_{int} - (W/\tau_E)$. For the same heating power, the energy content W doubles if τ_E doubles. Like in a house, a large stove (a large external heating power) is required in case of bad thermal insulation. This is not an economic approach for maintaining a desired temperature. In analogy, as τ_E is such a crucial parameter for confinement concepts and operational plasma conditions, the study of transport and confinement is at the core of fusion research.

Power balance (Eq. (23.7)) can be cast into the form $n\tau_E = f(T)$. This relation is plotted in Figure 23.2 for various assumptions – the break-even condition when the fusion output matches the external heat input (''the stove is puffing'') and the ignition conditions (''the stove transits into self-sustained burn''). The curves have a minimum due to the maximum in the fusion cross section (see Figure 23.1). The other cases shown in Figure 23.2 reflect the role of impurities. An unavoidable ''impurity'' is He, the fusion ash. Its concentration is determined by a particle balance, which is similarly structured in Eq. (23.7). The fusion process itself provides the particle source: 1 MW of fusion power creates 3.5×10^{17} α-particles/s. Under steady-state conditions, He is transported to the plasma edge and removed out of the system. The rate of this process is represented by a particle confinement time τ_p. As an empirical fact as well as expected from transport theory, τ_p and τ_E are linked: τ_p and τ_E vary together with confinement regimes and show similar parameter dependencies. In Figure 23.2, $\rho = \tau_p/\tau_E$ is assumed constant and varies between 3 and 40. Impurities are released by plasma wall interaction from limiters, divertor target plates, and the surrounding

3) In this simplified approach, we assume only bremsstrahlung radiation.

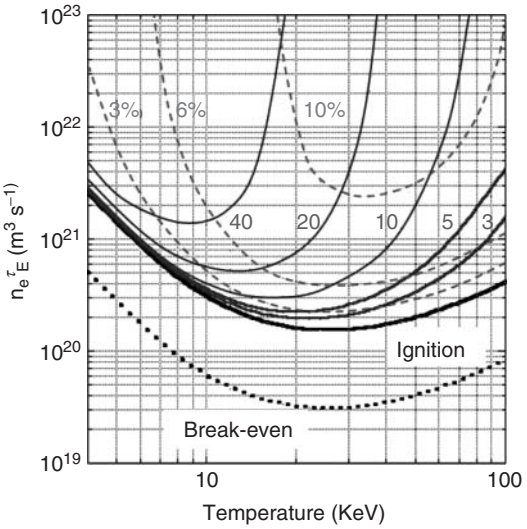


Figure 23.2 Plotted is the fusion product, electron density n_e times the energy confinement time τ_E, against the plasma temperature. The thick solid line represents the ignition condition; the dotted line, the break-even one. The marked curves represent the ignition condition with impurities. The cases for helium are shown by solid lines and the respective numbers are the ratio of particle to energy confinement time $\rho = \tau_p/\tau_E$; the dashed lines show the impact of carbon radiation ($Z = 6$) at different concentrations.

walls as preferential impurity sources. The dashed set of curves in Figure 23.2 shows the impact of carbon radiation on the ignition conditions when the carbon concentration is varied from 3 to 10%.[4] In the ideal case, the ignition conditions require a temperature of about 20 keV and a confinement product $n_e\tau_E$ of $\sim 2 \times 10^{20}\,m^{-3}s$. With impurities, the fusion conditions become more critical because of dilution of the fuel and excessive core radiation. Both fuel dilution and radiation lead to an increase in the $n_e\tau_E$ ignition limit. The maximal He concentration of 50% translates into a sharp upper temperature limit for ignition, which depends on ρ (see Figure 23.2). This moves the $n_e\tau_E$ minima toward lower temperatures. Radiation losses cause the temperature minima to move to higher values. In this case, there are two limits whenever the radiation losses match the α-particle source term. Because of the characteristic energy variation of the cross section (see Figure 23.1), this happens at two temperatures.

Figure 23.2 makes it plausible that the He ash has to be exhausted in a reactor and that fusion plasmas have to be clean, which requires judicious material selection. The technical tool for the exhaust of energy and particles is the divertor, which acts as the "ash pan." With divertor, the plasma wall contact for exhaust is moved into a separate chamber (see Figure 23.3) so that material contact close to the plasma surface is avoided. The curves in Figure 23.2 are based on carbon $Z = 6$ radiation. Carbon is a favored first-wall material in fusion experiments. Because of its strong hydrogen absorption, it leads to high tritium inventories (see Section 23.10.1) and has to be discarded. The favored inner wall material for a fusion reactor is tungsten. The radiation curves in Figure 23.2 represent tungsten with the impurity concentrations reduced by a factor $(Z_W/Z_C)^2 \sim 150$.

4) Carbon is used here because the wall of fusion experiments is often fully covered with graphite tiles, a material that does not melt. Carbon as a plasma impurity radiates at the plasma edge, which can be tolerated or is even welcome.

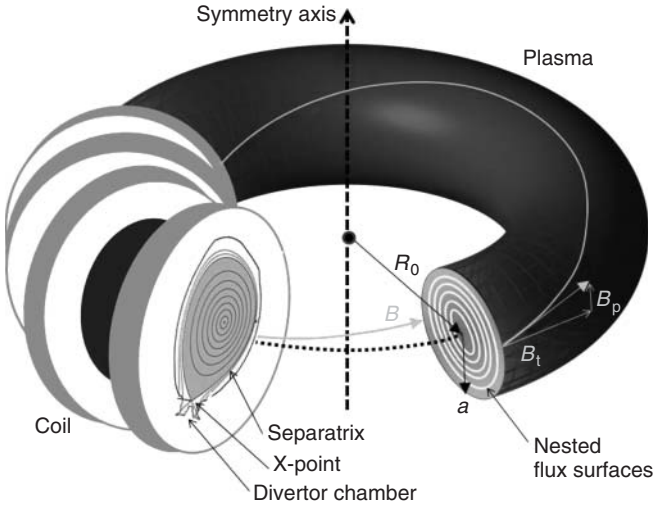


Figure 23.3 Schematic of a toroidal plasma with four toroidal coils indicated. Characteristic quantities such as the toroidal and poloidal fields, the nested flux surfaces on the right cut, and the separatrix of an elongated plasma cross section with the X-point on the left cut, as well as the divertor chamber for exhaust (left cut), are indicated.

There are two ways to satisfy the confinement product $n_e\tau_E$. One is based on large density and low τ_E, the other one is based on opposite pairing – large τ_E and low density. The first case is the goal of inertial confinement fusion research, the topic of Chapter 24. The second case is the topic of this chapter.

23.5 The Tasks of Fusion Research and the Basic Features of Fusion Energy

The major tasks of fusion R&D on the basis of magnetic confinement is to confine a hot, high-pressure DT plasma of about 1000 m^3 volume by strong superconducting coils, producing a field of ~10 T. Magnetic confinement has to provide plasma equilibrium and stability and low transport rates. The plasma has to be heated externally by complex techniques to about 20 keV; for example, by injecting ~100 keV hydrogen atoms[5] or by using electromagnetic waves. Under reactor conditions, the α-particle heating, emerging from fusion processes in the plasma core, takes over and the system enters the phase of a self-sustained burn. The external heating can be turned off. Techniques to remove effectively helium, the ash of the fusion process, have to be developed. Also the plasma has to be kept clean. Impurities would strongly interact with the fuel and contribute with unwanted electrons and give rise to excessive radiation (see Eq. (23.7)). Deuterium can be obtained from seawater by electrolysis. In case of fusion, electrolysis is an inexpensive way to produce deuterium. Tritium, produced in the blanket, has to be stored and distributed in such a way that the total tritium system inventory

5) A plasma of the line density of ITER requires a D^0 injection energy of ~1 MeV.

is kept small, specifically the fraction of volatile tritium. The blanket is a complex technical component. In addition to its tritium breeding function, it has to absorb the neutron heat, about 2–3 MW m^{-2}, and transfer the power safely to a cooling system, water, or gaseous helium. For this purpose, fusion technology has to develop materials to withstand high heat and neutron fluxes resistant to material damage. The final production of electricity seems to be the simplest task, which will employ standard steam techniques. The electrical system efficiency will be in the range of 0.35–0.4. For higher efficiencies, the operational temperatures have to be increased, which has limits because of the complex compound structures used in fusion technology and the loss of material quality toward higher temperatures.

The benefit of this development is a technique, which is capable of delivering large amounts of electricity without serious environmental side effects. The products of the fusion reaction have no environmental relevance, specifically fusion will not produce CO_2. The primary fuels, D with a concentration of 0.016% in water and Li, are basically inexhaustible and rather homogeneously distributed over the Earth, specifically when we consider Li extraction from seawater.[6] Fusion power is intrinsically safe: the fusion process is based on binary collisions and not on an avalanche process, which necessitates continuous control; the energy density inside the plasma is low; the plasma process stops at the slightest external interference; and the involved core energies cannot destroy the containment.

It has to be ensured, however, that tritium is well confined. Tritium, being radioactive, easily enters the biosphere via HT (tritiated hydrogen) and HTO (tritiated water) formation. In Section 23.11, we discuss the tritium safety problem in more detail. Another aspect to consider is fusion waste. Although fusion produces neither fissile products such as caesium or iodine nor plutonium or minor actinides (MAs), the fusion neutrons activate the minorities inside stainless steel structural material, leaving radioactive waste that has to be stored. It is important to note, however, that the time for storage is typically about 100 years and is not determined by long decay half-lives. This aspect is discussed further in more detail later.

6) The economy of this process has not been considered in all details.

23.6 The Basics of Magnetic Confinement

Like in the sun, the conditions in the core of the plasma have to fulfill the fusion conditions as formulated in Section 23.4. Under the conditions of magnetic confinement, the pressure gradients have to be balanced by the Lorentz force to establish an equilibrium state of force balance. In a magnetic field, charged particles move in a helix due to the Lorentz force $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, where q is the charge; $\mathbf{v}$, the individual particle velocity; and $\mathbf{B}$, the magnetic field. In the direction perpendicular to the field, the excursion of the particle is limited to the Larmor radius ρ_L. Because of the specific form of the Lorentz force as vector product of velocity and field, the particle momentum parallel to the field is not changed. Therefore, in a homogeneous field, magnetic confinement is insufficient because it applies to the perpendicular velocity component only. One has to involve and to accept systems with field inhomogeneity to improve confinement.

Magnetic confinement systems differ in the way they cope with the confinement parallel to the field, thus defining classes of confinement systems.

In linear devices – the first confinement class – parallel confinement is improved utilizing the mirror effect. The basis of the mirror effect is the magnetic moment $\mu = W_\perp/\mathrm{B}$ of a magnetized charged particle, which is a constant of motion. An electron or ion, which is thermally agitated with velocity $\mathbf{v} = (v_\perp, v_\parallel)$ and moves from a zone of low field $B_{\min}$ into that with higher field, increases its perpendicular velocity $v_\perp$ as a consequence of μ being constant. On the other hand, the kinetic energy of the plasma species is also a constant of motion with the corollary that the parallel velocity $v_\parallel$ decreases. If the field is sufficiently large, $v_\parallel \to 0$. At this field, the mirror field $B_{\max}$, the particle comes to a full stop. Thereafter, it is accelerated back to the low-field zone by the force $-\mu \times \mathrm{grad}_\parallel \mathrm{B}$. If the field system is built symmetrically, the game repeats itself at the other side with the outcome that the mirror effect causes the particles to bounce between two mirror points and to stay in the neighborhood of the field minimum, thus improving confinement. In an actual mirror machine, characterized by $B_{\min}$ and $B_{\max}$, two classes of particles, discriminated by $v_\parallel/\mathrm{v}$, have to be considered. If this velocity ratio is small, a particle is reflected by the magnetic mirrors, as described earlier, and represents a trapped one. Otherwise, it is slowed down along its trajectory toward higher field but does not reach the point with $v_\parallel/v = 0$. These particles escape the mirror. The boundary is given by $v_\perp^2/v^2 = B_{\min}/B_{\max}$ specifying a loss cone in phase space spanned by $v_\perp$ and $v_\parallel$ as coordinates.

The confinement of simple mirror machines is insufficient because the transparency of the mirror is too large. Still, we continue describing the confinement situation of the mirror because the concepts help us to understand the physics of more relevant confinement systems. Until now, we have ignored collisions between the plasma species. Of relevance are the collisions in phase space of the trapped particles occupying the low-field zone scattered into the empty loss cone. Because of their higher velocity, electrons are scattered more frequently into the loss cone than the ions. As a consequence, the plasma is polarized and charges up positively. The formation of an electric field keeps back the escaping electrons by electrostatic means and accelerates the ions. For a transport equilibrium, the so-called ambipolar electric field $\mathbf{E}$ enforces flux equality $\Gamma_\mathrm{e} = \Gamma_\mathrm{i}$ between the differently charged plasma species of different masses and different kinetic properties.

A confinement concept that avoids mirror losses is toroidal confinement realized in a plasma ring as shown in Figure 23.3. In this case, the magnetic coils are arranged in a closed ring thus avoiding end-losses. This geometry defines the class of toroidal confinement systems. An unavoidable feature is the inhomogeneity of the toroidal field with a higher field closer to the vertical symmetry axis. The field gradient points to the symmetry axis. This field inhomogeneity leads to rather unpleasant consequences: the Larmor orbits of the charged particles are not closed causing drifts. Ions and electrons drift parallel to the symmetry axis and leave their magnetic field line. The drift of electrons and ions is in the opposite direction, leading to an additional vertical electric field. A second parasitic drift appears now in the crossed-field

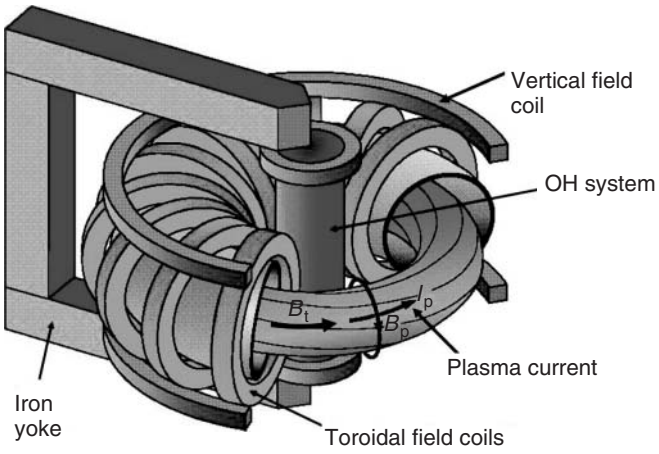


Figure 23.4 Principle setup of a tokamak with circular cross section.

arrangement of vertical electric and toroidal magnetic fields. Like in the Hall effect, an $\mathbf{E} \times \mathbf{B}$ drift appears perpendicular to both E and B_{tor} directions, which causes the plasma torus to expand radially. No force equilibrium is established.

The saving idea was to introduce rotational transform. A second field component B_{pol} is introduced with a perpendicular (poloidal) component. A field line with the components (B_{tor}, B_{pol}) winds around the torus in a helix (see Figure 23.3) and no longer stays in a plane rather maps out a toroidal volume. The vertical charge separation is avoided because the up-down sides of the torus are short-circuited by the helical field lines connecting these regions. Macroscopic radial equilibrium can be provided.

The essence of magnetic confinement is the development of a system of nested toroids with field lines, which encircle the tori toroidally and poloidally and with a field line density, which is higher at the inside than the outside (torus effect). The nested toroids, called *flux surfaces* (see Figure 23.3), are magnetically isolated and not connected by a field line with a net radial component.

The way the poloidal field component is produced defines confinement classes within toroidal systems. The simplest one is using the potential of a plasma to carry a current. A ring current flowing inside the plasma, the plasma current I_p, produces the poloidal field component B_{pol}. As the temperature of the electrons, defining the electrical conductivity of a plasma, is highest in the plasma core, the current density profile $j(r)$ peaks there. Such systems are called *tokamaks* [6] and are a Russian invention of the 1950s.[7] Figure 23.4 shows the principal setup of a tokamak. The major attraction of the tokamak is that I_p can be produced by a pulsed transformer, whereas the plasma ring surrounding a central primary coil acts as the secondary winding. In tokamaks, the ratio of $B_{tor}/B_{pol} > 1$. Another concept with plasma currents is the reversed field pinch (RFP) with $B_{tor}/B_{pol} \sim 1$ [7]. Tokamaks and RFPs belong to the category of internal confinement systems because part of the confining magnetic field is produced by currents flowing inside the plasma.

7) The acronym tokamak means toroidal chamber with magnetic coils.

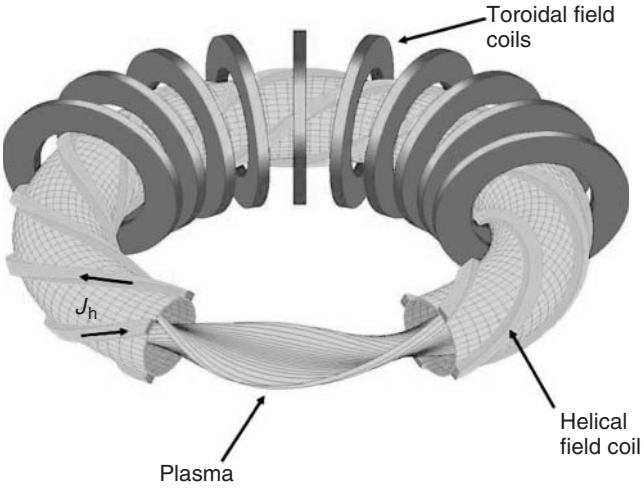


Figure 23.5 Principle setup of a classical $l = 2$ stellarator with helical coils.

An alternative way is to produce the poloidal field like the toroidal one – by external coils. In this case, the coils have to wind helically around the torus. The field composition resembles a multipole arrangement, with coils being first helically twisted and then bent to a torus. Depending on the current direction inside the helical coils – unipolar or bipolar – we speak of heliotrons/torsatrons or of stellarators. Both belong to the class of helical systems. Stellarators are a US invention of the 1950s. Figure 23.5 shows a sketch of a classical $l = 2$ stellarator with helical coils. Helical systems belong to the category of external confinement.

Toroidal systems share common descriptors. In the simplest case with circular poloidal cross section, the torus geometry is defined by the major radius R and the minor radius a (see Figure 23.3). The ratio $A = R/a$ is called the *aspect ratio*. The rotational transform ι of a flux surface is defined by the winding law of the field lines given by $\iota/2\pi = AB_{\mathrm{pol}}/B_{\mathrm{tor}}$. RFPs have a larger ι than tokamaks. In case of helical systems, B_{pol} is determined in a complex form by the external helical coil system, which is described by the number of coils; for example, three heliotron coils with unidirectional current yielding an $l = 3$ heliotron. In this case, the poloidal cross section is triangular. In case of a stellarator with four helical coils and bipolar currents, we have two pairs of coils and an $l = 2$ stellarator (see Figure 23.5). Its poloidal cross section is elliptical. Another important stellarator descriptor is the coil pitch. The coil winding laws are such that the coils ultimately close. The field pattern of helical systems is periodic in toroidal direction. The toroidal periods can be $m = 5$ (Wendelstein 7-A, Germany [8]) or in case of tight windings, for example, $m = 10$ (Large Helical Device, LHD, Japan [9, 8]).

Simple tokamaks have a circular cross section. Equilibria at higher plasma currents are possible when the cross section is elongated or is even triangular. Modern tokamaks benefit from the higher current

- reaching higher plasma density – a critical ignition and burn parameter (see Section 23.4);

- higher stability in terms of the ratio of $\beta = \langle p_{\mathrm{kin}} \rangle / (B^2/2\mu_0)$, average kinetic pressure to magnetic field pressure;
- better confinement, expressed by a larger energy confinement time τ_{E}.

We understand the importance of the plasma core in providing conditions for a high fusion yield. Also the surface of the confinement system, the so-called last closed magnetic surface, is of utmost importance. The plasma surface is defined by a separatrix (see Figure 23.3), which separates (therefore the name) the nested flux surfaces inside from a belt of open field lines at the outside in the so-called scrape-off layer (SOL). The separatrix if formed by a null in the poloidal field B_{pol}, which itself is determined by the superposition of the poloidal field of the plasma current with that of additional external poloidal field coils. An X-point is formed (see Figure 23.3), where the field is exclusively toroidal. The rotational transform of the separatrix is $\iota = 0$. Outside the separatrix, the field lines move around the plasma, are diverted at the X-point, are separated from the main plasma, and finally are guided into a separate chamber, the divertor chamber (see Figure 23.3). This is the basic geometry for exhaust and for the establishment of the plasma wall interaction inside the divertor chamber, as also foreseen for the ITER and a future fusion reactor. The α-particle power (apart from the fraction, which is radiated off at the plasma periphery, in the SOL and the divertor plasma) is transferred onto target plates located inside the divertor chamber. The target plates have to withstand high heat fluxes in the range of $10\,\mathrm{MW\,m^{-2}}$; the divertor chambers are equipped with pumping systems to exhaust He and to control plasma density and composition. Impurities released from the contact zones with the plasma are mostly retained inside the divertor. The magnetics of divertors is still an area of R&D with new concepts emerging, dubbed Super-X [10] or snowflake divertor [11].

Also, modern stellarators are not simply set up by coils described by l and m. Like in the tokamak case, shaping is essential to improve and optimize plasma performance. The most effective optimization is possible within the concept of multihelicity stellarators with a mix of poloidal and toroidal l and m parameters. In this case, the plasma cross section varies toroidally changing cross section according to the dominant local field Fourier coefficients. The benefit of such a configurational complexity is described in Section 23.9.2. It is difficult to realize multihelicity stellarators using helical coils. Instead of separate sets of toroidal and helical coils, the so-called modular coils are used. These are basically individual toroidal coils, which, however, are not planar like in the case of a tokamak rather show specifically shaped lateral excursions (see Figure 23.11), that line up such that a poloidal field component is generated. Unlike the tokamak, the stellarator does not need additional coils to produce a separatrix. Stellarators are limited by a separatrix as a natural feature of their magnetic system.

Both lines, internal and external confinement, are pursued because of complementary properties. The plasma current of tokamaks represents a strong ohmic heating source for the plasma, P_{OH}, and one can easily reach temperatures around a kiloelectronvolt with it. For RFPs, the ohmic heating alone may be sufficient to reach reactor conditions. An important advantage of the tokamak is its continuous toroidal symmetry. Hence, the toroidal angular coordinate Φ is an ignorable one;

the consequence is that the associated canonical momentum p_Φ is a conserved quantity. As discussed later, this aspect enters the quality of magnetic confinement.

We have discussed the role of field inhomogeneities in case of the mirror device. Toroidal systems also carry a mirror because of the radial field gradient. In a tokamak, there are two classes of particles like in the mirror device. Particles with small $v_\parallel/v$ are affected in their orbit by the increasing magnetic field on their way along the field line toward the inner sector of the torus. If $v_\parallel/v$ is too small, the particles are reflected at $B_{\max}$. In this case, they neither freely propagate along the field line nor experience the full rotational transform, rather they stay at the outside between two mirror points. They are dubbed ''trapped particles.'' Their drift in the field gradient is not fully compensated and they show, therefore, deviations from the flux surfaces much larger than the Larmor radius. Their movement is periodic, however, and they are confined. This is also ensured for the other class of particles. They move freely from the torus outside to the inside and back. They are called *free particles*. Their deviation from the flux surface is small but modified by the curvature of the toroidal field system. The strength of the axisymmetric tokamak is that particles move along periodic orbits with finite deviations from the flux surface.

The negative side of the tokamak is the need to inductively produce the current like in a large transformer with the plasma ring representing the secondary winding. This limits the plasma pulses. At the end – maybe even after a few hours – the primary transformer system is exhausted and has to be recharged within minutes. It is not yet clear whether a pulsed system with these features qualifies for a power plant in the gigawatts of electrical power range. Therefore, noninductive current drive techniques are developed using the injection of fast particles or directed electromagnetic waves. The optimization of external current drive for tokamaks is a dynamic R&D field.

Other problems emerging from the presence of a strong plasma current are current-driven instabilities. In a tokamak, the existence of an equilibrium is intimately connected with the existence of the plasma itself and its quality to carry a current. There are nonlinear relations between equilibrium, transport, and stability properties, which allow the development of instabilities causing degradation up to the limit of self-destruction of the plasma in a disruption. In case of a reactor, 20 MA or so may thus be switched-off within a fraction of a second. The electromechanical dynamics of such a perturbation leads to high voltages or high currents, I_{ind}, induced in structural components leading to strong $I_{\mathbf{ind}} \times \mathbf{B}$ forces, which have to be safely handled.

An interesting aspect of the RFP is that the current is maintained self-consistently by a dynamo process inside the plasma. Strong magnetic turbulence adds up in a coherent way to produce a plasma flow with velocity v_{dyn}, which induces an electric field $\mathbf{E_{dyn}} = \mathbf{v_{dyn}} \times \mathbf{B}$, which itself drives poloidal current giving rise to most of the core toroidal magnetic field. The catch of this simple concept is the lack of nested flux surfaces being replaced by a rather chaotic field pattern caused by a high magnetic turbulence level allowing only low confinement. High transport levels are the major concern for RFPs. Although the programmatic validity of the RFP might be questioned, its plasma has rich physics resembling in some aspects

those observed in stellar plasmas allowing to study astrophysics phenomena in the laboratory. It has also interesting aspects of self-organization of thermodynamically open systems, which are presented in Section 23.7.

The plasma geometry of stellarators is invariably three-dimensional. No equilibrium can be constructed with continuous symmetry in the current-free case when $\mathrm{rot}\mathbf{B} = 0$. In addition to the toroidal curvature effect, helical systems show a further field inhomogeneity, which is introduced by the helical coils (or by the aligned lateral coil excursions of modular coils). The field beneath the helical coils is larger than the one between these coils. As a consequence, there is a third class of particles complementing the free and toroidally trapped ones – the helically trapped particles. These particles are highly localized in phase space and benefit only a little from the confining effect of rotational transform with the consequence of a net radial drift out of the system. They lead to intolerable losses under reactor conditions.

A more fundamental treatment of particle confinement starts from the three-dimensional field geometry of a helical system, the lack of continuous symmetry replaced by discrete toroidal periodicity. In such a case, there is no toroidal (tokamak, RFP), poloidal (mirror device), or helical (linear stellarator) symmetry established as, in principle, possible for magnetic confinement. The corollary of the lack of symmetry is that no component of the canonical angular momentum is conserved and that collisionless losses in phase space occur. As this ''loss cone'' is continuously filled by collisions, helical systems have difficulties to meet the ignition and burn conditions unless the problem with the lack of symmetry is removed. This is the story of the so-called quasisymmetric stellarators as described in Section 23.9.2.

23.7 Characteristics of High-Temperature Fusion Plasmas

A fusion plasma has to fulfill many requirements – equilibrium, stability, good confinement, plasma purity, and other qualities. This requires a rigorous strategy of optimization. It involves not only high-level physics issues such as the understanding of turbulent plasma transport and its ''taming'' but also technical issues such as the qualification of compound structures able to withstand the contact with plasma. Here, we will suffice ourselves by only discussing one aspect, plasma transport and the resulting confinement, as it is the most critical issue and possibly the most interesting one from the science point of view. Reference [12] represents a comprehensive review of the physics of magnetic confinement.

Plasma transport has two branches – laminar and turbulent. The laminar processes are based on the Coulomb collisions between charged species, which cause particles to jump from one flux surface to another and – according to the pressure gradients – from the inside to the outside as the net effect. In a homogeneous plasma, the characteristic step parameters of this diffusive process are the Larmor radius ρ_L and the Coulomb collision time τ_{Cb}. In toroidal geometry, these two parameters are modified because of the induced drifts. Of relevance are the trapped particles because of their large step size of $\sim \rho_L B_{tor}/B_{pol}$, the Larmor radius in the poloidal field. The effective collision time is determined from the collisional flow of particles across the boundary in phase

space separating free and trapped particles. The laminar transport processes, which are described by a complete theoretical edifice, the theory of neoclassical transport [13], do not play a major role in the tokamak losses. This is different in case of stellarators. Under fusion-relevant conditions, the collisionless losses of helically trapped particles become dominant. Like in the case of mirror devices, their losses are not intrinsically ambipolar, causing a radial electric field, which itself affects particle orbits, specifically those of ions. The transport equation becomes nonlinear: the electric field reduces the losses and also leads to different confinement branches featuring widely different plasma transport characteristics.

Bifurcations in laminar transport can happen under 3D conditions but more so with turbulent transport irrespective of the system symmetry. Like the sun, technical plasmas are highly turbulent [14]. Thermodynamically, plasmas represent open systems characterized by nonlinear relations, which also govern turbulent transport. Pressure gradients and plasma currents represent the free energy, which drive plasma turbulence. There are destabilizing properties such as plasma resistivity or geometrical aspects such as the colinearity of pressure and field gradients at the outer half-sector of the torus, or the so-called rational flux surfaces, where a field line does not ergodically map out the flux surface rather closes upon itself after a number of toroidal and poloidal circulations. Resistivity decouples the conducting medium (the plasma) from the magnetic field; the diamagnetic nature of the magnetized plasma causes it to move outward into the areas of low magnetic field. This tendency is supported by the pressure gradient at the torus outside (and opposed by that inside). In the third case, the topology of a global instability as an eigenmode of the bound geometry coincides with the topology of the rational field line closing upon itself. Under these circumstances, a mode or an instability does not have to invest energy for the distortion of the field topology and can grow with less prohibition.

The transport relevant turbulence scales with the Larmor radius. The nonlinear relations between stabilizing and destabilizing contributions give rise to onset conditions and define critical gradients for temperature and pressure. If, for example, the temperature gradient grows beyond a critical threshold value, the level of turbulence rises strongly. Formally, the diffusivities depend on the driving force gradient. Therefore, gradients are clamped to the critical ones.

Here, we can understand why fusion requires the size of a reactor to demonstrate its scientific feasibility. In order to meet the high temperature of 20 keV in the plasma core ensuring a high fusion yield, steep gradients are necessary. These gradients are, however, governed by critical onset conditions and limited to critical values. The situation is reminiscent of a sandpile. In an effort to increase its height, critical gradients send excess sand in little avalanches to the base of the pile. Only by broadening the base, the peak can be raised obeying the overriding critical condition. In principle, the same is observed in fusion plasmas. In order to meet the core conditions, the volume has to be increased leaving fixed the normalized temperature gradients $\nabla T/T$. Thus, the plasma volume increases and more fusion power is produced with the upshot that fusion can only be demonstrated at the power level of a few hundred megawatts.

It is therefore important to overcome the conditions leading to profile similarity and to improve confinement. This would

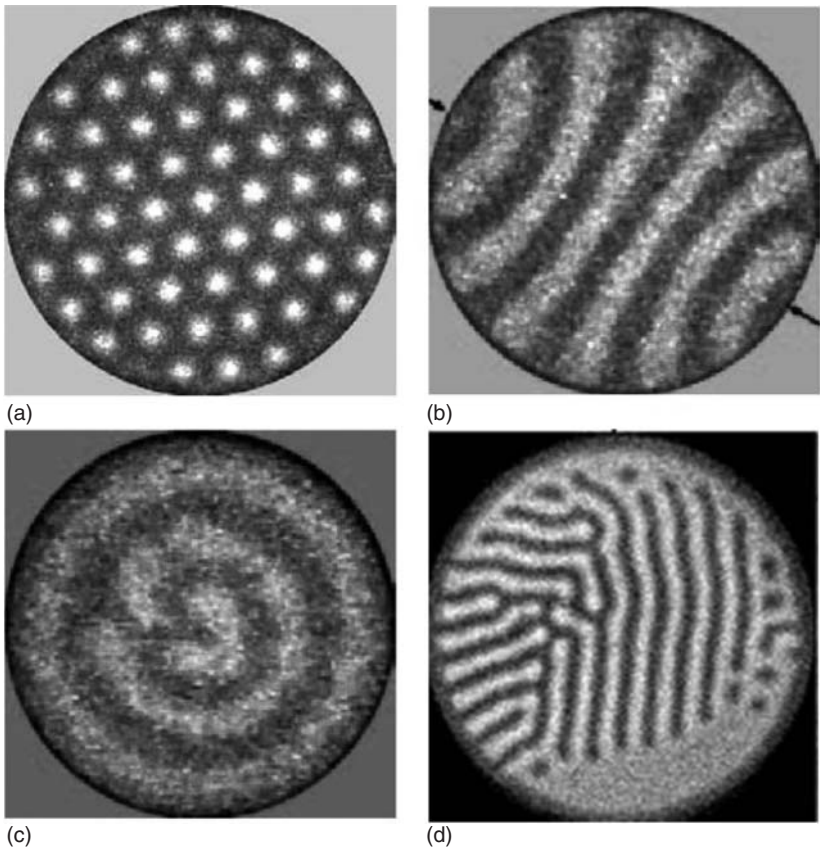


Figure 23.6 Examples of current density distributions of dielectrically suppressed diode discharges: (a) a hexagonal pattern, (b) a pattern with periodic stripes, (c) a rotating spiral pattern, and (d) a case with bright filaments. These data are obtained by the Research Group Purwins of the UNI Munster [15]. (Courtesy of H. Purwins.)

allow realizing compact and economic systems. Here, a characteristic feature of plasmas comes to rescue, which emerges from its nature as open system, operated far from thermodynamic equilibrium viz. to self-organize. We demonstrate self-organization in space first by a most simple plasma arrangement – a diode, which is operated with an ac voltage. In order to prevent the formation of an arc, the circular electrodes are protected by dielectric layers. This arrangement allows quasi-two-dimensional dielectric-barrier discharges. The ac-current density does not develop homogeneously across the electrodes rather organizes itself in patterns, which depend on the diode current as control parameter. Figure 23.6 shows examples of such patterns [15]. This phenomenon is not unique to plasmas, rather a universal phenomenon described by reaction–diffusion theory in the frame of nonequilibrium thermodynamics [16].

An interesting spontaneous transition into a state with better confinement is observed in RFPs. Under normal conditions, strong multihelicity magnetic turbulence governs transport and confinement. The turbulence level is large enough to ergodize the magnetic field within the confinement zone. If the system is driven hard, for example, by increasing the plasma current, it suddenly transits into a regime with a lower turbulence level. Spontaneously, in the core, the ergodized field is replaced by nested flux surfaces adopting a quasi-single-helicity (QSH) equilibrium. Figure 23.7a exemplifies such a case. The temperature in this zone shows a distinct increase (see Figure 23.7b). The plasma self-heals a defect improving confinement thereby. Figure 23.7c shows this configurational change with an island in the poloidal cross section containing nested flux surfaces giving rise to the benefit of a locally measurable increase in temperature [17].

Also in tokamaks and stellarators, bifurcations into regimes with higher confinement are observed, which are of programmatic relevance and allow realizing the fusion conditions under simpler and more realistic conditions. Figure 23.8 shows the transition from the L-regime, the low-confinement regime, to the H-mode, its high-confinement counterpart [18]. The example is taken from ASDEX Upgrade tokamak. Also in this case, the transition is a threshold process, which

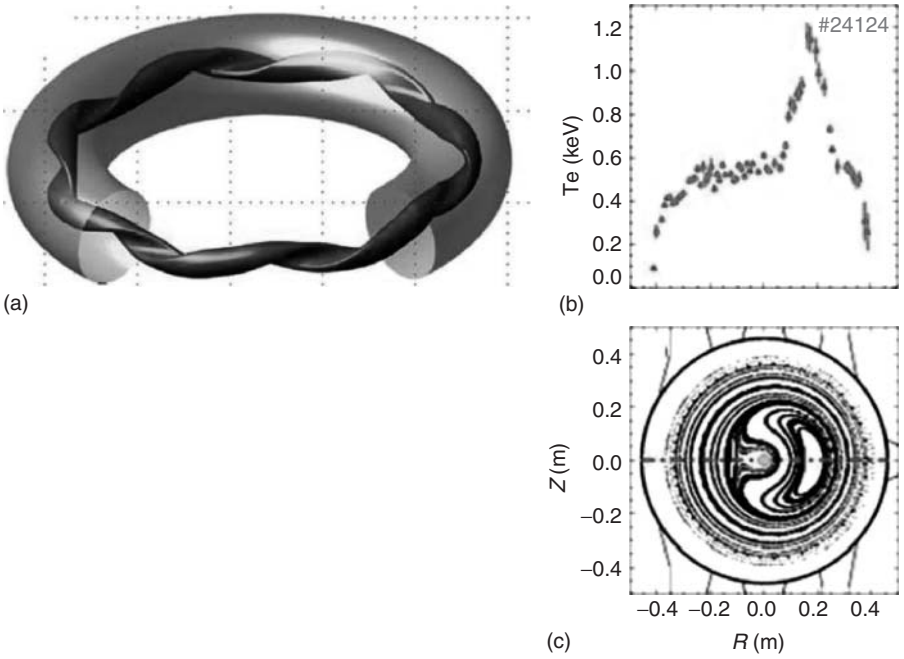


Figure 23.7 (a) The development of a single-helicity state in the plasma core of an RFP. (b) Electron temperature with a peak in the radial range of the helical equilibrium. (c) Poloidal cross section with nested flux surfaces and chaotic field lines closer to the edge. (Courtesy of Consortia RFX.)

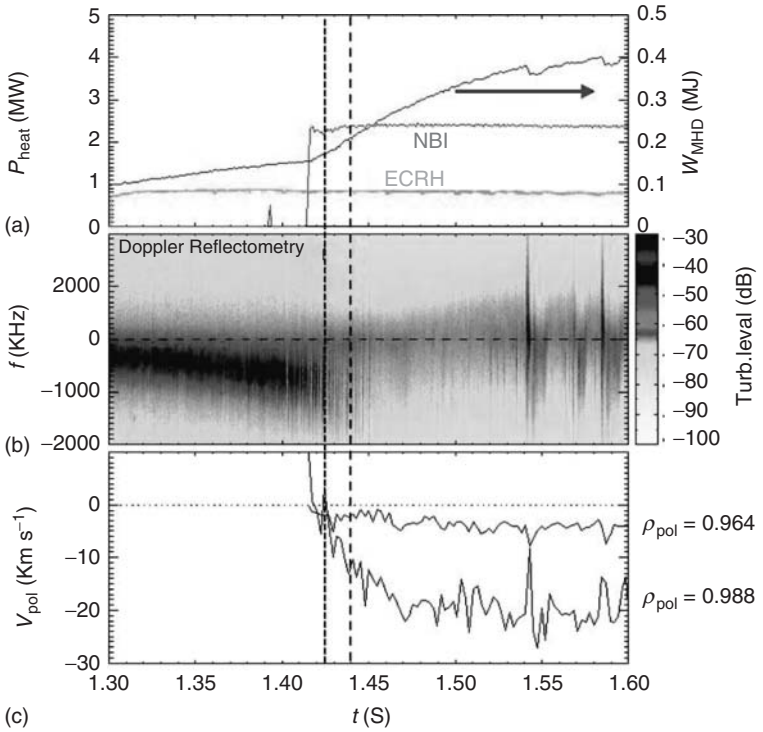


Figure 23.8 Plotted are different plasma traces from ASDEX Upgrade tokamak in a time window from 1.3 to 1.6 s during a discharge with auxiliary heating (NBI, *neutral beam injection heating*; ECRH, electron cyclotron resonance heating) with an H-transition (vertical lines). At about 1.425 s, the discharge transits into a period with limit-cycle oscillations; the final H-transition occurs at 1.44 s (dashed line). (a) The energy content together with the power of the two heating methods, (b) frequency and amplitude (dark grey, logarithmic scale) of the turbulence, and (c) the poloidal plasma rotation for two normalized radii ($\rho = r/a$). Poloidal rotation increases strongly at the plasma edge, about 1 cm inside the separatrix. (Courtesy of the ASDEX Upgrade team.)

requires a stronger drive developing after surpassing a heating power threshold. The transition occurs as soon as neutral beam injection (NBI) heating has been added (see Figure 23.8a). The two-dimensional turbulence at the plasma edge changes its dynamic behavior (see Figure 23.8b), whereas the spectral content condenses toward large scales limited only by the minor plasma radius. The interaction of turbulent eddies sum up to a coherent poloidal flow field (see Figure 23.8c). The sheared flow now acts back onto its driving agent and tears apart – decorrelates – the turbulent eddies. The turbulence is strongly reduced and steep gradients develop at the plasma edge forming – what is called – a *transport barrier*. When the turbulence is suppressed, the steep gradients for their part are maintained by equilibrium conditions where the steep ion pressure gradient is balanced by a negative radial electric field, which itself is equivalent to a strong and sheared poloidal flow. The state of low turbulence is preserved. The turbulence

quench acts as a trigger mechanism to a state with low turbulence and improved confinement. The prospects and expectations of the ITER are based on the H-mode confinement qualities. The generation of flows in interaction with 2D turbulence is reminiscent of the mechanisms that give rise to the belts of the planets Jupiter and Saturn and that play a role in atmospheric and oceanic flows.

The potential of self-organization is largest in RFPs because both confining fields are generated in a self-consistent manner. The stellarator is the opposite, because both fields are produced from the outside under full control. There is an interesting parallelism between the development of the QSH, the quasi-single-helicity state of RFPs and the H-mode. Both systems have to be driven hard from the outside. The free energy has to be increased in one case by the current and in the other case by the plasma pressure. In both cases, turbulence is reduced in selected regions – in the core or at the edge, respectively – and also the coherent interaction of turbulence gives rise to plasma flows. A transient process leads to a new equilibrium situation where the essential electric field is maintained either by the helical plasma deformation or the increased ion pressure.

23.8 Status of Fusion Energy Development

Fusion research was started worldwide after World War II. The EU fusion program is jointly coordinated by the Euratom contract, which provided the legal basis and the means to build the presently largest and most relevant fusion device, the Joint European Torus, JET, a tokamak, operated close to Oxford, UK. This device is closest to the ignition conditions. In the last decade, the Asian fusion program has made impressive progress with modern tokamaks in Korea (KSTAR [19]), China (EAST [20]), and India (SST-1 [21]). Japan has a long tradition in fusion development and houses the largest helical system at present, LHD, a heliotron [9]. Japan is now constructing a large tokamak, JT-60 SA [22]. All Asian tokamaks are realized with superconducting coils and address the critical and not yet solved problem of steady-state operation, specifically critical to tokamaks because of their intrinsically pulsed nature. Helical systems are operated in Australia, Germany, Japan, Spain, Ukraine, and the United States. RFPs are operated at many places at the scale of university experiments. The largest devices operate in Italy (RFX [23]) and the United States (MST [24]).

In separate experiments, fusion has already achieved the necessary temperature and density. Temperatures of 40 keV can be produced in a routine manner basically by pressing a knob. Densities have been achieved by a factor of ~10 more than the required one. The critical parameter is τ_E, where a factor of 4 is missing for reaching ignition. It is a major effort of the worldwide fusion program to improve confinement quality beyond the one achieved with the H-mode. Recently, it was shown that transport barriers, as observed in the H-mode at the edge, can also be produced additionally in the core region in the form of internal transport barriers [25]. Intense fusion research still pays off.

The US experiment TFTR (Tokamak Fusion Test Reactor) and JET have carried out DT experiments without any technical problems. Of course, the tritium consumption is in the range of grams and not kilograms as necessary in case of a reactor. α-particle heating has also been

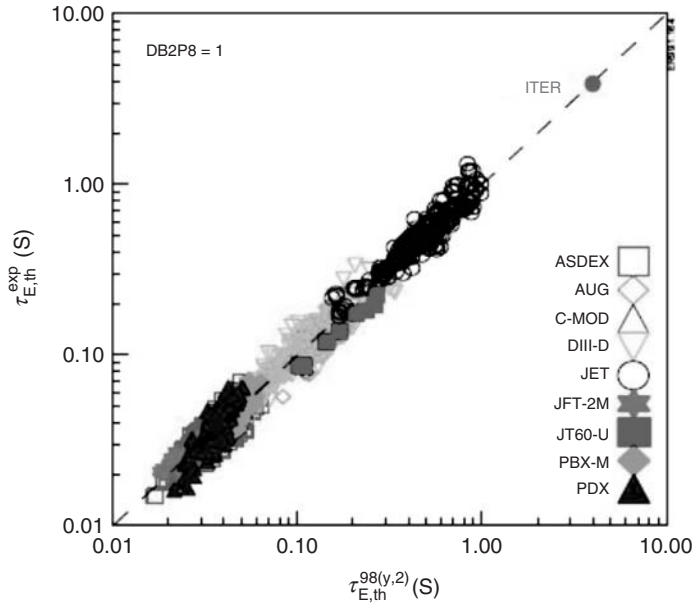


Figure 23.9 Multimachine thermal energy confinement time $\tau_{E,th}^{exp}$ against the scaling results of the 98(y,2) ITER scaling. The expected ITER confinement time is also shown.

demonstrated. Most importantly, JET has demonstrated in a short pulse a fusion power of 16.1 MW in DT experiments in the H-mode equivalent to a power amplification factor $Q = P_{fus}/P_{ext} \sim 0.65$.

The fusion community was able to support unanimously one concept for the first fusion test reactor, ITER. This device is based on the accumulated worldwide knowledge with a strongly shaped tokamak plasma incorporating a divertor for exhaust. The first-wall materials are in agreement with the best knowledge available at present.

The international efforts of fusion research are best demonstrated by the joint database, which provides the scaling of the energy confinement time τ_E defining the operation point of the ITER. As τ_E is determined by turbulent processes, a precise prediction of the ITER τ_E value is still not possible, although the basic mechanisms are understood and the equations and the formal structure for turbulence modeling have been developed. Its multiparameter dependence has to be identified by many experimental results subject to a rigorous statistical regression analysis. The outcome of such a study is shown in Figure 23.9. An experimental τ_E value depends on many parameters: the plasma current I_p, toroidal field B_{tor}, heating power P, density n, isotopic mass M, and the key geometry parameters, major radius R, elongation ε, and triangularity of the plasma cross section κ.

$$\begin{aligned}\tau_E^{ITER}(s) &= 0.0365 I_p^{0.97}(MA) B^{0.08}(T) \\ &\times P^{-0.63}(MW) n^{0.41}(10^{20} m^{-3}) \\ &\times M^{0.20} R^{1.93}(m) \varepsilon^{0.23} \kappa^{0.67}\end{aligned}$$

The expected τ_E value of the ITER is also shown (3.7 s). It is in agreement with the

objectives of the ITER and defines several crucial parameters of its design.

Stellarators can operate at higher density than tokamaks, which is an advantage for high power divertor operation. They show similar parameters as tokamaks of similar size and field. They do, however, not yet reach the high temperatures and confinement times of large tokamaks because of the smaller scales of the present-day helical systems.

23.9 The Next Development Steps

23.9.1 ITER

The ITER, shown as a computer drawing in Figure 23.10, will be the first fusion reactor producing fusion energy aiming at a power gain of $Q = 10$. The technical specification of the ITER is given in Box 23.1. The ITER has major physics issues to explore and to clarify, which necessitate a burning plasma state. The most important ones are (i) the confinement characteristics with dominant self-heating, (ii) burn control and helium exhaust, and (iii) plasma stability in the presence of a fast ion component, the α-particles. Crucial are also the properties of helium transport from the core to the edge and the safe removal of He within the divertor. Only a helium concentration of 2% is tolerable in the reactive core (see Figure 23.2). The He fluxes have a convective and a diffusive component. It is expected, but not ensured, that the convective part will move He ions to the outside. The α-particle pressure also contributes to the free energy with the potential to drive instabilities. In this case, the Alfvén waves could be induced by a velocity resonance within the spectrum of the slowing-down α-particles, which could act back onto the driving mechanism and reduce the confinement of energetic particles needed for plasma heating. Another crucial area of exploration with the ITER is the external current drive to prepare the conditions for steady-state tokamak operation. This effort combines suitable techniques with judiciously tailored plasmas, which maximize internal currents (bootstrap current) thus minimizing the need for external contributions.

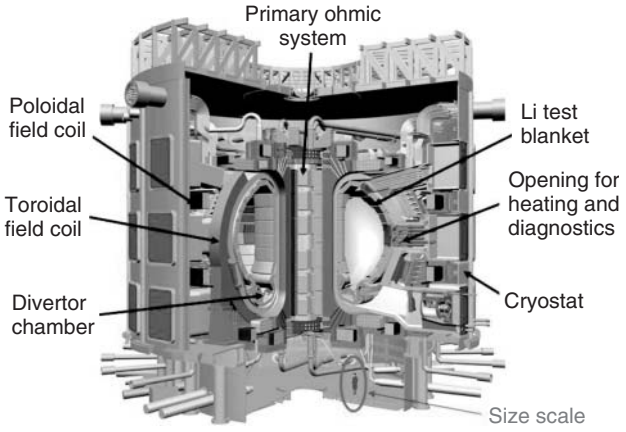


Figure 23.10 Computer drawing of the ITER.

Box 23.1: ITER

The ITER is based on the tokamak concept. It is the first fusion reactor to produce a large amount of fusion power (500 MW) on the basis of dominant self-heating by 100 MW of α-particles. The ITER will be a fully licensed nuclear device. The magnetic field on axis will be 5.3 T produced by superconducting coils using Nb_3Sn. The current of the ITER will be up to 15 MA. The geometry of the ITER is as follows: major radius $R_0 = 6.2$ m, minor radius $a = 2$ m, elongation $\kappa = 1.7$, and triangularity $\delta = 0.33$. The ITER will be equipped with a divertor. External heating for the ITER will employ neutral beam heating (33 MW) and both electron and ion-cyclotron heating (20 MW each).

The formal objectives of the ITER are a fusion amplification factor $Q \geq 10$ with an average neutron wall load of $\geq 0.5\ \mathrm{MW\,m^{-2}}$. The pulse length is limited to 500 s. In another scenario, quasi-steady-state will be demonstrated, but under the reduced requirements of $Q \sim 5$. The ITER will start operation around 2020.

There is a large field of technological contribution expected from the ITER. As only real projects solve practical problems and provide the necessary technologies, the ITER plays a unique role in fusion technology also. The ITER will demonstrate the use of superconductivity in the harsh environment of fusion and the viability of remote handling techniques for the complex assembly and maintenance tasks inside the plasma vessel, and it will develop ambitious external heating and current drive concepts such as the injection of high-energy hydrogen atoms at 1 MeV energy and in the 100 MW power range. Critical issues are the exhausted α-power handling at a level of $15\ \mathrm{MW\,m^{-2}}$ and the demonstration of tritium production using test blankets in a form that allows clear modeling predictions for those used in DEMO and a later power reactor.

23.9.2 Wendelstein 7-X and the HELIAS Reactor

As we have seen, stellarators need confinement optimization. The most rigorous form led to the concept, which underlies the Wendelstein 7-X (W7-X) project to be realized in Greifswald, Germany, which is presently (2012) in the assembly phase (see Figure 23.11). The design parameters of W7-X are given in Box 23.2. The theoretical background of the optimization is that (i) the deviation of particles from the flux surface as well as equilibrium and

Box 23.2: Wendelstein 7-X

W7-X is a multihelicity stellarator with optimized properties following the quasi-isodynamicity concept. Its 50 modular coils are fabricated from NbTi superconducting materials with 3 T maximal field on axis. Heating and cooling of the system will allow 30 min pulses. Major radius of the device is 5.5 m. The aspect ratio is 10. The major heating system is a steady-state electron cyclotron system of nominal 10 MW power at 140 GHz. Neutral injection and ion-cyclotron heating are in preparation. For exhaust, W7-X is equipped with an island divertor. The device is presently assembled in the IPP branch Greifswald; start of operation is scheduled for 2014.

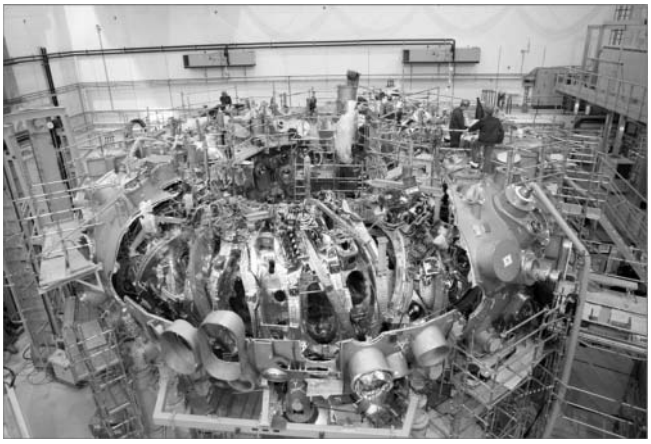


Figure 23.11 A view into the experimental hall of Wendelstein 7-X (as per December 2011).

stability properties depend on the variation of the magnetic field strength $|\mathbf{B}|$ in the flux surface [26] and (ii) $|\mathbf{B}|$ can be made two-dimensional in an otherwise three-dimensional toroidal geometry [27] as it is a given fact for external confinement. Systems with these properties are called *quasisymmetric systems*; meanwhile, several smaller stellarators are built and operated based on these concepts (HSX in Wisconsin [28] and Heliotron J in Kyoto [29]). The design of W7-X follows the quasi-isodynamic principle. In an isodynamic system [30], all particles, heat, and charges would ideally flow in a flux surface without radial component. Such a system cannot be realized in the frame of toroidal confinement. Quasi-isodynamicity, a viable approximation, foresees rigid equilibrium and stability toward high plasma pressure and good confinement for thermal and energetic particles (α-particles) avoiding the helical ripple losses as described in Section 23.6. One optimization feature is the utilization of the mirror effect. The toroidal periodicity of W7-X corresponds to a pentagon. Straight sectors are connected by corners. The toroidal curvature is therefore localized to the corners. In order to minimize the drift losses in the zones of field inhomogeneity, the field strength $|\mathbf{B}|$ is increased in the corners. W7-X resembles a system of linked mirrors. The most critical particles to drift off the system, those with a large $v_\perp/v$, are expelled from the sectors with curvature. Complete and rigorous optimization is still not verified. However, some of the basic physical qualities have already been tested in the smaller and partially optimized stellarator device W7-AS [8]. These favorable properties, if they become reality, are on top of the intrinsic stellarator characteristics – steady-state operation free of current-driven instabilities. W7-X, being a superconducting device, is supposed to demonstrate these features for pulses up to 30 min.

Stellarator reactor studies have the main emphasis on stellarator-specific aspects such as the coil system and the complex 3D geometry of blanket and neutron shield. They are basically extrapolations of the present concepts; in case of W7-X, it is the HELIAS (Helical Axis Advanced Stellarator) reactor concept [31]. The design is dictated by two criteria: the size must be large enough to accommodate the blanket and shield between plasma

and coils and the confinement time, which roughly grows with the plasma volume, must be large enough to ensure ignition and self-sustained burn. Not explored in much detail yet are the necessary properties subsumed under the term *RAMI* – reliability, availability, maintainability, and inspectability. The 3D nature of stellarators will complicate the technical realization of the RAMI criteria. Otherwise, stellarators have to demonstrate the same qualities as foreseen for the ITER and described in the previous section. Of course, stellarators will also benefit tremendously from the additional physics and the technology provided by the ITER. An intermediate step might be necessary in the stellarator reactor evolution, such as a ''stellarator ITER,'' but this needs to be explored.

A future stellarator reactor will not need a current drive system. About 200 MW current drive power is required for a steady-state tokamak reactor. The plant electricity consumption and the recycling power are high for a typical electrical efficiency of 0.4. This could increase the cost of electricity by 25% [32].

23.9.3 DEMO

The DEMO will be an electricity-producing fusion reactor prototype, which will deliver about 1 GW electric power in the steady-state or quasi-steady-state. It will be slightly larger than the ITER. Specifically, it will operate at higher density (P_{fus}, being a binary process, scales with n^2) and will require somewhat higher confinement and stability margins. These topics are on the agenda of the ITER research program. The ITER will also play an important role for DEMO in the field of licensing. This is a totally new area, which will be exercised for the first time with the nuclear safety authorities of France.

The DEMO could be a tokamak or a stellarator or both. One or two decades from now, the global energy demand may become so critical that more means will be provided than that is presently the case for energy research, and hence, fusion R&D could continue on a much broader basis.

23.10 Technical Issues

23.10.1 Tritium Production

The critical quality of the breeding blanket is the tritium breeding ratio (TBR). For each triton burnt in the fusion process, a new one has to be formed. Because of unavoidable losses (decay in storage, initial feeding for other fusion reactors, spurious tritium inventories in the exhaust and recycling system), TBR has to be >1 to achieve supply self-sufficiency. This quality depends on plasma physics properties such as the *burn-up fraction* (fusion reaction rate/tritium fueling rate) and design aspects of the blanket, the nuclear processes inside the materials used, and the way a 14 MeV neutron slows down causing interactions with lithium governed by the cross sections in Figure 23.1 till it is finally absorbed or lost. The relevance of process (23.5) is the high cross section down to the lowest neutron energies. In the ^{6}Li reaction, each n has the chance to generate a tritium nucleus. In addition, this process contributes nearly 30% to the fusion energy yield. The importance of process (23.6) is the n-multiplication, which is possible for neutrons with an energy above the

threshold of 2.5 MeV (see cross section in Figure 23.1).

A fusion reactor needs an initial supply of tritium from an external source to start operation till it reaches tritium-self-sustainment. The external sources are heavy water fission reactors (e.g., the CANDU (CANada Deuterium Uranium) reactor), which use heavy water for moderation and cooling as they are operated in Canada, India, and Korea. The worldwide amount of tritium is determined by the production and mostly the radioactive decay and may eventually rise to 25 kg in 2036 when Canada plans to stop the last of the present CANDU reactors under operation. A fusion power plant will initially require 7 kg tritium to start. The exact value depends on the tritium burn-up fraction, which, however, cannot be predicted with confidence. This shows us that the tritium availability gives rise to rather critical supply conditions for the later market penetration of fusion power plants in and beyond 2050. At an optimistic breeding ratio of 1.15, a 3 GW_{th} power station produces 25 kg $year^{-1}$ surplus tritium to start other fusion reactors. But, also in the future, tritium can and will be produced by other means. Canada, the main tritium supplier at present, plans to build advanced versions of the CANDU reactor type. Thermal reactors, including those of the Generation-IV family, can breed tritium via the process (23.5).

In any case, it is urgently necessary that the actual TBR values are confirmed under realistic conditions. Fusion has not yet tested a breeding blanket because a serious test requires a burning plasma supplying 14 MeV neutrons. Pure lithium would have the highest TBR close to 2 but is not considered because of the safety aspects of this alkali metal and its unfavorable flow conditions as a conducting liquid in the magnetic environment of a fusion device. For the ITER, two breeding blankets will be tested *in situ*: one based on liquid PbLi eutectic and the other based on solid Li ceramic with a lower TBR. Both will allow testing the TBR under realistic conditions but the foreseen test blankets will not allow the ITER to produce the tritium it actually needs. It has to be added from external sources.

At present, the expected TBR parameters are based on mock-up experiments and numerical calculations of the n-transport. The calculations and their benchmark are of utmost importance because blanket designs optimized for a high TBR have to be done by calculations. Figure 23.12 shows the volume T-production in a LiAlPb blanket mock-up as a function of the penetration depth of initially 14 MeV neutrons from an accelerator [33]. The overall rise of the curve depicts the increasing role of ^{6}Li when neutrons slow down. There is good agreement with calculations. Some discrepancies come about by the lack of detailed knowledge in the cross sections of relevant nuclear processes. In order to increase the tritium yield, neutron multipliers can be used, which are based on the (n,2n) reactions with thresholds lower than 14 MeV. Candidates are beryllium ($^{9}Be + n \rightarrow 2\ ^{4}He + 2n$) and lead ($^{208}Pb + n \rightarrow\ ^{207}Pb + 2n$), which is already part of the LiPb blanket eutectic. Pb has the higher cross section (see Figure 23.1). The idea to enrich natural Li with ^{6}Li because of its high cross section specifically for thermal n-energies has to be traded off against the virtue of process (23.2) to increase the number of neutrons. For a fusion blanket meeting all the three tasks, a realistic value of the TBR is about 1.1.

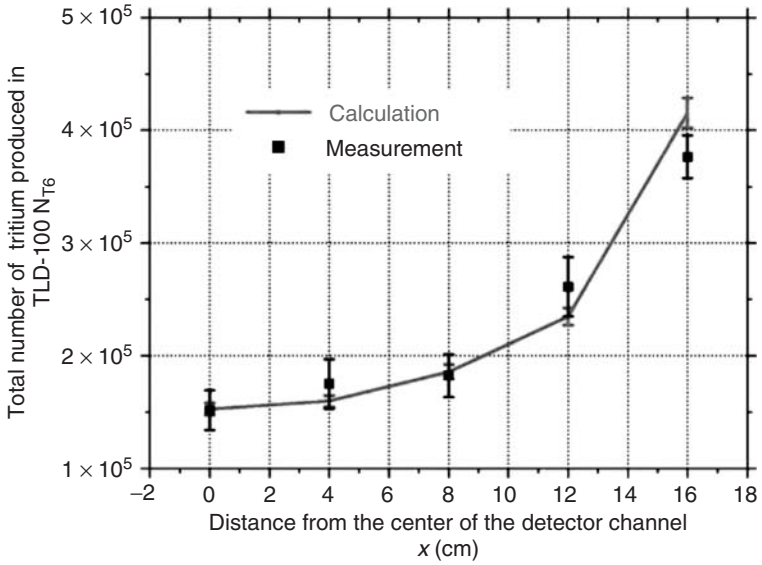


Figure 23.12 Tritium production in the LiAlPb mock-up experiment for the source fluence of 2.15×10^{13} neutrons measured with LiF-TLD (thermoluminescent dosimeter), which were calibrated with a Li glass scintillator, and tritium production calculated with MCNPX (Monte Carlo N-Particle eXtended) for the same neutron source fluence. The *x*-axis shows the distance from the center of the detector channel to one end of the channel [33]. (Courtesy of H. Freiesleben and Elsevier publisher.)

23.10.2 Fusion Materials

A fusion reactor is a technically challenging device. One has to build a system, which allows 150 million K in the core and 2 K about 2 m away. The maximum magnetic field of the system is ~12 T at 1.8 K using Ni_3Sn superconducting filaments. The magnetic system energy is in the range of 50 GJ. High neutron fluxes have to be handled in the range of 2–3 MW m^{-2}, resulting in degradation of the mechanical and thermomechanical properties such as ductility, tensile strength, fracture strength, and creep strength. The tolerable neutron fluxes are defined in terms of displacements per atom (dpa). The situation of the ITER is uncritical with 1–3 dpa. The DEMO and reactor materials have to stand as much as 100–150 dpa.[8] The corollary is that parts of the blanket have to be replaced after 3–5 years of operation. In addition, (n,α) reactions produce helium and (n,p) reactions produce hydrogen inside the material, also leading to equally critical material fatigue. As a consequence of material degradation, the ductile-to-brittle transition temperature (DBTT) moves toward lower values thus reducing the thermodynamic efficiency of the system.

A crucial field is the development of low-activation material. The usual minorities of stainless steel responsible for selected material properties have to be reduced if they are transmuted into radioactive isotopes. These minorities are predominantly Co, Ni, Nb, and Mo. Low-activation

8) The production of 0.5 TW h electricity in a fusion reactor gives rise to about 1 dpa.

materials based on ferritic martensitic steels are predominantly developed in Europe and Japan. In case of Eurofer [34], the less critical minorities are Cr, W, V, and Ta. Structural materials based on vanadium alloys or SiC ceramics would be even superior from the waste point of view but they require further development. Such a material selection program pays off for fusion because the structural materials and not the reaction products, which do not leave any choices, determine the radioactive inventory.

The ITER can be built by standard technologies and materials, for example, austenitic steels. Material challenges beyond the ITER necessitate a material R&D program. For this purpose, a 14 MeV neutron source is required to assess all material consequences of the high transmutation rate of 14 MeV neutrons. This can be realized in the form of a spallation source with a deuterium accelerator and a tritium target or by a neutron source based on thermonuclear reactions. IFMIF, the International Fusion Materials Irradiation Facility [35], is presently designed as 14 MeV n-spallation source in the frame of the IAEA by Europe, Japan, Russia, and the United States. It could also be a plasma neutron source based on the spherical tokamak (ST) [36] or on a specific version of a mirror machine [37].

23.10.3 The Availability of Fusion Fuels

A total of 3.55×10^{20} fusion processes per second yield a fusion power of $1\,GW_{th}$. Expressed in atomic mass units (u) for easy scaling purposes, this corresponds to $18.7\,kg\,year^{-1}$. The annual fuel demand is therefore 37.4 kg of deuterium and 56 kg of tritium. Assuming that the majority of T-production occurs via ^6Li, which has a cross section more than three orders of magnitude larger than that of ^7Li at low energies (see Figure 23.1), the annual ^6Li consumption is 112 kg, which corresponds to $1.5\,ton\,year^{-1}$ of natural Li. For a TBR of 1.1, we would, on the basis of these simple estimates, need about $20\,kg\,year^{-1}$ of beryllium (9 u) or $400\,kg\,year^{-1}$ of lead (207 u), respectively. The actual consumption is larger and depends on details of the blanket design and plasma operation.

We use the raw material data compiled in [38] to roughly assess the energy obtained from fusion, considering the availability of the materials the fusion process consumes. Seawater contains about 2.3×10^{13} ton of deuterium. This is adequate forever, and we do not have to discuss this fuel branch any further. The reserves (resources) of lithium are about 7 (22) Mton. In sum, they would deliver about 6×10^7 TWh of energy.[9] This compares with 20 000 TWh of the present world electricity demand. The water on Earth would supply a factor of 10^4 more Li than is known on the Earth. The Li fuel is therefore not a problem of availability rather of workability. For 1 GW of electricity, one has to extract lithium from about 300 l of water per second. For the electricity requirement in Germany, this implies the processing of the flow of a small river. To compare with, the river Rhine flows at about $2000\,m^3\,s^{-1}$. A corresponding technology to extract lithium from water might indeed be reasonable.[10]

About 80 000 ton of beryllium is expected to be mined from Earth. It can provide about 1.5×10^7 TWh of energy and would

9) We assume an electricity conversion efficiency of 0.35.

10) The energy balance is positive even if the evaporation of water were necessary for the extraction process.

limit fusion energy to about 700 years for the present demand. But Be can be substituted by lead. The lead reserves are about 79 Mton and the resources measure up to 1.5 Gton, providing $\sim 10^{10}$ TWh electricity.

Structural materials and those for the superconducting coils are not considered here. Also helium, which might become sparse if the careless use of it continues, is not considered. The actual materials used to build a fusion reactor will depend on the overall technical progress, for example, on the possible application of high-temperature superconductors.

In summary, lithium is the limiting raw material in fusion. As Li is presently also the material of choice for batteries, the removal of ^{6}Li may have to be considered. ^{7}Li can be used for batteries without performance reduction. The mining of Li from seawater might be feasible. In this case, no practical limitation for the use of fusion energy is expected from the fuel point of view.

23.11 Safety Issues

23.11.1 Operational Safety

In Section 23.5, we have summarized the intrinsic features of fusion, which can be exploited to realize a safe and justifiable energy source. The safety aspects of a fusion reactor are based, in addition, on reactor models and accident scenarios. The ITER goes presently through a licensing procedure with a detailed safety analysis and will provide practical evidence for many safety issues.

The basic safety features of fusion energy are that the reactor containment cannot be destroyed by an internal accident. The hazards of fire are reduced by using lithium in the blanket in ceramic form. Graphite will not be used inside a reactor to avoid large tritium inventories. In case of a loss of coolant accident (LOCA), only the afterheat in the structural materials has to be considered. Its power density is about a factor of 100 lower than that in case of fission. The temperature increase after an LOCA does not reach melting temperatures and it develops slowly reaching a maximum after about a week. The major safety concern is the release of tritium. Tritium decays with a half-life of 12.3 years. It is a β-emitter with, as a consequence, low radiotoxicity. Tritium is highly volatile and quickly diluted in nature, in the atmosphere and the soil. The human body purifies itself both from gaseous tritium compounds and from HTO. The typical timescale is 10 days. The tritium that is embedded in bones is minute but more firmly bound and has to be considered separately; for example, it can be scaled to the radioactivity of natural ^{40}K, that is, $>100\ \mathrm{Bq\,kg^{-1}}$, in human bone tissue. These conditions are fundamental, which ease the consequences of a maximal accident, for example, after an earthquake. In such an event, up to 1 kg of tritium could be released. Under the worst weather conditions, a comparatively small area of approximately square kilometers has to be evacuated for a short period. The collateral damage due to the presence of a fusion power plant will be much less than the primary damage caused by a hypothetical catastrophe.

23.11.2 Fusion Waste

Fusion energy produces waste the amount of which is comparable to that of fission.

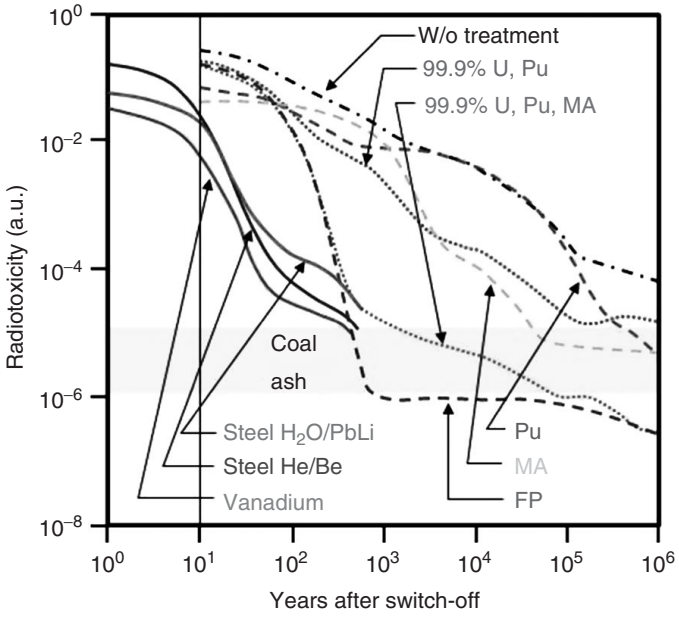


Figure 23.13 Decay of radiotoxicity of the waste from fission and fusion reactors. Shown are the cases of a light-water reactor without any waste treatment, with the removal of Pu, minor actinides (MAs), and fission products (FPs). These traces and timescales are compared with those of fusion with various structural materials and coolants. (The plot is partially based on data made available by the KIT.)

The radiotoxicity of the waste is, however, totally different because the fission products, plutonium and MAs, with long half-lives are absent in fusion (see Figure 23.13). As described earlier, low-activation structural materials are developed for fusion. Most realistic at present is the expectation that the activated waste produced by fusion will decay typically within 100 years and can again be used thereafter. If longer-term waste is created, it is a small amount for which shallow geological deposition is sufficient. It is realistic to expect that this fraction can be avoided by the material development as discussed in Section 23.10.2.

23.11.3 Proliferation Issues

Fusion involves tritium and the 14 MeV neutrons allow producing fissile materials ^{239}Pu and ^{233}U out of ^{238}U and ^{232}Th, respectively [39]. Typically within a month, several kilograms of fissile material can be produced, which comes close to the amount needed for an explosive weapon. Tritium itself is used in atomic bombs, boosting ignition via fast neutrons. Hydrogen bombs can be built without tritium by using lithium-6 deuteride as the explosive. Tritium is available and is used in medicine and industry, and it is not an IAEA-controlled commodity. Tritium

proliferation is not an issue for fusion technology.

The most recent analysis of the proliferation hazards of fusion being used for breeding fissile material is given in Ref. [40]. We follow the main arguments provided in that source:

- Fusion power plants are large and can hardly be hidden.
- Nuclear materials are not needed in fusion developed as an energy source. Therefore, they can be easily detected in case of an internationally monitored power plant.
- Fusion technology is complex. A simpler technique is to utilize fission; in this case, illegal production of fissile material is less probably noticed.
- A fusion breeder needs a large and vulnerable infrastructure. Clandestine fission material production can be stopped by destroying the infrastructure without any nuclear risks – if such a measure should indeed become unavoidable.

23.11.4 Fusion as Neutron Source

So far, we have discussed fusion as an energy source. But it can also be considered as a powerful neutron source. While a thermal fission reactor delivers $\sim 5 \times 10^{-3}$ neutrons/MeV for additional use, fusion with n-multiplication produces ~10 times more fast neutrons per megaelectronvolt. A spallation neutron source is, however, not able to deliver this quality. Such an n-source might be required for the development of fusion energy, for example, to test and optimize in-vessel components such as blanket or heating or diagnostic hardware, to produce and handle tritium and minimize its inventory and its permeation to the outside, or for material testing. Nonfusion applications could be in other research fields, where neutrons are used, or in the production of isotopes for medical purposes. Fusion neutrons also allow transmuting fission waste, specifically the long-lived MAs.

The requirements for a fusion n-source – for whatever purpose – are distinctively lower than those of pure fusion. The physics specification of a fusion n-source is less challenging than that of the ITER. Its fusion power is in the range of 100 MW with $Q \sim 1$; its plasma parameters are close to those achieved in present-day experiments. Also technical and material requirements are alleviated because of the lower neutron flux densities.

A fusion n-source demands a compact build to be economic [36]. Within the tokamak family, a separate branch has been developed in the recent years, the ST. An ST is characterized by a low aspect ratio $A \sim 1.5$. It can meet plasma stability requirements for high I_p/B_t values, and it allows high plasma pressure (and power density), $p \sim I_p/aB_t$, for typical operational conditions. The compact geometry and the profile of the released neutrons also ensure good coupling between them in a surrounding blanket containing lithium for tritium production as well as for the fission waste to be processed.

The fast neutrons of a fusion n-source are able to transmute the radioactive waste of fission into stable or short-lived isotopes integrally shortening its decay time and even allowing to change the whole waste processing and disposal strategies (e.g., by realizing waste disposals with access). A general strategy could be to let the fission products with low n-absorption cross sections decay naturally and process MAs by a fusion n-source. One obvious

advantage of utilizing the fusion process is that no additional waste is produced within the transmutation process and the neutrons for transmutation are produced independently without requiring the waste itself as n-source. This is a difference between fusion and fission.

A 1 GW_e light-water reactorLWR) produces per operational year typically 25 kg of MAs. Because of the high fission cross sections of MAs in the fast-neutron energy range, a fusion neutron source, as specified earlier, could transmute these forms of waste for five LWRs. The benefits of the decay of radiotoxicity of partitioning and transmutation of fission radioactive waste are shown in Figure 23.13. The figure stresses, however, once more the benefits of clean fusion energy from the waste point of view even after optimal posttreatment has been granted to fission waste.

23.12 Outlook

The scientific efforts to develop controlled nuclear fusion started in the 1950s. The original hope might have been to develop fusion energy as soon as fission was realized after the military use had demonstrated the vast energies released thereby. Although Fermi demonstrated the first self-sustained and controlled fission chain reaction in December 1942 with the Chicago Pile-1 at a ~1 W power level and at the safety precautions provided by an (empty) college football stadium, we now know that fusion needs an appropriate confinement, expressed in confinement time τ_E, which can only be realized at large scales. The consequence is that the first test reactor, the ITER, is a large and expensive industrial product, which right away produces 500 MW of fusion power. Such a device needs full nuclear licensing and corresponding safety features. The scientific and technical demonstration of fusion can only happen in a device, which is close to a commercial reactor.

The feature of fusion unknown in the pioneering period was the high level of plasma turbulence, which transports energy at a high rate from the core to the edge thereby limiting confinement. As a consequence, the fusion development strategy necessitates a chain of devices increasing in size and closeness to the final reactor parameters. Therefore, there are several decades between formulating a decision to realize controlled nuclear fusion and the final outcome of such a decision. In the 1970s, after the first oil crisis, the US fusion community predicted 40 years as developing time for fusion, following a request from its government. The 40 years have passed by and the fusion reactor is not yet ready. This delay is in people's memories, and one of the jokes on fusion is that it takes always 40 years irrespective of the date the question on its realization is posed. This report presents hopefully enough evidence to contradict this view. What is not considered at all is the financial side of the fusion strategy. Of course, the 40 years development prediction had a sound financial basis. But as soon as the oil crisis was overcome in the United States and worldwide, the budgets for energy research decreased. Fusion did not get much more than about 10% of what had been specified and demanded. Without support, no delivery on time!

The situation now is different from that in the 1970s. Human population has grown from 4 to 7 billion and will grow further. The oil prices strongly fluctuate with a continuous increase in the average prices. An additional aspect is environment and global warming as the corollary of

anthropogenic CO_2 release. The need to develop powerful clean energy sources has strongly increased and has become a global need. The development of inherently safe fission power plants of the Generation-IV family will take decades. The same time frame governs the development of proper storage for electricity of fluctuating renewable sources if storage is based on chemical processes and not pumped water storage, whose scaling to larger capacities is subject to local conditions. It is meanwhile also clear that bio-energy has only a very limited capacity in the range of 5% of the global primary energy consumption.

In the next 40 years, the oil peak will happen with unimaginable consequences on the economic activities of a globalized world. Facing these perils, energy research has to be tremendously intensified. Only the results of energy research may be the insurance that 9 billion people on Earth can live in peace. Facing the few options mankind has and the doubts on their potential, acceptability, and sustainability, none of the three possibilities may be discarded or delayed.

The ITER is scheduled to start operation in 2020. The initial experimental period will be used to develop plasma scenarios with good prospects to meet the overall ITER goals. The DT operation is scheduled for 2027. The step following ITER will be DEMO, the fusion demonstration reactor. It will be a complete reactor, operating steady-state or pulsed, under quasi-steady-state conditions, which should produce electricity and demonstrate also the economic competitiveness of fusion energy.

W7-X will start operation in 2014. The quality of the optimized stellarator concept will enter the discussion on the continuation of the fusion development when the ITER has started. In this period, different DEMO concepts can be pursued and elaborated; a concept decision can be taken later. It could easily be that both lines are rigorously pursued because the ITER has successfully met all its goals and with W7-X a convincing steady-state concept has been demonstrated. In case of success, at the beginning of the 1930s, the development of fusion power can be actively placed into national electricity development strategies. In this period, also a clearer picture will exist about the prospects of the Generation-IV fission reactors and the CO_2-free operation of renewable energy sources, thanks to the availability of chemical storage systems. With the availability of fusion energy, the process of implementation of this technology starts. This period will depend strongly on the availability of tritium to start new reactors. In any case, fusion will not supply exclusively electricity to mankind. Fusion will be an option where ensured base-load power is required and where safe operation of a large electricity source in critical areas is necessary, for example, within the boundaries of megacities.

The cost structure of fusion electricity is determined by large capital costs and largely ignorable fuel costs. Fusion reactors will contribute to the base load and require long full load hours for economic operation. They do not cause external costs because electricity production is CO_2-free and the fuel is obtained without any critical environmental damage. The electricity costs will be smaller than the usual state taxes and charges per kilowatt hour on electricity are, but they are expected to be higher than the cost of electricity from fossil fuels and present nuclear reactors. The details depend, however, on the development of CO_2 costs and those for additional fission reactor safety. It is also expected that fusion electricity will be cheaper than photovoltaic electricity.

Glossary

Alfvén waves: The slow hydrodynamic wave in a magnetized plasma where both magnetic field and density propagate together as either compressional or shear wave. The name goes back to Nobel Prize winner Hannes Alfvén, a plasma physicist and an astrophysicist.

Anomalous transport: A misnomer of the not-well-understood turbulent transport in plasmas.

Becquerel: The activity (number of decays per time unit) of a radioactive substance; $1\,\mathrm{Bq} = 1\,\mathrm{decay/s}$. The name goes back to Nobel Prize winner A. H. Becquerel.

Bifurcation: Here, a bifurcation is the transition of a nonlinear system into a regime with different dynamics. The bifurcation is caused by the change in a critical parameter.

Binding energy: The energy that has to be spent to split a unit into several parts. Here, it is the energy necessary to split a nucleus into two lower-mass fractions. In the reversed process, the binding energy is released when two nuclei fuse together.

Blanket: The fusion blanket surrounds the plasma. It has three tasks – to absorb neutrons and transfer the energy to cooling media, to breed tritium out of lithium, and to shield the superconducting coils from fusion neutrons.

Bootstrap current: In this chapter, we discuss inductive currents. In plasma with pressure gradients, the so-called pressure-driven currents appear. One of them is the self-generated toroidal current – therefore the name.

Burn-up fraction: It is the probability of a tritium atom fueled into the plasma to fuse with a deuteron before being transported to the plasma edge and lost.

CANDU reactor: A Canadian fission reactor that uses heavy water as the moderator. This reactor type produces tritium as by-product.

Coulomb collisions: Collisions between charged particles as described by the Rutherford cross section or - in integral form in a plasma – by the Coulomb cross section. Assemble properties such as collision times and mean-free paths are derived from them.

DEMO: DEMO is the fusion demonstration reactor, which is the prototype for commercial reactors. DEMO should demonstrate electricity production at market prices.

Diode: A diode here is a parallel-plate electrode arrangement. An ac or a dc voltage between the electrodes produces a plasma discharge with a heating current. The electrodes are shielded by a dielectric material to prevent the formation of an arc. This arrangement allows the so-called dielectric-barrier discharges.

Divertor: A divertor is a separate chamber inside the plasma vessel used to exhaust the plasma power and particle flows. It requires a specific magnetic setup where the plasma is limited by a separatrix and surrounded by a scrape-off layer with open field lines. The field lines are separated from the main plasma by an X-point and guided by specially shaped magnetic fields into the divertor chamber. Here, the field lines cross the target plates where the plasma is neutralized on incidence and the plasma power flow is transmitted to mechanical structures with cooling channels. Inside the divertor chamber pumps are foreseen for particle and He ash control.

dpa: The displacement per atom caused by impinging neutrons inside the blanket and the wall.

Eigenmode: Eigenmode is a normal mode of a system with fixed boundaries. In this frame, oscillatory movements are possible with characteristic frequencies. The eigenmodes of a toroidal geometry are characterized by the poloidal and the toroidal mode numbers m and n.

Energy confinement time: The time taken by the plasma to replace its energy content under steady-state conditions. It is a measure of the transport processes inside the plasma.

Eutecticum: A mixture of materials (e.g., metals) with a specifically low melting temperature, which can be applied for various purposes in the liquid state.

Flux surface: This is a term of *magnetic confinement*. Field lines map out a geometry, which embraces a constant magnetic flux. As a flux surface does not lose flux, the field lines are tangential to it. Toroidal systems comprise nested flux surfaces. The most inner one is a degenerated one, a closed loop – the magnetic axis.

Generation-IV reactor: Presently, Gen-II reactors are worldwide in operation. Gen-III or Gen-III+ reactors are under construction in various countries. Gen-IV reactors are designs that are not yet commercialized. Several concepts are pursued with different favorable properties.

Heliotron: The Japanese version of a torsatron (see torsatron).

H-mode: The H-mode (high-confinement mode) is an operational plasma regime with good confinement in contrast to the L-mode (low-confinement mode). Presently, the H-mode is the basis for all fusion reactor studies.

IFMIF: The acronym for International Fusion Materials Irradiation Facility, which is presently in the design phase. This 14 MeV neutron source is based on the spallation principle and should serve the research on fusion materials.

ITER: The acronym for International Thermonuclear Experimental Reactor – the first fusion reactor built in Cadarache, France, and based on the tokamak principle.

JET: The acronym for Joint European Torus, operated in Culham, close to Oxford, UK. JET is the world's largest tokamak.

Lepton: Leptons are the small elementary particles of the standard model with spin $1/2$.

Limiter: A limiter resides inside the plasma vessel and defines the plasma size. It is the first obstacle (either a diaphragm or a toroidal rail) the plasma comes into contact with. The limiter is exposed to high particle and power fluxes and represents a source for plasma impurities. In modern systems, it is replaced by a magnetic separatrix.

LWR: Light-water reactor.

Magnetic mirror: A magnetic mirror is a term from magnetic confinement. It denotes a section with increased magnetic field. A magnetic mirror has the property to repel (mirror) charged magnetized particles with low parallel velocity.

Minor actinides: These are transuranic elements, such as americium, curium, that are produced by fission. Plutonium is called a *major actinide*.

Mock-up: A mock-up is generally a scaled model for specific demonstration purposes.

Neoclassical transport: The collisional transport including the geometrical effects

on particle orbits of the inhomogeneous toroidal magnetic field.

Particle confinement time: Equivalent to the energy confinement time but for particles.

Phase space: In regular space, a point defines a position, whereas in six-dimensional phase space, it denotes three spatial and three momentum (or velocity) coordinates. Here, it is defined by the velocities parallel and perpendicular to the magnetic field.

Plasma: A gas at high temperatures where the atoms are ionized. It consists of ions and free electrons and is neutral at the macroscopic scale. Plasmas are highly conductive. More than 95% of matter is in this state.

Poloidal direction: The poloidal direction is orthogonal to the toroidal direction and denotes the short way around the torus.

Radiotoxicity: Radiotoxicity characterizes the hazards emerging from radioactive waste with decaying elements. The radiotoxicity hazard of a radioactive element is classified considering its activity, the forms of radiation, its energy, and the forms of ingestion.

Reaction rate: It denotes the speed of a reactive process. It is determined by the density of the reaction species and the reaction rate coefficient constructed from the interaction cross section, the particle relative velocity, and the velocity distribution function.

Reversed field pinch: The RFP is a toroidal confinement concept with plasma current with a strong poloidal component produced by magnetic turbulence. The name comes from the equilibrium property that the toroidal field reverses direction at the edge.

Self-organization: Self-organization is a feature of strongly nonlinear, nonequilibrium systems as they occur in physics, chemistry, and biology. Such systems develop transitions from less to more ordered organizational forms. Often, such transitions are initiated by interacting fluctuations leading to a coherent dynamic response, which affects the driving agent.

Separatrix: See divertor.

Spallation source: A neutron source using an accelerator to accelerate hydrogen isotopes incident onto a target to knock out neutrons. IFMIF is a spallation source using accelerated deuterons impinging on a lithium target.

Stellarator: A toroidal confinement system with external coils to produce a confining field system. The classical stellarator combined toroidal and helical coils. Modern stellarators consist of nonplanar modular coils.

Tokamak: Toroidal confinement system with a strong toroidal field and a strong plasma current to produce the poloidal field.

Torsatron: A helical confinement system with helical coils only, which produce toroidal and poloidal field components.

Tritium breeding ratio: TBR describes the efficiency of the blanket to produce tritium out of lithium from the impinging neutrons. It is defined as a ratio of the tritium produced and the one lost by fusion processes in the core of the reactor.

Wendelstein 7-X: W7-X is an optimized stellarator with superconducting modular coils, which is under assembly in Greifswald, Germany.

X-point: See divertor.

References

1. McCracken, G. and Stott, P. (2012) *Fusion: The Energy of the Universe*, 2nd edn, Academic Press, Oxford ISBN-13: 978–0123846563.
2. Kikuchi, M., Lackner, K., and Tran, M.Q. (eds) (2012) ISBN: 978 92 0 130410 0 *Fusion Physics*, IAEA.
3. Wagner, F. (2012) Fusion energy by magentic confinement; IPP-report 18/3.
4. Lawson, J.D. (1957) *Proc. Phys. Soc. B*, **70**, 6.
5. Ikeda, K. (2007) Progress in the ITER physics basis. *Nucl. Fusion*, **47**, 1–414.
6. Wesson, J. (1987/2011) *Tokamaks*, Oxford University Press. ISBN: 978 0198509227.
7. Bodin, H.A.B. and Newton, A.A. (1980) *Nucl. Fusion*, **20**, 1255.
8. Hirsch, M. *et al.* (2008) *Plasma Phys. Controlled Fusion*, **50**, 053001.
9. Iiyoshi, A. *et al.* (1999) *Nucl. Fusion*, **39**, 1245.
10. Valanju, P.M. *et al.* (2009) *Phys. Plasmas*, **16**, 056110.
11. Ryutov, D.D. *et al.* (2008) *Phys. Plasmas*, **15**, 092501.
12. Boozer, A.H. (2005) *Rev. Mod. Phys.*, **76**, 1071.
13. Hazeltine, R.D. and Meiss, J.D. (2003) *Plasma Confinement*, Dover Books on Physics Series. ISBN: 978 0486432427.
14. Yoshizawa, A., Itoh, S.I., and Itoh, K. (2002) *Plasma and Fluid Turbulence: Theory and Modelling*, Taylor & Francis, Bristol. ISBN: 9780750308717.
15. Purwins, H.-G. (2011) *IEEE Trans. Plasma Sci.*, **39**, 2112–2113 (Special Issue – Images in Plasma Science 2011).
16. Grindrod, P. (1996) *The Theory and Applications of Reaction Diffusion Equations: Patterns and Waves*, Oxford Applied Mathematics and Computing Sciences Series ISBN: 978 018500049.
17. Martin, P. *et al.* (2007) *Plasma Phys. Controlled Fusion*, **49**, A177
18. Wagner, F. (2007) *Plasma Phys. Controlled Fusion*, **49**, B1.
19. Lee, G.S. *et al.* (2001) *Nucl. Fusion*, **41**, 1515.
20. Wu, S. *et al.* (2007) *Fusion Eng. Des.*, **82**, 463.
21. Saxena, Y.C. *et al.* (2000) *Nucl. Fusion*, **40**, 1069.
22. Ishida, S. *et al.* (2011) *Nucl. Fusion*, **51**, 094018.
23. Rostagni, G. (1995) *Fusion Eng. Des.*, **4**, 301.
24. Prager, S.C. *et al.* (1990) *Phys. Fluids B*, **2**(6), 1367.
25. Wolf, R.C. (2003) *Plasma Phys. Controlled Fusion*, **45**, R1.
26. Boozer, A.H. (1980) *Phys. Fluids*, **23**, 904.
27. Nührenberg, J. and Zille, R. (1988) *Phys. Lett. A*, **129**, 113.
28. Anderson, F.S.B. *et al.* (1995) *Fusion Technol.*, **27**, 273.
29. Obiki, T. *et al.* (2001) *Nucl. Fusion*, **41**, 833.
30. Palumbo, D. (1968) *Nuovo Cimento B*, **53**, 507.
31. Igitkhanov, Y. *et al.* (2006) *Fusion Eng. Des.*, **81**, 2695.
32. Ward, D. (2009) *Fusion Sci. Technol.*, **56**, 581.
33. Klix, A. *et al.* (2008) *Fusion Eng. Des.*, **83**, 1813.
34. Ehrlich, K. (1999) *Phil. Trans. R. Soc. Lond. A*, **357**, 595.
35. Moeslang, A. *et al.* (2006) *Fusion Eng. Des.*, **81**, 863.
36. Kuteev, B. *et al.* (2011) *Nucl. Fusion*, **51**, 073013.
37. Burkakov, A.V. *et al.* (2010) *Plasma Phys. Controlled Fusion*, **52**, 124026.
38. Bradshaw, A.M. *et al.* (2010) *Fusion Eng. Des.*, **86**, 2770.
39. Manheimer, W. (2004) *J. Fusion Energy*, **23**, 2005.
40. Glaser, A. and Goldston, R.J. (2012) *Nucl. Fusion*, **52**, 043004.

24
Heavy Ion Inertial Fusion

Rudolf Bock and Ingo Hofmann

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

24.1 Introduction

The consequences of global warming are widely known today, and strategies are needed for the realization of alternative energy generation: inexhaustible, sustainable, and free of CO_2. The concept of thermonuclear fusion by *inertial confinement* offers such an opportunity. Its realization in a gross technical process is the subject of a worldwide research effort for more than four decades. The main physics challenge of this technique is the extreme energy concentration and efficient confinement.

The concept of inertial confinement fusion (ICF) is based on the fact that beams of lasers and high-energy heavy ions offer the opportunity to concentrate large amounts of energy in matter at extremely high density on a very small spot at a very short timescale. This concept, proposed in 1962 for laser beams, is pursued since the mid-1960s with laser facilities of increasing power. Its principle is that the deuterium–tritium fuel is enclosed in a tiny hollow sphere, the pellet, that is compressed by ablation and heated up to ignition temperature. Energy is produced by burning the fuel according to the highly exothermic reaction $D + T \Rightarrow {}^4He + n$. By imploding the pellet, the fuel becomes extremely compressed and is kept together for a short time by inertial forces. The stage of stagnation is long enough for burning a substantial fraction of the fuel. This principle of ignition and burn has been investigated during the last decades by large-scale simulation codes and by means of powerful laser facilities.

In the 1970s, a new technique was proposed by A. Maschke, who realized that high-energy heavy ion beams might also be transported to and focused on a pellet with sufficient confinement conditions. As compared to laser pulses, heavy ion beams have some intriguing features that are particularly relevant with respect to the economy of energy generation in a reactor combined with a heavy ion driver. The main advantage in this respect is the physics of energy deposition in the pellet as well as the efficiency and repetition rate of particle accelerators. It is the main subject of this report.

The heavy-ion-driven inertial fusion (HIDIF) has been investigated since 1976, when this concept was presented by Maschke and Martin on the famous ERDA (Energy Research and Development Administration) workshop in Oakland/California [1]. In Germany, Heavy Ion Inertial Fusion studies for energy generation (IFE) were launched in 1979 in the framework of government-funded

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

research program by an international collaboration and it was continued later in a European frame. The driver is based on a linear accelerator combined with sets of storage rings. Alternatively, heavy ion inertial fusion investigations were pursued in Berkeley based on a different driver concept, the induction linac.

The following report is based on the results of two comprehensive conceptual design studies for a heavy-ion-beam-driven inertial fusion reactor, which have been worked out in a collaboration between German research groups and the University of Wisconsin, the HIBALL study (1981–1985) [2], and the HIDIF study (1995–1998) [3] elaborated by a European Study Group under the leadership of CERN and GSI. Results of the Berkeley studies were reported in a review paper [4].

24.2 Overview: Basic Issues and Key Parameters for an IFE Plant Driven by Heavy Ion Beams

Before we proceed with the presentation of the building blocks of an IFE plant, it seems to be reasonable to first have a look at the basic facts and orders of magnitude.

24.2.1 Basic Facts and Numbers

The concept of inertial confinement is to compress small amounts of DT fuel enclosed in a spherical capsule to high density and to heat it to up to ignition temperature. Within the extremely short time interval the fuel is kept together by its inertia, a considerable amount of fuel is burnt according to the fusion reaction

$$\mathrm{D} + \mathrm{T} = {}^4\mathrm{He} + \text{neutron} + 17.6\ \text{MeV}$$

As compared to other fusion reactions, this reaction has a very high specific energy release of $q_{\mathrm{DT}} = 3.37 \times 10^{11}\ \mathrm{J\,g^{-1}}$.

As in magnetic confinement, the *confinement parameter* $n\tau$ has to satisfy the Lawson criterion $n\tau > 10^{15}\ \mathrm{s\,cm^{-3}}$, where n is the plasma density and τ is the confinement time. Obviously, both n and τ required for inertial fusion are extremely different from those of magnetic confinement. Because of the short confinement time τ of about 10^{-10} s for inertial fusion, an extremely high fuel density is necessary, which needs to be achieved by beams of lasers or heavy ions. In inertial fusion, instead of $n\tau$, the product ρR is more commonly used. In case of a spherical configuration of the burning DT fuel with a mass density ρ and an initial radius R, the confinement time is given by the time of the rarefaction $\tau = R/c_{\mathrm{s}}$, which spreads with the sound velocity c_{s} [5].

The *fraction of fuel* burnt within the short time of stagnation is obtained as

$$\varphi = \frac{\rho R}{(H_{\mathrm{B}} + \rho R)}$$

with the *burn parameter* $H_{\mathrm{B}} = 4c_{\mathrm{s}}(m_{\mathrm{D}} + m_{\mathrm{T}})/\langle\sigma v\rangle$, m_{D} and m_{T} being the D and T mass, respectively, and $\langle\sigma v\rangle$ being the DT reaction rate. This expression for φ takes into account depletion of fuel by burn. With the burn parameter H_{B} of about $7\ \mathrm{g\,cm^{-2}}$ and 30% of burn-up, the ρR needs to be about $3\ \mathrm{g\,cm^{-2}}$. Keeping ρR fixed, the fuel mass scales with ρ^{-2}. In case of 1 mg fuel, which is a reasonable amount for such a microexplosion, a ρR of $3\ \mathrm{g\,cm^{-2}}$ requires a fuel density of $300\ \mathrm{g\,cm^{-3}}$ leading to a thermonuclear energy release of 100 MJ. As the reactor cavity has to stand the microexplosion, the amount of DT per pellet is limited to about 10 mg. This is the reason for the

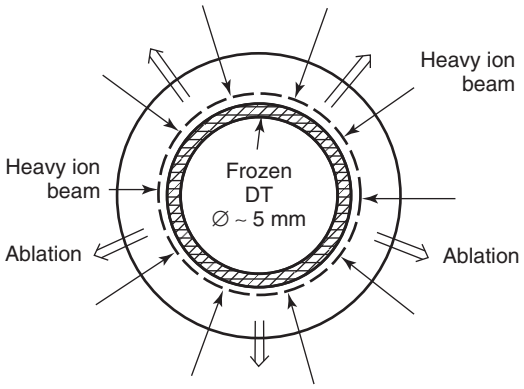


Figure 24.1 Spherical shell ("the pellet") with cryogenic DT fuel, compressed by ablation. The hollow sphere is filled with deuterium–tritium fuel at cryogenic conditions. The outer shell, glass or metal, heated by heavy ions is evaporated (direct drive). By the radial pressure, the fuel is compressed and heated up to ignition temperature and ignited by shock waves.

extremely high fuel compression, which is an essential condition to make ICF work.

24.2.2 Working Principle of the IFE Plant

On the basis of these numbers, the proposed procedure for energy generation is as follows: the plant consists of two main components, the driver and the reactor vessel, and both are clearly separated, which is an enormous advantage for its operation.

The DT fuel is filled in a hollow sphere, the pellet; cooled down to cryogenic temperatures with the fuel frozen inside the shell (Figure 24.1); and injected into the reactor chamber. Having reached the center of the chamber, it is ignited by a large number of synchronous heavy ion beam pulses. The total energy for ignition needs to be about 5–10 MJ. Isentropic compression is achieved by adequate pulse shaping of the heavy ion beams. Simulation codes predict that an energy gain of the order of 100 or more is obtained (Figure 24.2) [6].

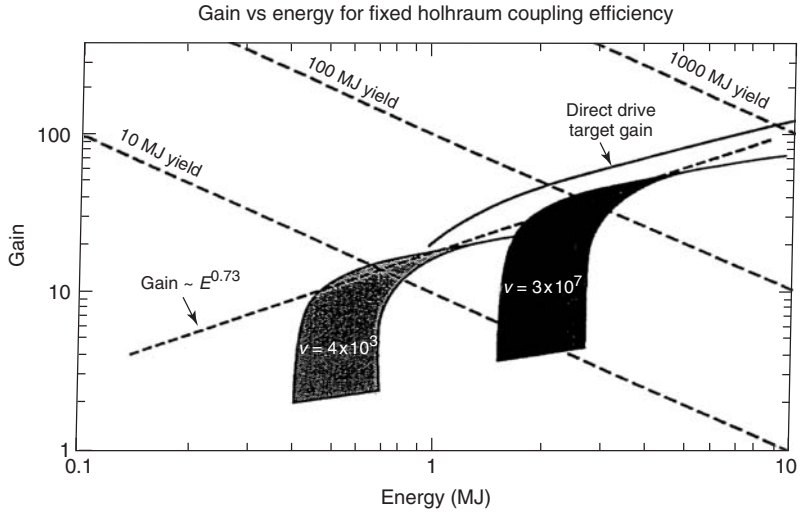


Figure 24.2 Pellet gain depending on driver pulse energy (MJ).

The first wall of the reactor vessel consists of a fluid wall of lead lithium *protecting* the solid structures of the vessel and *breeding* tritium fuel by the reaction $^{6}\mathrm{Li} \Rightarrow \mathrm{T} + {}^{3}\mathrm{He}$. In addition to wall protection and breeding, the fluid is used to transport the produced *energy* out of the chamber to the conventional electricity-generating systems.

24.2.3 Efficiency Considerations

On the basis of the numbers given in the previous sections, the energy balance is as follows: for a single heavy ion beam pulse of 10 MJ and a gain of 100, the output energy per pulse is 1000 MJ, which is deposited in the fluid wall of the reactor. With this amount of energy and a repetition rate of 10 Hz, the thermal power of the plant is $10\,\mathrm{GW_{th}}$, corresponding to an electrical energy of about $3\,\mathrm{GW_e}$. The overall efficiency of the plant is shown in Figure 24.3. It demonstrates the importance of a *high efficiency* and *high repetition rate* of the driver and exhibits the advantage of heavy ion beams.

24.2.4 Advantages and Problems

One of the advantages of inertial confinement as compared to magnetic confinement is obviously the separation of driver and reactor vessel, which allows optimization of both the main components rather independently. Compared to laser fusion, the deposition of beam energy to the pellet is classical and well understood. Based on experience, the repetition rate of the driver accelerator is not a problem and its efficiency can be achieved by the specific accelerator design. However, the beam intensity and the high beam quality at space charge limits will be the main object of future research. A qualitative rating of specific features of various driver scenarios is given in Table 24.1.

24.3 Target Physics and Target Design

Understanding the physics of the target is the key to realize the concept of energy generation by inertial fusion. The results of investigations on beam energy deposition, dynamics of the compression, instabilities, radiation symmetry, and ignition and burn are the ingredients for the target designers. Our knowledge of all these issues is based on a wealth of theoretical investigations and simulations, as well as experiments at the existing big laser facilities and with heavy ion accelerators. They contribute to our understanding of the extremely complex and highly coupled processes during compression and burn and, very importantly, allow predictions on the energy gain

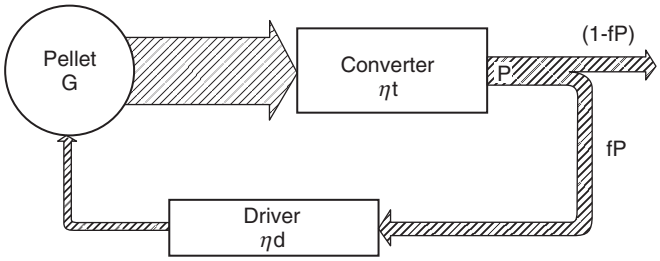


Figure 24.3 Efficiency considerations for an IFE plant.

Table 24.1 Qualitative rating of various inertial fusion driver candidates.

Requirements	CO_2 laser	KrF laser	Free electron laser	Heavy ion accelerator	Light ion accelerator
Efficiency (10–20%)	?	?	+	++	+
Focusing (10^{14}–10^{15} W cm^{-2} at 5 m)	+	+	+	+	?
Target coupling (10%)	?	++	++	++	+
Repetition rate (10–20 Hz)	+	+	++	++	+
Cost ($200 J^{-1} at 2 MJ)	+	+	+	+	++

+, Meets requirements.

for various target designs. After several decades of research, many scientific questions and a number of basic technical problems have been solved and thus our expectation that the concept will work has been confirmed. During the recent years, there was considerable progress with new powerful laser facilities – in particular, NIF came into operation – but the breakthrough for ignition is still ahead.

24.3.1 Interaction of Heavy Ions with Matter

Concepts of IFE are based on the fact that lasers as well as heavy ion beams offer the opportunity to concentrate energy in matter at an extremely high density on a very short timescale. By the interaction, cold matter is changed into plasma and it is essential to take into account the specific properties of energy deposition in both cold and ionized matter (Figure 24.4). Different from laser beams, which are absorbed close to the target surface, heavy ions deposit their kinetic energy in the target shell with a well-defined range. The stopping power for heavy ions in matter is extremely high and the specific coupling of energy is well understood. As compared to very light projectile, such as protons, the specific energy deposited by very heavy ions, such as bismuth, is higher by a factor of 10^3. So, heavy ion beams are extremely well suited to efficiently heat the pellet.

24.3.2 Basic Issues of Target Development for Heavy Ion Beams: Direct and Indirect Drive

A necessary condition to achieve ignition and high gain is the radiation symmetry of the pellet. Owing to the efficient energy coupling of heavy ions with the target, irradiation asymmetries may be generated right at the initial phase of ablation and compression. Smoothness of the target surface is an important condition. This is especially true for a limited number of beams, where the symmetry requirement is difficult to achieve in a *direct drive* scheme (as shown in Figure 24.1). In an *indirect drive* scheme, first proposed at Livermore in 1993, the spherical fusion capsule is enclosed by a cylindrical casing of high-Z material in which the kinetic energy of the heavy ion beam is changed by two converters on opposite sides – according to the Stefan–Boltzmann law – into soft X-ray radiation inside the casing (Figure 24.5 a). The spherical fusion capsule is radiatively imploded by the much

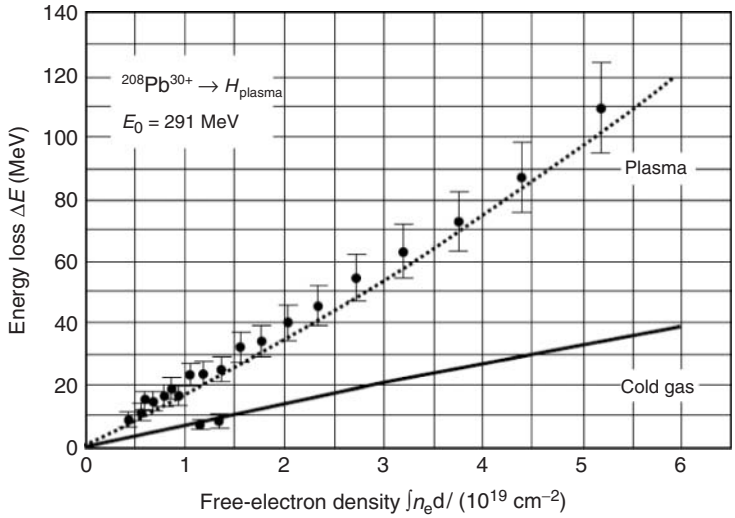


Figure 24.4 Stopping of heavy ions in cold matter and plasma. Measurements of the energy loss of lead ions in cold hydrogen gas and fully ionized plasma. The enhanced energy loss in plasma is due to the higher effective charge state of the lead ions and to the more effective energy transfer in collisions with free electrons.

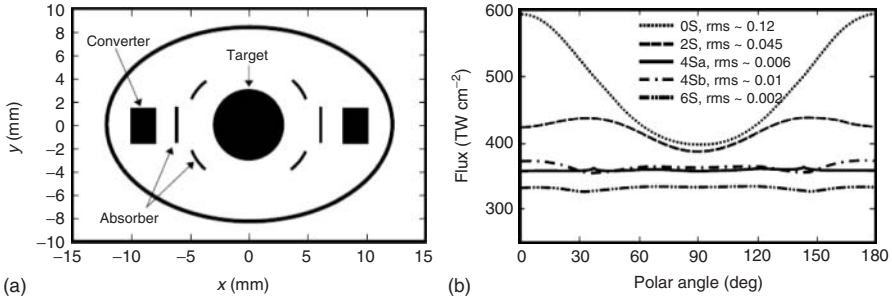


Figure 24.5 Indirect drive targets. (a) The kinetic energy of the heavy ions is converted into hohlraum radiation, which compresses the inner shell filled with fuel. For two-sided target illumination, it is necessary to introduce shine shields between converters and the target in order to achieve sufficient radiation symmetry. (b) Simulations show the effects of symmetrization achieved by additional shields.

more symmetric hohlraum radiation. Many different arrangements of converters and shields of indirectly driven targets have been investigated by simulations. It has been shown that by the introduction of additional shine shields inside the casing, the radiation symmetry can be considerably improved and, as shown by simulations, kept down to the required level of about 1% (Figure 24.5b) [7].

A different concept was suggested by Tabac and Callahan-Miller where the

converter target is distributed in solid angle to achieve the desired radiation symmetry. The case-to-capsule ratio is an important number that affects the energy gain of the fusion pellet. Two-dimensional integrated simulations show, however, that for an optimized heavy ion target, a gain of 130 can still be achieved in a distributed radiator target. Optimized high-gain targets have been designed and investigated by a number of research groups contributing to the *European Study Group on Heavy Ion Driven Inertial Fusion* (*HIDIF*) [8].

24.3.3 Fast Ignition

In 1994, a different concept for ignition was proposed by Tabak *et al.* [9]. Fast ignition allows igniting the target with a considerably reduced amount of energy for the primary driver beam. The usual ignition concept in a spherical compression is based on forming an ignition spark in the center of the shock-compressed fuel. Caused by this spark, a propagating burn wave is launched and propagates throughout the fuel. Fast ignition is based on the idea of decoupling the compression and ignition processes. The fuel is first compressed by an intense heavy ion beam. Ignition is then independently achieved in a separate process by an external laser pulse of ultrashort duration and extremely high intensity. In this case, the beam has to propagate through overdense plasma. Interaction processes of the ultrashort high-intensity laser beam will create a relativistic electron beam that travels toward the highly compressed fuel core. The physics of this process is not yet fully understood, but experimental as well as theoretical investigations worldwide are addressing these problems vigorously, as laser beams with intensities exceeding $10^{19}\,\mathrm{W\,cm^{-2}}$ are available. Recent experiments have demonstrated that such laser beams impinging on a curved target surface can create a high-energy intense proton beam that can be used to propagate toward the highly compressed fuel and finally ignite it.

24.3.4 The Physics of Dense Plasmas

For the design of an efficient high-gain IFE target, the basic processes of beam–plasma interaction and the hydrodynamic response of matter need to be understood on a microscopic level. The detailed knowledge of heavy ion energy deposition in a target that is partly ionized and in a state of high energy density is obviously a prerequisite for the design of the target as well as for realistic simulations of target hydrodynamics. For experiments to measure the conversion of kinetic beam energy into radiation in the regime of radiation-dominated plasmas, the available intensity of heavy ion beams is by far not yet sufficient. Laser facilities such as the PHELIX laser at GSI combined with the new Heavy-Ion Synchrotron Facility FAIR, which is under construction, will allow beam–plasma interaction phenomena in high-density, high-temperature plasmas, such as shock wave compression and equation-of-state physics, to be investigated experimentally. Inertial fusion research offers the opportunity to study outstanding fundamental problems in the field of highly compressed matter and coupled plasmas. Figure 24.6 shows various problems that can be addressed with present and future laser and accelerator facilities.

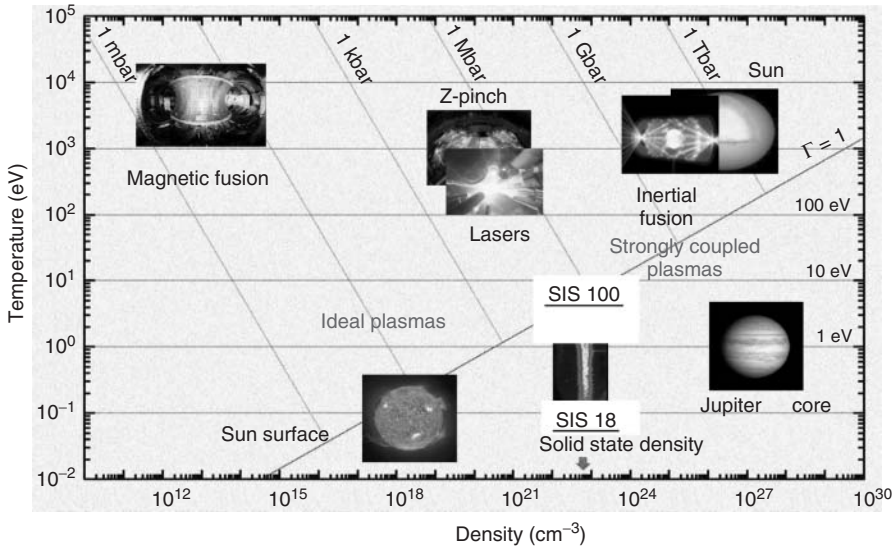


Figure 24.6 Temperature–density diagram for plasmas. The various fields of research on dense plasmas are indicated. On the left-hand side, the ideal plasmas, typical for magnetic confinement, are shown. Indicated on the right-hand side are the field of dense plasmas, relevant to inertial fusion with their exciting implications in astrophysics, equation-of-state physics of dense matter, and strongly coupled plasmas, which are outstanding topics of present fundamental research.

24.4 The Heavy Ion Inertial Fusion Driver

24.4.1 Introduction

Since the mid-1970s, heavy ion accelerators have been considered as possible candidates to drive inertial fusion targets. The first proposal of using high-energy accelerator technology for inertial fusion energy goes back to A. Maschke who presented it at the ERDA workshop in 1976 [1]. Attractive properties of heavy ions in this context are (i) the classical behavior in transport and energy deposition in the fusion pellet and (ii) the highly developed technology of proton and ion accelerators in high energy physics and in applications, such as the spallation neutron sources, where highest intensities and megawatt beam power are required. Examples for such megawatt accelerators are the SNS at Oakridge, USA, and the JPARC facility in Japan. The FAIR accelerator facility under construction at GSI, Darmstadt, will deliver pulsed beams of heaviest ions with 10–30 kJ of energy.

The primary development goals in heavy ion fusion design studies have been to demonstrate that the extremely high peak power needed for target compression and ignition can be achieved by such accelerators. Such high peak power is the proven strength of laser for ignition experiments, such as the US National Ignition Facility (NIF) [10] or the French MEGAJOULE project. Although heavy ion accelerators cannot compete with lasers for ignition experiments, they should be viewed as a technology with high potential

for later application in commercial energy production, where efficiency, repetition rate, and reliability are crucial issues.

Two complementary accelerator scenarios have been considered as potential inertial fusion energy drivers:

- The RF linac with storage rings approach, which benefits from large operating experience with linear accelerators, synchrotrons, and storage rings. As mentioned, it has been in the focus of two system studies coordinated by the GSI, Darmstadt, the HIBALL study in 1980–1984 and the HIDIF study in 1998.
- The induction linear accelerator concept was developed in the United States. Several smaller scale experiments with this new technology (for ions) have been carried out at LBNL, Berkeley.

24.4.2 Basic Principles

24.4.2.1 Requirements for Fusion Energy Production

Target parameters suitable for heavy ion drivers are not significantly different from those of laser drivers, although different optimizations may be needed. The driver efficiency and the actual energy gain of the target may play an important role in driver performance. A reasonable ratio between circulating energy and thermal energy output from the reactor can be reached only by a sufficiently large driver efficiency. For a given target gain, reducing the driver efficiency from 25 to 5%, for example, would increase the recirculating power from 12% to the unacceptably large value of 60%. An efficiency of 20–25% can, in principle, be reached with conventional accelerator technology used in high energy physics for protons or heavy ions, as was demonstrated in the HIBALL study. A typical assumption in heavy ion system studies has been to use a target gain G of 80–100, which is confirmed by different gain models. Extensive code simulations have indicated that such a gain should be possible with an input energy of 5 MJ under ''conservative'' assumptions on entropy increase and hydrodynamic coupling efficiency during the implosion. For effective compression, this energy has to be delivered during 10 ns, which sets the standard power requirement for the accelerator driver to the value of 500 TW.

The specific power that measures the power delivered per unit mass and thus determines the actual temperature and pressure rise in the target beam stopper is given by

$$P = \frac{(E \times I)}{(R \times F)} \ [\mathrm{W\ cm^{-2}}]$$

where E is the total kinetic energy; I, the particle current (A); R, the range ($\mathrm{g\,cm^{-2}}$); and F, the focal spot area ($\mathrm{cm^2}$). In most studies, ranges between 0.1 and 0.3 $\mathrm{g\,cm^{-2}}$ have become standard, although still shorter ranges and lower energies would be preferable. The specific power required to achieve 300 eV plasma temperature in the target is estimated as $P \geq 10^{16}\,\mathrm{W\,g^{-1}}$, which clearly gives preference to a small range, and thus low kinetic energy and heavy ions. For heavy ions, the range is basically a result of the classical Coulomb stopping and anomalous effects are expected to be absent.

24.4.2.2 Intense Heavy Ion Beams

Besides the preference for heavy ions, the optimum choice of their kinetic energy is still open. It can be narrowed down by requesting vacuum transport in the accelerator as well as from the accelerator

Table 24.2 Typical accelerator and target parameters.

Ion species	$^{209}Bi^{1+}$
Kinetic energy	10 GeV
Total energy	5 MJ
Final pulse duration	10 ns
Final momentum spread	3×10^{-3}
Emittance at target	20 mm mrad
Spot radius	3 mm
Ion range	$0.3\,g\,cm^{-2}$
Specific power	$1016\,W\,g^{-1}$

to the reaction chamber and the target, which calls for not too low energies.

A crucial issue is the possibility to transport and focus the heavy ion beams within the different accelerator structures and between the accelerator and the target with appropriate focusing to prevent the beam from diverging due to its emittance and space charge. On this basis, the number of beamlines to the target chamber is estimated as, typically, 50–100. A typical set of beam parameters for an indirectly driven target is summarized in Table 24.2.

24.4.3 Driver Scenarios

The main task of the driver accelerator is to increase the kinetic energy and multiply the current extracted from the ion sources to the level of kiloamperes. The required current multiplication is dictated by the target physics and amounts to about five orders of magnitude. In the RF-linac storage ring scheme, this is performed by a sequence of beam manipulations in the rings, following the acceleration in the RF linac. For the induction linac, acceleration and current compression always occur in a single structure such that the beams always stay close to the space charge limit, according to their kinetic energy.

24.4.3.1 RF Linac and Storage Ring Systems

In the following, we outline characteristic features of a driver satisfying similar requirements as in Table 24.2, but reduced to 3 MJ energy to satisfy the needs of target ignition with an energy gain of <10. A schematic layout is shown in Figure 24.7 from the HIDIF study [3].

- *Linac*: It has the task of providing a pulse of 10 GeV Bi^+ ions sufficiently long to fill all storage rings and thus accumulate the desired total energy. For a pulse current of 400 mA this requires about 1.5 ms to fill all accelerators for one fusion shot. The total length of the RF linac is estimated to 3.4 km.
- *Storage rings*: The linac beam is stacked in the horizontal phase space plane to fill each of the 12 storage rings with ions. The stacked beam leads to stored energy of 250 kJ per ring. The ions fill the RF buckets to create 12 bunches, each 250 ns long.
- *Final compression and focusing*: An important task of the final transport is to remove the time difference of bunches in one ring by delay lines and to provide the voltage – here, by means of induction bunchers – to achieve the final compression of the bunches to a duration of 6 ns as is required for the targets.

24.4.3.2 Induction Linear Accelerator Systems

The induction accelerator concept was first invented by Christophilos in the 1950s for high-current (kiloamperes) electron beams. The acceleration is achieved by the electric field of ferromagnetic cores generated by induction. This concept was proposed by Keefe [11] in the mid-1970s at

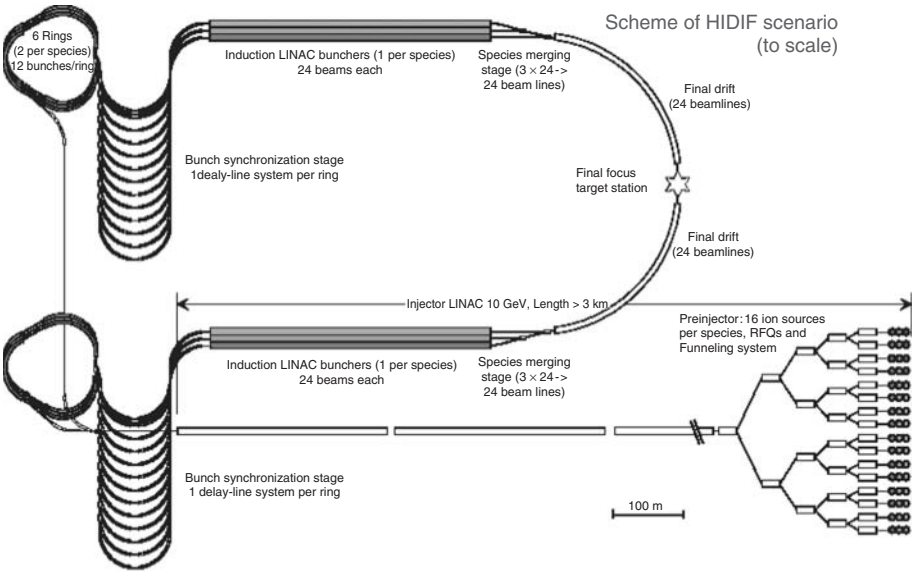


Figure 24.7 Detailed layout of the HIDIF accelerator scenario.

LBNL, Berkeley, for inertial fusion drivers with heavy ions. Different from the RF linac and the storage ring concept, the induction linac is a single-pass accelerator consisting of many induction elements in which the heavy ion bunches are simultaneously accelerated and longitudinally compressed. The main accelerator has a typical length of several kilometers and is preceded by an injector, followed by a drift line in which the heavy ion bunches receive the final compression of the required final pulse length of about 10 ns. Efficiencies of >25% are expected for this type of accelerator. An overall scheme is shown in Figure 24.8.

24.5 The Inertial Fusion Reactor

24.5.1 Reactor Design

It is the great advantage of the inertial fusion reactor design that the chamber walls can be protected from radiation by a thick fluid wall of molten salt in which the energy of the fusion neutrons is deposited and the breeding of tritium fuel is achieved. Among various materials proposed for this purpose, lithium lead ($Li_{17}Pb_{83}$), Flibe (Li_2BeF_4), and Flinabe ($LiNaBeF_4$) have been considered because these chemical compounds (i) are not combustible,

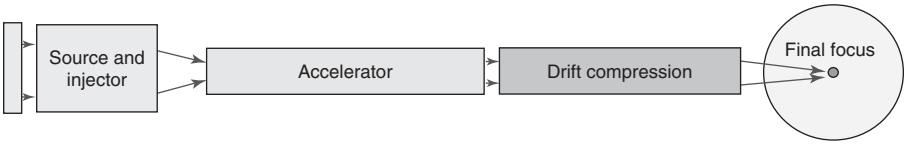


Figure 24.8 Overall scheme of a heavy ion induction linac driving a fusion target.

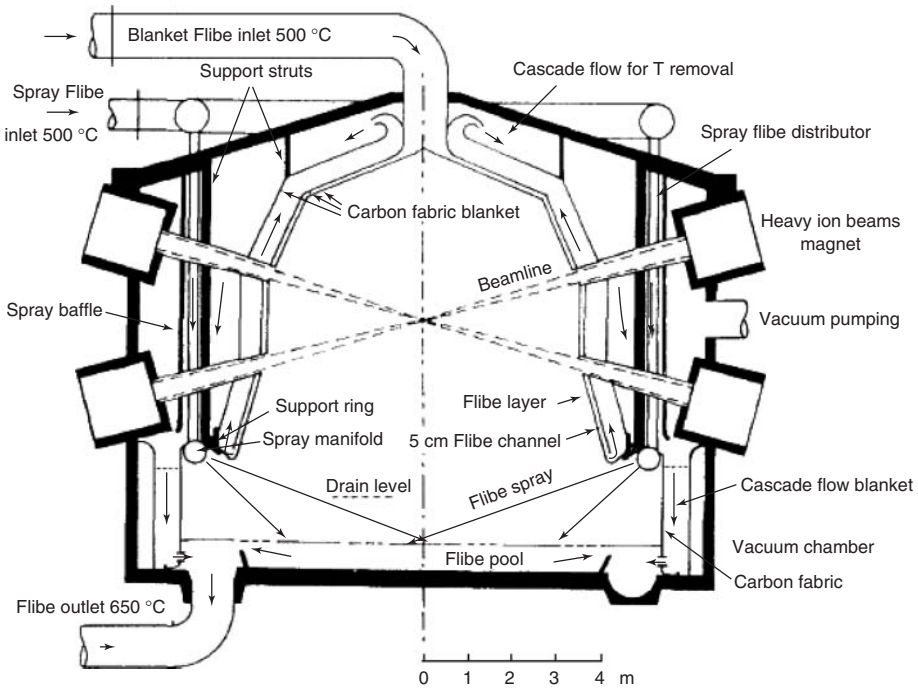


Figure 24.9 OSIRIS reactor chamber design.

(ii) have a very low solubility for tritium, (iii) have a low activation rate, and (iv) in particular, have an extremely low vapor pressure at the operational reactor temperatures. Lithium lead has been chosen for the HIBALL design, and Flibe and Flinabe were selected for the recent reactor concepts. The heat is transported by this liquid to the conventional electricity-generating facilities. Several reactor designs have been developed by various groups: with the HIBALL design (1981, 1985) [2], a direct drive concept, the economic feasibility of a heavy ion beam driver concept was shown for the first time. It was followed by the Japanese design HIBLIC. More recent reactor concepts were developed in the United States for laser and for heavy ion beams, namely, HILYFE II and OSIRIS [12]. A cross section of the cylindrical reactor vessel of OSIRIS is shown in Figure 24.9. Some characteristic power plant parameters are given in Table 24.3.

Table 24.3 Power plant parameters.

Driver energy per pulse	7 MJ
Target gain	57
Target yield	400 MJ
Pulse repetition rate	6 Hz
Driver efficiency	38%
Fusion power	2400 MW
Thermal power	2832 MW
Conversion efficiency	44%
Gross electric power	1246 MW
Auxiliary power	50 MW
Pumping power	27 MW
Driver power	110 MW
Net electric power	1058 MW
Driver cost	2.78 G$
Other plant cost	2.27 G$
Total plant cost	5.05 G$
Cost of electricity (COE)	7.18 ct kWh^{-1}

The following description of reactor components is based mainly on the RPD 2002 design of OSIRIS [12], which is the most recent and first self-consistent reactor design driven by 120 heavy ion beams. A two-sided illumination is used, with 60 beams from each side. Flinabe is used as the molten salt blanket for protecting the first structural wall from neutron and blast damage as well as shielding the beam ports and superconducting final-focus magnets from radiation. Many flow dynamical investigations and experimental studies have been performed for various geometrical arrangements of jets in the new designs taking into account also the neutron-induced isochoric heating of the blanket.

Several conditions must be considered to allow operation at high repetition rate. Following the explosion of the target, the resulting Flinabe splash from the previous pulse must be cleared away to allow target injection and beam propagation through the chamber for the next shot. From the study of Flinabe cold-spray performance and flow rates, sufficient condensation should be achieved for a 6 Hz operation.

24.5.2
Target Techniques

Much work has been done for target design and for the development of the target preparation technique. The DT fuel is filled into microshells by diffusion. Filling and layering techniques have been developed to prepare cryogenic fuel layers in the shells with the necessary smoothness and high precision. Target fabrication techniques as well as a target factory have been developed mainly in the United States [13], Russia, and Japan. A helium gas gun has been proposed to inject targets horizontally into the chamber. Target acceleration of $1000\,\mathrm{m\,s^{-2}}$ up to $100\,\mathrm{m\,s^{-1}}$ takes 100 ms and requires a 5 m long gun barrel. At this acceleration, the cryogenic state of the fuel can be maintained during the whole injection process.

24.5.3
Tritium Technology, Neutronics, and Safety Aspects

Tritium has low solubility in Flibe and tends to diffuse out. A vacuum disengager has been designed to remove the tritium from the molten salt after breeding. Two disengagers in series reduce the tritium concentration in Flibe by a factor of 10^5.

With HILIFE II, detailed investigations on neutron flow, tritium breeding ratios, activation, and radiation damage have been carried out with the following results: (i) the tritium breeding ratio is calculated to be 1.17, (ii) the energy released by nuclear reactions in the blanket and chamber structures increases the fusion power by 18%, and (iii) radiation damage rates in terms of displacements per atom (dpa) turn out to be sufficiently low.

24.6
Conclusions

Since the ERDA workshop in 1976, where Heavy-Ion Inertial Fusion was first discussed in a broader community, noticeable progress has been made during the following two decades. With the conceptual design of HIBALL – based on direct drive – for the first time, a viable concept for target, reactor chamber, and driver was developed and its technical feasibility within a reasonable economic frame was demonstrated. Heavy ion beams thus became a promising driver option for ICF.

Also for the enhanced requirements of indirect drive, the heavy ion accelerator continues to be a viable driver candidate. Since the demonstration of the HIBALL design, some more sophisticated concepts have been developed (HYLIFE, OSIRIS), showing the attraction of the inertial fusion concept and the progress in this field. Many experimental activities have been started for the study of existing problems at various places, in particular, at LBNL, Berkeley; GSI, Darmstadt; Russia; and Japan. The target, however, is still based mainly on computer simulations with assumptions that might not be realistic in all details.

For the heavy ion approach of IFE, it was the strategy that breakeven and ignition will be achieved with the next generation of laser facilities, the NIF in Livermore and Laser Megajoule in France. Because NIF is already in operation with full power since 2013, first results on target performance have been obtained, but not yet conclusively evaluated. The results of the ongoing evaluation of these experiments will offer a valuable data basis for inertial fusion, in general, and to reassess the heavy ion case, in particular.

Glossary

Driver: Driver is the facility delivering the beam for igniting the target. Various facilities have been considered, for example, accelerators for heavy ions and protons, laser facilities.

Fuel: The fuel is a solid (cryogenic) mixture of deuterium and tritium with an atomic mass ratio 1 : 1. Heavy ion beams provide the energy to heat the fuel up to reaction temperatures.

Fusion: Fusion reactions of light nuclei, such as hydrogen and other light elements, are exothermic and deliver large amounts of energy. The energy is carried mainly by the emitted neutrons and γ-rays. The reaction with the highest energy release is the fusion of deuterium and tritium.

Storage rings: The beams for igniting the target need to be of relatively high kinetic energy, high intensity, and extremely short pulses. For that purpose, the accelerated particle beams are merged (stacked) and stored in a set of rings and are finally bunched and compressed and simultaneously focused on the target with high precision.

Target: The target is a small shell (''the pellet'') filled with fuel and heated up to reaction temperatures by interaction with the heavy ion or laser beam. The form of the target may be a shell or cylinder or a more sophisticated arrangement of absorbers to raise its efficiency for energy release.

References

1. Bangerter, R.O. *et al.* (1976) ERDA Summer Study of Heavy Ions for Inertial Fusion. Report LBL-5543, Lawrence Berkeley Laboratory, Berkeley, CA.
2. (a) HIBALL: Badger, B. *et al.* (1981) HIBALL, A Conceptual Heavy-Ion Beam Driven Fusion Reactor Study. Report KfK 3202, Kernforschungszentrum Karlsruhe (KfK); and (1985) KfK 3840;
(b) Bock, R. *et al.* (1983) HIBALL II, an improved conceptual heavy ion beam driven fusion reactor study, Report TC392/29, IAEA.
3. Hofmann, I. and Plass, G. (1998) HIDIF, European Study Group on Heavy Ion Driven Inertial Fusion. Report GSI-98-06, Gesellschaft fuer Schwerionenforschung (GSI).
4. Bock, R., Hoffmann, D.H.H., Hofmann, I., and Logan, G. (2004) in *Landolt-Börnstein*, Vol. VIII/3B, Energy (ed. K. Heinloth), Springer, Berlin, pp. 529–554.

5. (a) J. Meyer-ter-Vehn, *Nucl. Fusion* **22**, 561 (1982); (b) J. Meyer-ter-Vehn and M. Murakami, *Particle Accel.* 37–38, **519** (1992).
6. (a) J.D. Lindl, *Phys. Plasmas* **2**, 3933 (1995); (b) J.D. Lindl, *Nuovo Cimento* **106A**, 1467 (1993); and (c) J.D. Lindl *Inertial Confinement Fusion*, Springer, New York (1998).
7. Maruhn, J. *et al.* (1997) *Fusion Technol.*, **31** (251), 265.
8. Atzeni, S. *et al.* (1998) HIDIF, European Study Group on Heavy Ion Driven Inertial Fusion. HIDIF-Report GSI-98-06, Gesellschaft fuer Schwerionenforschung (GSI), p. 161.
9. Tabak, M. *et al.* (1994) *Phys. Plasmas*, **1**, 1626.
10. NIF (a): E. Moses *et al.*, The national ignition facility, *Eur. Phys. J. D* **44**, 215 (2007); (b) S.H. Glenzer *et al.*, Demonstration of ignition radiation temperatures. *Phys. Rev. Lett.* **106**, 085004 (2011).
11. (a) Keefe, D. (1976) ERDA-Workshop, 1976 [1]; (b) Keefe, D. (1981) *Particle Accel.* **11**, 197.
12. Meier, W.R. *et al.* (1992) OSIRIS and SOMBRERO, ICF Power Plant Designs. Final Report WJSA-9201 DOE/ER/54100 (2 Vols.), W.J. Schafer Associates.
13. Monsler, M.J. *et al.* (1994) *Fusion Technol.*, **26**, 87331, 280 (1997).

Index

Encyclopedia of Nuclear Physics and its Applications, First Edition. Edited by Reinhard Stock.

d

f