IOP Concise Physics | A Morgan & Claypool Publication

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

IOP | ebooks

IOP Concise Physics | A Morgan & Claypool Publication

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

IOP | ebooks

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Morgan & Claypool Publishers

ISBN 978-1-6817-4513-8 (ebook)
ISBN 978-1-6817-4512-1 (print)
ISBN 978-1-6817-4515-2 (mobi)

DOI 10.1088/978-1-6817-4513-8

Version: 20170401

IOP Concise Physics
ISSN 2053-2571 (online)
ISSN 2054-7307 (print)

A Morgan & Claypool publication as part of IOP Concise Physics
Published by Morgan & Claypool Publishers, 40 Oak Drive, San Rafael, CA, 94903 USA

IOP Publishing, Temple Circus, Temple Way, Bristol BS1 6HG, UK

Dedicated to my parents.

Thank you both for everything.

Contents

Preface

This book aims to shed light on the fascinating nature of the universe. We'll begin with a brief history of time and the fundamental forces and particles at play. Mind-bending concepts follow, including Einstein's theories of relativity, which describe how spacetime stretches and warps. After that, we'll dive into quantum physics, which covers the mysterious ways that particles and light behave. Finally, we'll explore black holes, dark matter, dark energy, and recent discoveries in science, including gravitational waves that ripple through spacetime.

The Universe Untangled is written for everyone. A professor may find it useful for a course such as General Science, Philosophy of Physics, Philosophy of Science, or Astronomy. A high school teacher might recommend it to curious students. And any popular science enthusiast is sure to be inspired by its contents. So welcome, and enjoy!

Acknowledgments

A warm thank you to: Professor Stephen Barr, for all of his time and effort through the editing process; Professors Barry Walker, Stuart Pittel, Harry Shipman, and John Clem for their help along the way; Brian Greene, for inspiring me to pursue physics; all of my professors at Boston College, especially Professor Graf for his time and patience; and my family and friends who have provided feedback and support. Thank you all!

Author biography

Abigail Pillitteri

Abigail Lorraine Pillitteri has been proclaimed a modern day Renaissance woman. Now a supervising editor of educational science content, she has written for the U.S. Department of Energy and major publishers of next-generation educational products. Her contributions include the design and content production of educational games, digital mathematics lessons, professional development courses, and science worktexts. Independently, she has published three books of poetry, and she paints colorful conceptual artwork. Her free-verse poetry is rhythmic, emotionally raw, and honest, with occasional analogies to physics that connect the worlds of science and soul. Some of her paintings are also infused with scientific concepts, and her artwork has been displayed in events, galleries, and homes nationwide.

Abigail was born in small-town New Jersey to John and Gail Pillitteri, a home-remodeler and an artist/assistant teacher. She earned her Bachelor of Science in Physics & Philosophy at Boston College and a Master of Science in Physics at the University of Delaware. Her career path as a science writer and editor became clear during graduate school, where she was recognized for her keen writing abilities. Physicists Brian Greene and Neil deGrasse Tyson quickly became her figures of inspiration. She strives to make the miraculous concepts of physics accessible to all audiences, and to show the world how the realms of art and science are deeply connected.

IOP Concise Physics

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Chapter 1

The birth and growth of the universe

If we're going to start somewhere, we may as well start from the beginning. Let's rewind our clocks to what scientists believe is the earliest event in the history of our universe: the Big Bang. We'll see if we can wrap our heads around the beginning of time, and then we'll walk through some of the major events that have occurred in outer space since then.

Most people have heard of the Big Bang since it is our best explanation for how the universe came to be the way we observe it to be today. The words 'Big Bang' are a bit misleading, because this event took place microscopically and may not have produced any sound. Sound requires a medium, such as air, to travel through. But there was no medium around yet, so the event was probably silent. The name 'The Big Bang' actually came from an opponent of the Big Bang theory, Fred Hoyle, but it nevertheless stuck [1].

So what was the Big Bang? Imagine everything in the universe condensed into one tiny point of matter and energy. This miniscule, dense object not only contained the energy and atoms of everything that will ever exist, it also contained both space and time themselves. Scientists refer to this point as 'the Big Bang singularity'. It is a stretch of the imagination to try to picture this all-encompassing object, because our minds want to place it in some region of space at some particular instant in time. But space and time were born from this singularity with everything else in the universe. We cannot define the location of the singularity, for there was no space yet in which to place it. We cannot speak of what happened before it came to be because there was no such thing as time yet. There was no concept of 'before'. Time began at the Big Bang, the beginning.

For argument's sake, let's consider the possibility that the universe is infinitely old. If this were true, then an infinite number of events would have already occurred before we came into existence. Looking backward in time, we would see yesterday followed by the day before yesterday, and then the day before that, and then the day before that… forever. That means, if the universe were infinitely old, the present moment could never come to be. An infinite number of events would have to happen before it, which is impossible. Well, clearly, the present moment has been reached, as you are reading these words right now. Since the string of events in history was able to reach the present moment, then there had to be a beginning of time. The Big Bang is what set our universe's clock ticking.

Keep in mind that I'm talking about space and time in *our* universe. Perhaps something unknown to us existed, and from that unknownness, our universe sprouted. That very well may be the case, but we can't think of that unknownness in the same way that we think of space and time in *our* universe. Perhaps there was some other space and time outside of our space and time, or perhaps something very different from space and time existed and gave rise to the Big Bang. We really don't know. For our purposes here, we'll keep our discussion focused on space, time, matter, and energy in *our* universe.

Okay, so back to the Big Bang. The Big Bang theory cannot describe exactly what happened the instant that the Big Bang occurred, but it comes pretty darn close. Within the first super tiny fraction of a second (less than a trillionth of a trillionth of a trillionth of a second) the Big Bang singularity contained the same amount of matter and energy that we have today, but in a different form. Little is known about the state of matter this early in time. We do know that the universe expanded incredibly quickly and cooled down in a period known as 'inflation'. The universe continues to expand today, but the expansion was much more rapid during inflation.

At the end of this inflationary period, there were elementary particles, light, and energy. When I say 'elementary particles' I mean particles even smaller than atoms—particles that we can't break down into smaller pieces, as far as we know. Additionally, some mysterious stuff that we call 'dark matter' may also have existed. We'll save a discussion of the elementary particles and dark matter for later chapters [2].

All of these particles were moving at incredibly fast speeds in an incredibly hot environment. Some of the elementary particles came together to form protons and neutrons. Protons and neutrons are the particles that make up the nucleus of an atom. Different amounts of protons and neutrons will make up the nuclei of different atoms. (Side note: the word ‘nuclei’ just means more than one nucleus.) The smallest nuclei are the nuclei of hydrogen and helium atoms. They only contain a few protons and neutrons. So, as the universe continued to expand and cool, protons and neutrons were able to combine to form the nuclei of helium atoms within the first 20 minutes. Meanwhile, electrons were still floating around freely. It was not until about 379 000 years later that electrons began to orbit nuclei to form hydrogen and helium *atoms*—the first atoms of the universe (figure 1.1).

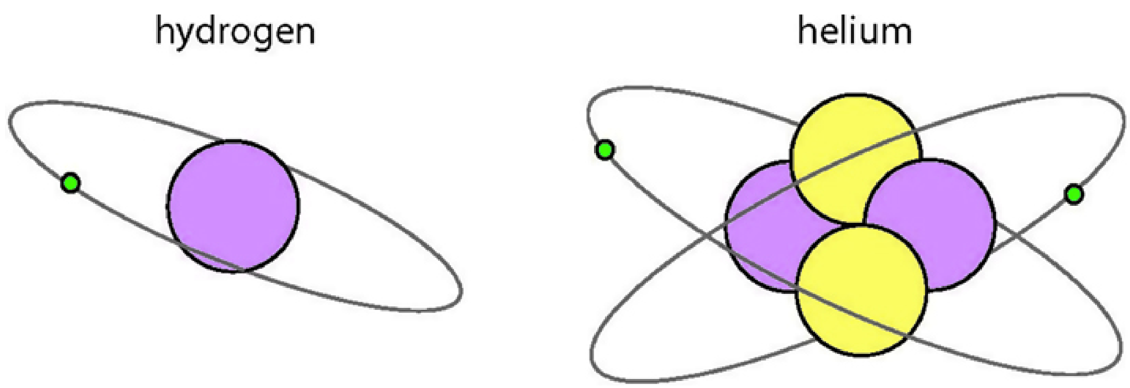

Figure 1.1. These are the common forms of hydrogen and helium atoms. The purple circles represent protons, the yellow are neutrons, and the green are electrons. (Not drawn to scale. Actually, you can assume all images will not be drawn to scale.)

Particles of light, called ‘photons’, continued to travel about on their own. These photons were traveling in the form of microwave radiation (we’ll discuss the different forms that light can take in chapter 5). They became what is known as ‘the cosmic microwave background radiation’. This cosmic radiation is still detectable today. In fact, it is one of the main pieces of evidence that supports the Big Bang theory for the history of the universe [3].

Once hydrogen and helium atoms formed, the universe continued to expand and cool. By the time 200 million years had passed after the Big Bang, the temperature dropped significantly, and particles became less energetic. In dense regions the gravitational attraction of matter was able to bring hydrogen and helium atoms together into clumps. These clumps of matter were distributed fairly uniformly throughout the universe. They were the beginning of galaxies like the Milky Way galaxy in which we live. Within the galaxies, matter in smaller dense regions continued to come together by gravitational attraction. The atoms became close enough together for nuclear fusion to take place. That's when nuclei of smaller atoms fused together into nuclei of larger atoms, releasing energy. These hot, energetic clumps of atoms were the first stars in the universe.

Inside the stars, hydrogen atoms continued to fuse into helium atoms, giving off light and heat in the process. Some of these stars fused all of their hydrogen into helium and then continued to fuse helium into heavier elements like carbon and oxygen. But in forming the heavier elements, not as much heat was given off. So pressure from the gravitational attraction inward began to overcome pressure from the thermonuclear heating pushing outward. In massive stars, this caused the core of the star to collapse. A massive star in this final stage implodes fantastically. The core collapses into a super dense object, and the exterior follows inward. Most of the star then bounces off of the dense core, flying out into the universe. This explosion of a star is called a supernova. Supernovae release their heavier elements into the universe.

The heavier elements are then caught in the galaxies by gravitational attraction. Those elements form new bodies such as planets. Earth itself was made from elements that formed inside stars of the early universe. And about 8 billion years after the Big Bang, the Earth and several other planets in our Milky Way galaxy began orbiting the sun. Thus, our solar system was formed.

The elements which originated in stars caused life to form on Earth. The star that we orbit today (the sun) continues to provide the light and energy necessary for us to survive. So, quite literally, we owe our existence to the stars. And since the elements that make up our bodies came from stars, in a sense we are stardust [4].

While I have left out many of the details, this summary brings us up to speed with the birth and growth of the universe over the past 14 billion years or so. Figure 1.2 summarizes these events. In the following chapters, we'll discuss how this amazing universe works, and we'll explore some of its most mysterious features. We'll unpack the forces that underlie all of nature. With an understanding of space, time, particles, and forces, we can get into Einstein's revolutionary theories of relativity as well as the mind-boggling nature of quantum mechanics. We'll explore some interesting questions from a physical perspective, like whether it is possible to travel through time. And we'll discuss some of the most enlightening physical discoveries of today.

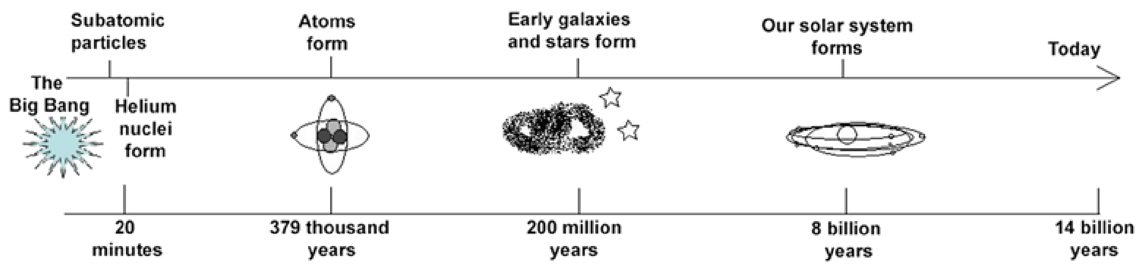

Figure 1.2. Timeline of major events since the Big Bang.

References

[1] Ryden B 2003 *Introduction to Cosmology* (San Francisco, CA: Pearson Education)
[2] Guth A H 1998 *The Inflationary Universe: Quest for a new theory of cosmic origins* (New York: Basic Books)
[3] Peacock J 1999 *Cosmological Physics* (Cambridge: Cambridge University Press)
[4] Goldsmith D 1990 *Supernova: The violent death of a star* (New York: Oxford University Press)

IOP Concise Physics

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Chapter 2

Fundamental forces and particles

There are all sorts of interactions that take place in the universe every day. Atoms fuse together inside stars, electricity provides light to your home, particles turn into other particles, the moon circles Earth. All of the physical interactions in the universe can be explained by one or more of the four fundamental forces: the strong force, electromagnetic force, weak force, and gravity[1]. It is quite amazing that so many different events may occur, yet only four forces are needed to describe them. It is even possible that only one force exists, and the four forces we know of today are just different manifestations of that same force.

Our discussion of forces will involve many different kinds of particles. It is less important that you memorize all of those different particles, and more important that you appreciate the different interactions. Our goal is to gain an understanding of how the physical universe works. We'll try not to get too far into the weeds with details and pay more attention to the bigger picture. But if you do memorize one type of particle, let it be the photon. A *photon* is a particle of light. It will come up countless times in this book (and in the universe at large!). The chapters that follow will get more into the mind-bending concepts of relativity and quantum physics.

The four fundamental forces

What exactly is a force? We have an intuitive understanding of forces on large scales, i.e. within interactions which are observable to the naked eye. For instance, when we push a shopping cart, we have to exert a force on it to move it across the floor. In this large-scale sense, a force is simply a 'push or pull' that a body experiences as a result of an interaction with something else. A force causes a body to change its state of motion [1].

When we speak of the *fundamental* forces, our definition is a bit different. A force is fundamental if it cannot be explained by any simpler interactions. The large-scale forces that we experience can be explained by the microscopic, fundamental forces. For instance, when two billiard balls collide, both balls feel a force that causes them to bounce off each other. This large-scale force is a result of what's happening microscopically: the repulsive electric force is acting between particles within the billiard balls.

The strong force, electromagnetic force, and weak force are mediated via force-carrying particles. That means, in order for one of these forces to act between two matter particles, other force-carrying particles must be involved. These force-carrying particles are usually called 'exchange' particles since they are exchanged between matter particles. The forces themselves are the result of the exchanging of particles.

You might think of these forces in the same way you think of the interaction you experience when you play a game of catch. When you throw a ball to your partner, he catches it and throws it back. The two of you are connected by this game. If one of you doesn't throw the ball back, there is no longer a game of catch. Similarly, matter particles play catch with exchange particles, and they are bound together by this exchange. If they stop exchanging particles, they no longer experience a connection. There is no longer a force between the two matter particles when the exchanging stops.

As you might remember from science class, matter is anything that has mass and takes up space. Some force-carrying particles don't have any mass at all, such as photons or gluons. So those particles aren't really matter; they are the particles that allow matter to experience forces. Matter particles include quarks and leptons. The most important difference between quarks and leptons is that quarks can experience the strong force, while leptons cannot.

The strong force

Scientists used to think that protons and neutrons were fundamental particles which could not be broken down any smaller. They later discovered that there are quarks and gluons inside. Quarks are the matter particles, and gluons are their exchange particles. Quarks playing catch with gluons results in the strong force, which is (not surprisingly) the strongest of the four fundamental forces. When quarks exchange gluons, the quarks become bound together. For instance, inside a proton, three quarks exchange gluons (figure 2.1). So a proton is not a fundamental particle at all. Rather, it is a particle made up of three quarks which 'play catch' with gluons!

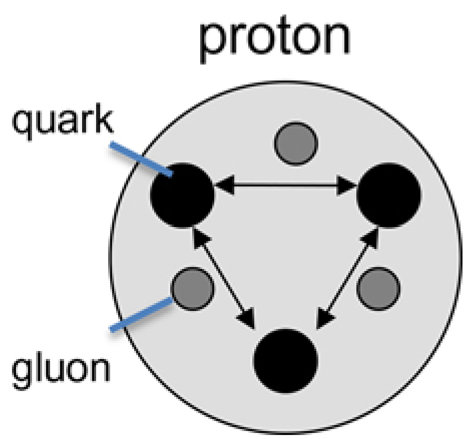

Figure 2.1. The proton contains three quarks (shown in black). They are bound together by exchanging gluons (shown in gray). This exchange results in the strong force. (This image is a simplification to show how quarks exchange gluons to form bonds. While all the quarks in the image appear to be of the same type and all the gluons also appear to be of one type, some are different from others in reality.)

The quarks inside protons also play catch with the quarks that are inside *other* protons and neutrons. That's how protons and neutrons attract one another: their quarks exchange gluons. The image below may help illustrate these interactions in the nucleus of an atom (figure 2.2).

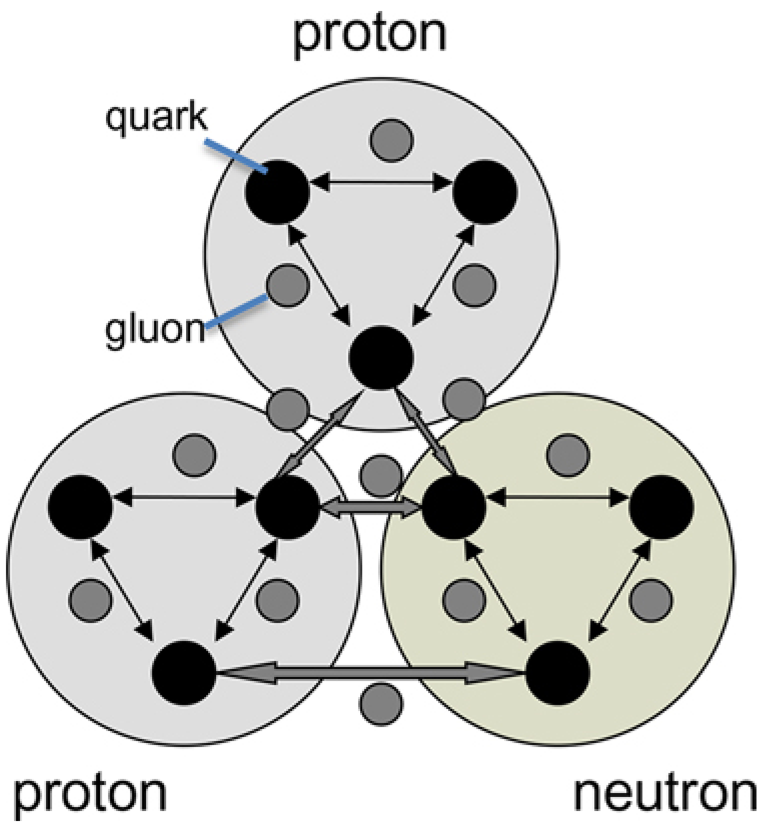

Figure 2.2. Protons and neutrons are bound together as their quarks exchange gluons. (This image is a simplification to show how quarks exchange gluons to form bonds. While all the quarks in the image appear to be of the same type and all the gluons also appear to be of one type, some are different from others in reality.)

Hang on, it looks like the protons can attract each other. This might seem strange to you if you remember the saying, 'same charges repel, and opposite charges attract'. Two protons have the same positive charge, so shouldn't they repel? Well, that saying applies to the *electric* force. Here, we are talking about a different type of interaction—the *strong* force. It is true that opposite charges attract and same charges repel by means of the electric force. But it is also true that quarks exchange gluons, causing protons to attract via the strong force. So both the electric and strong forces act at the same time. The strong force is a very short-range force, so it is only significant between two particles that are both inside the same nucleus. It is *very* strong, so it overpowers the repulsive electric force that the protons also experience.

Now, back to quarks and gluons. These particles are subject to 'confinement', which means they are confined within the larger particles that they make up. We can never observe a single one of them standing alone! There are six types, or 'flavors', of quarks, while gluons do not have flavor. The physicists who discovered the six flavors of quarks playfully named them strange, charmed, up, down, top, and bottom (bottom is sometimes called beauty instead).

Each quark or gluon can come in a variety of 'color states' (or 'colors' for short). There are rules for the possible combinations of quarks and gluons based on their colors. The 'color' of a quark or gluon is nothing like color as we think of it visually. Rather, the so-called colors are associated with a property called *color charge*. Color charge is kind of like a secret language that only quarks and gluons know how to speak. It is a unique property which allows quarks and gluons to be the only particles that can experience the strong force [2].

The electromagnetic force

The next strongest fundamental force is the electromagnetic force, a combination of electric and magnetic forces. It is about 100 times weaker than the strong force. This is the force that keeps the negatively-charged electrons orbiting the positively-charged nucleus of an atom. As we discussed, protons repel other protons via the electric force, but their quarks cause them to attract one another via the strong force. Since the strong force is so much stronger than the electric force at short distances, the protons are able to remain bound together within the nucleus.

We are also familiar with the magnetic force that causes magnetic materials to attract or repel one another. The magnetism of a material is actually caused by the motions of the charged particles in the material. Each electron, proton, and neutron is a small moving charge, and it generates a tiny magnetic field. Collectively, the magnetic fields of all the charged particles produce the magnetic field of the total material. A material's total magnetic field then produces a magnetic force which can act on other magnets.

We have mentioned that electrons, protons, and neutrons are moving charges which produce magnetic fields. Evidently then, there is a connection between electric charge and magnetism. The connection between electricity and magnetism was discovered through experiments involving currents through wires in circuits. Experimentalists discovered that running an electric current through a wire produces a magnetic field which circles around that wire (figure 2.3).

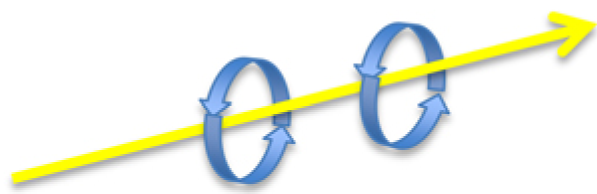

Figure 2.3. Electric current running through a wire (shown in yellow) produces a magnetic field around the wire (shown in blue). This is one of the effects that signifies an intimate relationship between electricity and magnetism.

They also discovered that a magnetic field can act on an electrically charged particle if the particle is in motion. This means that a moving electric charge (such as an electron) must have some kind of magnetic property. Additionally, a changing magnetic field produces an electric field, and a changing electric field generates a magnetic field. When all of these relationships were written in mathematical formalism, it became clear that the electric and magnetic forces were simply different aspects of the same force. We now call this force the electromagnetic force [3].

According to quantum mechanics, an electromagnetic force occurs due to the exchange of photons, the particles of light. The photons in these processes are considered 'virtual' photons. They are virtual because we cannot detect them in a laboratory. We know that they exist only from their effects. For instance, we observe that the proton and electron attract one another, and this attraction is due to the exchange of virtual photons. So the electromagnetic force is experienced by charged particles playing catch with photons. All particles which have an electric charge associated with them are subject to electromagnetic forces; they all can play catch with particles of light.

If you find it difficult to form a mental image of how these microscopic forces exist within the observable world around you, you're not alone. They are not visible to the naked eye, making them difficult to believe. Imagine looking at a marble under a very strong microscope. The marble appears to be one small solid ball to the naked eye. But under the microscope, you see that it is made up of many smaller balls (electrons) and those balls are orbiting around small lumps (the nuclei of atoms). Even more surprising to you are the distances between the electrons and the nuclei. They are so far apart that the majority of the marble is empty space! The moral of the story is, the quantum world hides many secrets since it operates on such small scales. When we zoom out to large scales, the quantum mechanical under-workings are invisible.

Antiparticles

Before we move on to the next fundamental force, there is something else worth mentioning. Our discussion of fundamental forces and particles would be incomplete without antiparticles. Antiparticles have the same mass but opposite electric charge to their corresponding regular particles. The interesting thing about antiparticles is that they can annihilate their corresponding particles. An electron's antiparticle is a positron; if an electron and positron collide, they annihilate and turn into a burst of photons.

Antiparticles play a role in one of the unsolved mysteries of the universe. Scientists have not figured out why we observe more matter than antimatter. If there were equal amounts of matter and antimatter created, then all particles would have annihilated in the early universe. Our universe would be full of radiation created by those annihilations, and nothing else. In other words, we wouldn't be here today if there were an equal amount of matter and antimatter. Something must have happened—some physical phenomenon, or some physical law must exist—to explain why there is more matter than antimatter. Scientists have their theories, but there is no conclusion on this one yet.

The Higgs boson

On the other hand, there *is* one mystery that physicists were able to solve recently, and it involved the discovery of the Higgs boson. Scientists wondered why certain particles, like the W and Z bosons, have mass. They expected those particles to be massless. The conclusion was that those particles could have mass if they were moving through some kind of 'condensate' that fits into their model. The so-called condensate would be something like a liquid, but invisible. This condensate was theorized to be made up of Higgs bosons. They would be everywhere in the universe. Other particles move through the condensate and do not lose energy, but they are affected in some way. Peter Higgs himself, the proponent of the theory, said moving through the Higgs condensate is not really like experiencing drag, but more like how light is bent when it moves through water. He admits the Higgs mechanism is difficult to describe. The important thing is that the Higgs boson was found, and the Standard Model of particle physics would be nonsense without it [4].

The weak force

The weak force, also called the weak interaction, is somewhat unique compared to the strong and electromagnetic forces. It is mediated by different exchange particles called W and Z bosons. This force differs from the others since the exchange particles involved are usually released from a single matter particle rather than transferred from one particle to another. The weak force causes a quark or a lepton to transform into another type of particle. A quark turns into another type of quark by releasing a W or Z boson, or a lepton turns into another lepton in the same manner.

One notable example of a weak interaction is known as beta decay. It is a process that turns a neutron into a proton. What happens is, a down quark inside a neutron emits a W boson. The emission causes the down quark to turn into an up quark. The result is that the neutron turns into a proton.

Meanwhile, the W boson decays into an electron and an antineutrino (a type of lepton). The resulting electron is called a beta particle, so that's why this is called beta decay (figure 2.4). Why do we care? Well, the older an atom gets, the more times it will have undergone beta decay. Thus, by detecting beta particles, scientists are able to determine the age of substances or objects. This process is known as 'radiometric dating'. It has been used to determine the age of many fossils and artifacts [1].

Figure 2.4. Beta decay is a weak interaction where a neutron turns into a proton. This happens when one of the down quarks in the neutron turns into an up quark by emitting a W boson. The W boson decays into an electron and an antineutrino.

Physicists have used particle accelerators to explore the weak force further. They discovered that, at high energies, there were similarities between the function of photons in electromagnetism and the W and Z bosons in the weak interaction. They were able to unify the electromagnetic force and the weak force into the electroweak force. At high energies, they also observed that the electroweak force becomes stronger and the strong force becomes weaker. The quarks of the strong force appear to move more freely than they did at lower energies. Their exchange particles (gluons) begin to operate like the photons, W, and Z bosons of the electroweak force. In theory, at the 'unification energy', the strong and electroweak forces also become unified.

We are not able to reach this energy in the laboratory because our particle accelerators are not large enough to accelerate particles to high enough speeds. However, the Standard Model mathematically predicts this unification, and it has not been proven wrong. Many scientists believe that, at the beginning of the universe, all of the four fundamental forces were unified into one. In fact, if we look far back in time, the Standard Model predicts the unification of three forces: the strong force, the electromagnetic force, and the weak force. But theorists have not yet found a way to incorporate the fourth force: gravity.

Gravity

Gravity is the weakest of the fundamental forces. For particles like protons and electrons, the gravitational force is practically negligible. The other forces overwhelm it. Gravitation has some interesting properties of its own, though. It operates at infinite distances (although this is also true of the electromagnetic force). That means you are experiencing the gravitational effects of bodies on the other side of the galaxy and beyond. And all particles experience gravitational effects of massive bodies. Table 2.1 summarizes the elementary particles and the different forces that they experience.

Type of Particle	Name	Antiparticle	Relation to Forces	Unique Properties
Quark	Up	Antiup	Experiences strong, electromagnetic, weak, & gravity	Color charge, cannot be isolated
	Down	Antidown		
	Strange	Antistrange		
	Charmed	Anticharmed		
	Top	Antitop		
	Bottom	Antibottom		
Lepton	Electron	Positron	Experiences electromagnetic, weak, & gravity	Lepton number
	Muon	Antimuon		
	Tauon	Antitauon		
	Electron neutrino	Positron neutrino	Experiences weak, & gravity	Lepton number, no electric charge
	Muon neutrino	Antimuon neutrino		
	Tauon neutrino	Antitauon neutrino		
Guage boson (exchange particle)	Gluon	Gluon	Mediates strong	Massless, no electric charge
	Photon	Photon	Mediates electromagnetic	
	W^- boson	W^+ boson	Mediates weak	Can change quark flavor
	Z boson	Z boson	Mediates weak	No electric charge
Higgs boson	Higgs boson	Higgs boson	Accounts for mass	No electric charge, recently discovered

Table 2.1. A table summarizing elementary particles and fundamental forces that have been discovered.

Right now, gravity is described by two different theories that don't match up. One is Einstein's general theory of relativity, which claims that gravity is not a 'real' force like the other three fundamental forces; rather, gravity is the result of the curvature of spacetime induced by the presence of mass. This will be discussed in detail in chapter 4. The other theoretical framework is quantum mechanics, which predicts that the force of gravity is transmitted by an exchange particle called the graviton. We have not yet discovered the graviton, but physicists like to think that it exists since the other three fundamental forces all operate due to exchange particles. The existence of the graviton is also predicted within the mathematics of quantum theory. However, when physicists try to combine the general theory of relativity with quantum mechanical theory, the equations break down. They produce non-physical answers, like infinity when a finite number is expected. One of the great tasks of modern physics is to find a theory which can describe gravity on both microscopic and large scales, unifying general relativity and quantum mechanics [5].

Einstein's general theory of relativity was nothing less than revolutionary. Before Einstein's lifetime, the Newtonian description of gravity was accepted as universal law. Newton's equations accurately describe gravitational effects on Earth, but when we look to very large scales, his equations make inaccurate predictions. General relativity makes the correct predictions, and better yet, it illuminates the unity of space and time: they operate as one fabric which can curve and warp when matter is present (figure 2.5).

Figure 2.5. Since we can only see three spatial dimensions, spacetime curvature cannot be realistically envisioned. This is an artist's (the author's) depiction of spacetime curvature around spheres. It is a three-dimensional representation of a four-dimensional phenomenon (three spatial dimensions and one time dimension).

A few years before Einstein wrote his general theory of relativity, he devised the special theory of relativity, which was equally astounding. We'll begin our discussion of Einstein's theories with the special theory of relativity in the following chapter. If you think that time passes at the same rate for all people and objects everywhere, or that space is uniform and unchanging, you're in for an enlightening surprise.

References

[1] Hewitt P G, Suchocki J and Hewitt L A 2008 *Conceptual Physical Science* 4 ed. edn (San Francisco, CA: Pearson Education)

[2] Feynman R P 1985 *QED: The strange theory of light and matter* (Princeton: Princeton University Press)

[3] Newton R G 2007 *From Clockwork to Crapshoot: A History of Physics* (Cambridge, MA: Harvard University Press)

[4] Griggs J 2012 Peter Higgs: Boson discovery like being hit by a wave. [Online] Available at: https://www.newscientist.com/article/dn22033-peter-higgs-boson-discovery-like-being-hit-by-a-wave/?full=true&print=true

[5] Greene B 2003 *The Elegant Universe* (New York: Vintage Books)

[1] Note that I said the *physical* interactions can be explained by these four forces. As for interactions like love, or the creation of ideas, well those might have more going on than we think. We can reduce them to chemical interactions in the brain, but I would guess that's an oversimplification. Only time and discovery will tell.

IOP Concise Physics

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Chapter 3

The truth about space and time: Einstein's special theory of relativity

Reference frames

All motion is relative. You might feel as though you're sitting still right now, but you know that the planet you're on is both spinning on its axis and soaring through the solar system in a path around the sun. Your state of motion may only be defined in relation to something else. So when we speak of motion, we have to define the *reference frame* from which the observations are being made. In your reference frame, you may consider yourself stationary. But in the reference frame of the sun, you are in motion. That means the sun would look at you and say, 'of course you're moving, you're sitting on Earth while it turns like a merry-go-round, and the whole merry-go-round orbits around me!' So when we talk about something moving, we always have to consider what the object is moving *with respect to*. I am sitting still with respect to the ground, but I am moving with respect to the sun and the distant stars.

From your own reference frame, it looks as though the sun is rising into the sky in the morning and falling below the horizon in the evening. This is why so many astronomers were misled when they first formulated theories about the motions of objects in outer space. They assumed that the so-called 'heavenly bodies' orbit around Earth because, from our reference frame, that's the way many objects appear to move. Nowadays, scientific measurements establish our motion in space with respect to the distant stars. Looking toward Earth from a star, you would observe that Earth in fact rotates on its axis and revolves around the sun, while the sun orbits around the center of the Milky Way galaxy.

Here's an on-Earth example to illustrate how motion is relative, which will also prepare us for the less intuitive aspects of special relativity. Let's say our friend Troy is inside a train that's traveling 60 miles per hour in a straight line. Another friend Patty is standing on a platform as the train goes by. Troy is in the train, Patty is on the platform. The train came from a different terminal, and it is not stopping at Patty's stop. So it is passing Patty's platform at a constant speed. Obviously from Patty's point of view, she is stationary while Troy passes by her. On the other hand, from Troy's point of view, he may claim that *he* is stationary while Patty and the platform speed by him at 60 miles per hour. His train moving past the platform at constant speed is equivalent to the platform moving past the train at that same speed.

Let's put it this way: if we take away the train, the platform, and everything else except for Patty, Troy has no way of telling who is in motion. After he passes the platform, he feels as though he is stationary while Patty recedes away from him. Similarly, for Patty, she feels as though she is stationary and Troy is moving away from her. Each of them says, 'I'm sitting still while my friend moves away from me at 60 miles per hour'. The motion can only be defined relatively. Figures 3.1 and 3.2 may be helpful.

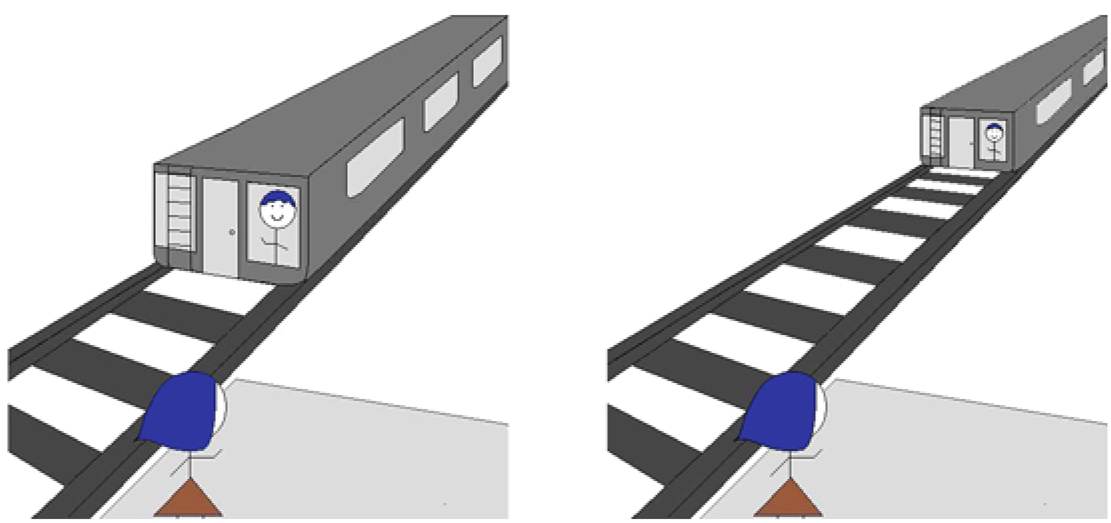

Figure 3.1. From Patty's point of view, Troy's train appears to be receding from her at 60 miles per hour.

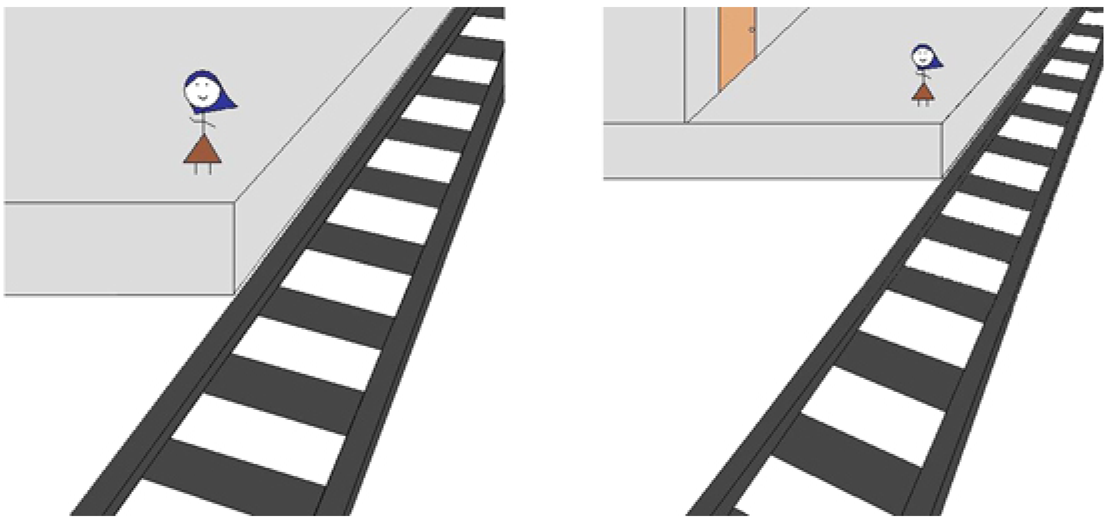

Figure 3.2. From Troy's perspective, the platform appears to recede from him at 60 miles per hour.

Reference frames that are at rest or moving at a constant velocity with respect to each other are called *inertial reference frames.* As long as neither reference frame is accelerating or decelerating, then they are in uniform motion. If the train were to accelerate or decelerate, Troy would be in a *non-inertial* reference frame, and it would be obvious to him that there was a change in speed. For instance, once the engineer[1] hits the brakes, Troy becomes aware of the train's motion. He is forced to lean forward in his seat as the train slows down. His body doesn't want to slow down—it wants to keep moving forward as it was before the conductor hit the brakes. He is experiencing what is known as a *fictitious force*. Meanwhile, Patty does not experience any such forces. She feels as though she is sitting still, as she has felt all along. Troy's experience of a fictitious force, causing his body to lean forward in his seat, is evidence that he is now in a non-inertial reference frame [1].

The speed of light in all reference frames

So far these ideas regarding relative motion may seem fairly obvious. But things get interesting when we consider the motion of light. Scientists have measured the speed of light in vacuum to be about 671 million miles per hour (that's about 300 000 kilometers per second). Light moves at nearly this speed in air as well, so we'll use this number in our example. Let's say Troy is inside his train, which is again moving at a constant velocity of 60 miles per hour. This time, Troy shines a flashlight toward the front of the train. He measures the pulse of light's speed with a measuring device and it reads that the speed of the light is 671 million miles per hour. Patty on the platform also has a measuring device. How fast does she say the light is moving?

You might think that Patty measures the speed to be 671 million miles per hour *plus* the speed of the train, which was 60 miles per hour. But alas, she measures the speed of light to be 671 million miles per hour, the same as Troy's measurement! What's going on here? How is that possible?

It turns out that *the speed of light is the same for all observers*, regardless of their states of motion. This is a fact, and it has been tested countless times through various experiments. It seems that light breaks the rule we have stated, which is that all objects' speeds are relative and can only be defined with respect to something else. Does that mean our rule is incorrect? No, it means that light is an exception to the rule. There is something special about the speed of light. In fact, it is so special that space and time have to change in order to maintain that the speed of light is measured to be the same for all observers!

Speed, by definition, is distance over time. When we say a car travels at a speed of 60 miles per hour, we mean 60 miles (distance through space) are covered in one hour (duration of time). For all observers to see light travel the same speed, their measurements of distance and time have to change. That means that *space and time are not the same for all observers!*

Time: no longer the same for everyone

We'll use a train in an example again, but we'll speed things up a bit so that the effects for Patty and Troy are more obvious. Troy is on a crazy train moving very fast this time, maybe 80% as fast as the speed of light, with respect to the platform. Let's assume Troy and Patty both have superhuman eyesight, and Patty notices Troy's clock on the wall in the train. She looks at the time passing on her wristwatch and compares it to the time passing on Troy's clock. From her perspective, Troy's clock seems to be running slow. She holds up her wristwatch to show Troy the difference. Troy looks at Patty's wristwatch and compares it to his clock. From Troy's perspective, it looks like *Patty's* watch is the one that's running slow.

There are two important points to this example. One is that they are both correct in saying the other person's clock is running slow. We have discussed how motion is relative, and how Troy could claim that he is stationary while Patty recedes away from him since they are in inertial reference frames. From Troy's perspective, Patty is in motion so her watch is running slow; from Patty's perspective, *Troy* is in motion so *his* clock is running slow[2].

Here is the point: *a clock in motion with respect to an observer runs slow compared to the observer's clock.* The clocks run at different rates so that the speed of light remains the same in both reference frames. We'll look to Troy and Patty for an explanation.

This time, Troy has a unique type of clock called a light-clock. It consists of a light source on the floor of the train and a mirror on the ceiling of the train. The clock starts when the light source emits a pulse of light. The pulse of light then bounces between the floor and the ceiling. Each time the pulse of light hits the floor or the ceiling, the clock ticks. Recall that the speed of light is constant and the same for all observers, so Patty and Troy agree on the speed of the light pulse. Now from Troy's perspective, the train is stationary, so he sees the pulse of light initially emitted straight upward. He watches the light continue to bounce straight up and down between the floor and the ceiling (figure 3.3).

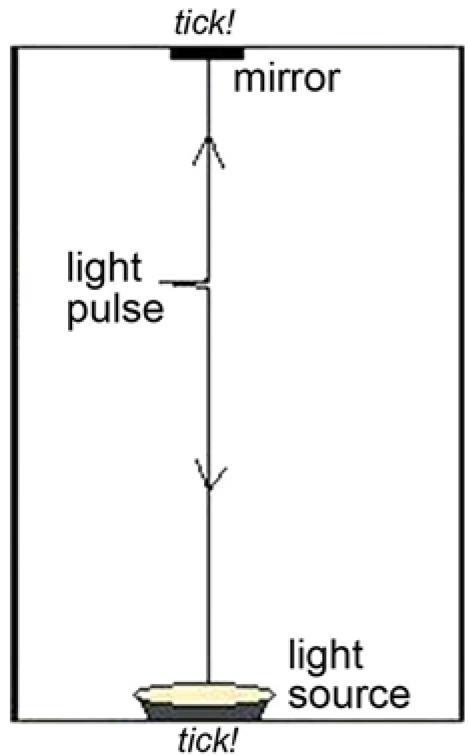

Figure 3.3. Troy's view of his light-clock in the train. The pulse of light is represented by a spike of light intensity between the light source and the ceiling. The light pulse appears to bounce straight up and down since, to Troy, the train is stationary.

But from Patty's perspective, the pulse of light travels along a longer, diagonal path. From her viewpoint, the light source is moving with the train, so it emits the light upward but also slightly sideways in the direction that the train is moving. The moving train does not give the light any additional speed; it only contributes to the light's direction. The result is that the pulse of light appears to travel in a diagonal path from the light source to the ceiling as the train moves forward (see figure 3.4). It is obvious that this diagonal path is longer than the straight-up-and-down path which Troy observes. Since Patty sees the light take a longer path, it takes a longer time to move from the light source to the ceiling. The light pulse bounces off the mirror and travels back toward the floor. It again appears to take a longer path, and thus more time, than what Troy observes. Patty sees that *the time interval between ticks is longer than if the clock were stationary*; *the moving clock ticks more slowly*.

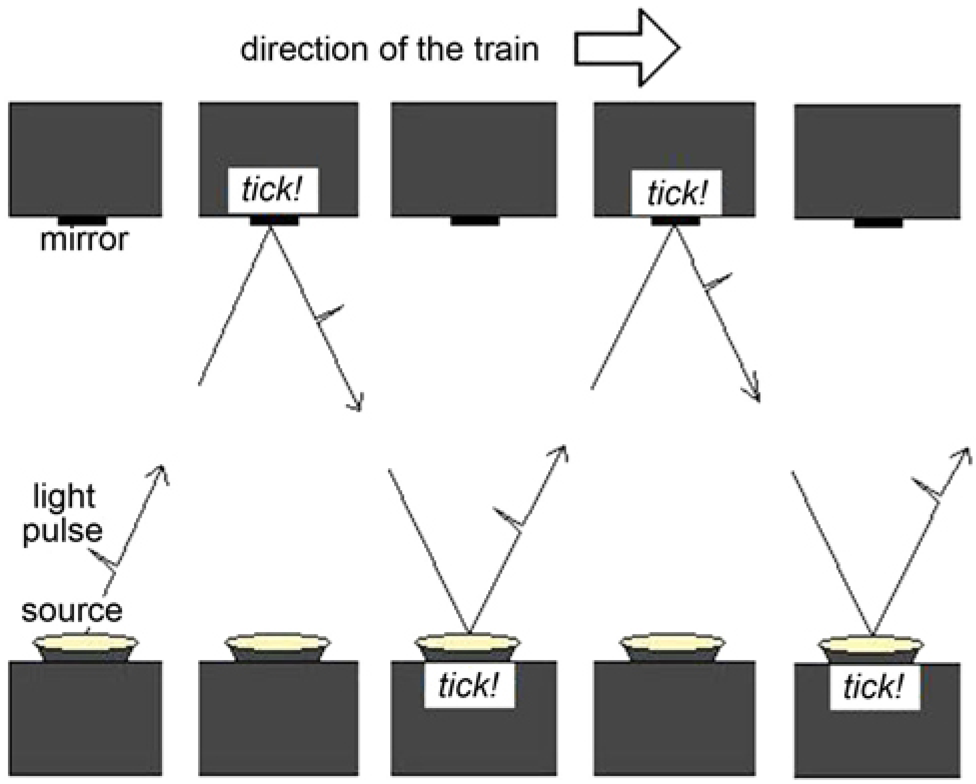

Figure 3.4. Patty's view of the light-clock in the train. Each frame shows a snapshot of the light-clock as it moves to the right with the train. From this perspective, the light pulse appears to travel along a diagonal path. Patty sees Troy's clock ticking less often than he sees. She sees his clock running slow.

Patty sets up her own light-clock on the platform for comparison (see figure 3.5). Indeed, she observes that there is a longer interval between ticks of Troy's clock than her own. That means Troy's clock appears to tick *less often* than hers. Troy's clock appears to tick more slowly.

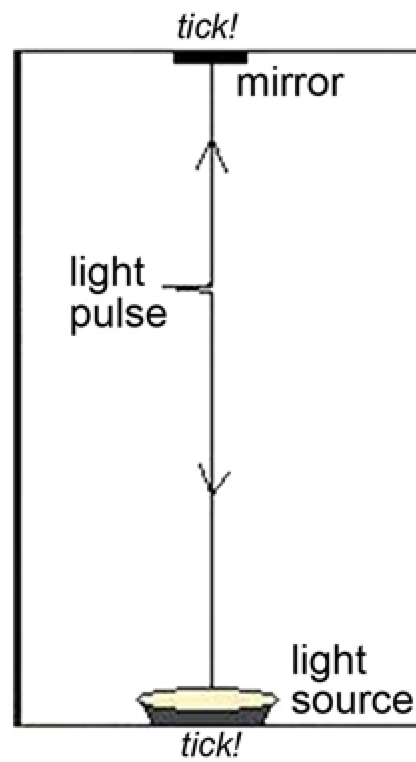

Figure 3.5. Patty's view of her light-clock on the platform. Since she sees her light-clock as stationary, the pulse of light appears to simply bounce straight up and down.

What happens if we look at Patty's clock from Troy's perspective? Well, he is in uniform motion, so he considers himself stationary while the platform recedes away from him. As the platform recedes, Patty's light appears to travel along a diagonal path. Troy sees the time interval between ticks of Patty's clock to be longer than the interval between ticks of his own clock. He sees Patty's clock running more slowly than his own, contrary to Patty's belief (see figure 3.6). Since their motion is relative, their time is also relative [2].

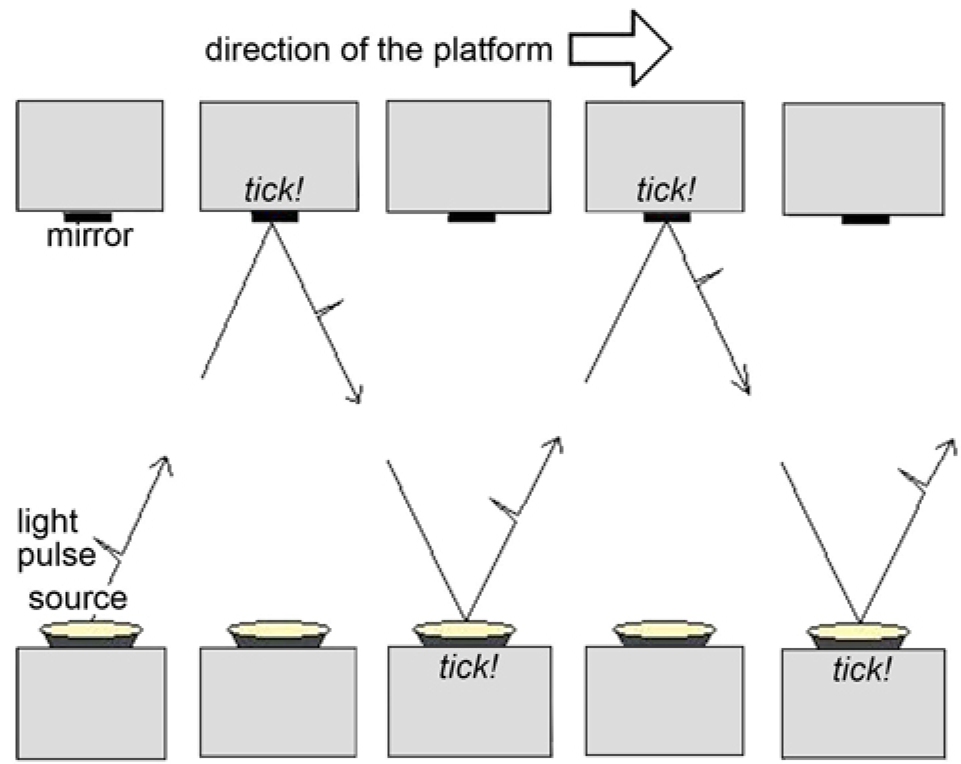

Figure 3.6. Troy's view of the light-clock on the platform. It is no coincidence that Troy observes Patty's light-clock the same way that Patty observes Troy's light-clock. Each of them thinks that the other is the one in motion. That is the essence of special relativity.

When Troy or Patty sees the other person's clock running slow, it's not just some illusion with the clock. It's an observation about *time itself*. When Patty raises her arm to wave to Troy, Troy sees her arm move in slow motion. Similarly, from Patty's perspective, everything in the train seems to move in slow motion (except for light, of course!). This slowing of time due to relative motion is known as *time dilation*.

You might be wondering what happens when Troy gets off the train. If Patty says Troy's clock runs slow, and Troy says Patty's clock runs slow, what is the final result? In other words, if they were exactly the same age when Troy left, then when he finally gets off at his destination, who is older? Well, to get off his train, the train has to slow down. Troy can no longer consider himself stationary because he can feel that the train is in motion as it slows down. *The act of decelerating makes the train a non-inertial reference frame.* When he finally comes to a stop, his clock has ticked fewer times than Patty's did for the duration of the trip. So Troy is now younger than Patty. While Troy was traveling at constant velocity, his perspective was symmetrical to Patty's. But the act of decelerating breaks that symmetry. It allows us to take Patty's perspective as the stationary frame of reference [1].

You might also be wondering why time dilation does not influence our lives on a day-to-day basis. Based on these principles of special relativity, I would expect that whenever I take a train to visit my brother in New York City, he would complain that I always show up late. He would say that my clock runs slow compared to his. But the slowing of time for me in motion, as seen from the perspective of my stationary brother, is insignificant unless I am traveling hundreds of thousands of miles at a speed much faster than any non-crazy train can travel. When we move about at everyday speeds, from a crawling pace (0.7 miles per hour) to jet plane speeds (600 miles per hour), the slowing of time is nearly immeasurable.

However, time dilation does affect some of the modern technology we use every day. For example, a Global Positioning System (GPS), often used in navigating cars and planes, would not be accurate if the effects of relativity were ignored. The 24 satellites that are used in a GPS system move at speeds of around 8700 miles per hour (about 14 000 kilometers per hour) as they orbit Earth. The relativistic effects become more significant at such high speeds. Your GPS is a receiver of radio signals from these satellites. The receiver uses the speeds of the signals and the time that the signals were emitted from the satellites to calculate your location in relation to those satellites. When the effects of time dilation are accounted for in this calculation, a GPS system can show your location to an accuracy of about 10 meters. Without accounting for time dilation, the calculation would be so erroneous that the GPS would be useless [3].

Space: also relative

We've discussed how time itself changes in order for the speed of light to be the same for all observers. The slowing of time for objects in motion, as seen by an external observer, is called *time dilation*. It turns out that space also changes for objects in motion. From an outsider's perspective, the length of a moving object actually shrinks along the direction that the object is moving. This phenomenon is known as *length contraction*. In a sense, *space is contracted along the direction of motion*.

Here's an example of why that happens. This time, Troy sets up his light source so that the pulse of light travels horizontally through a train car instead of vertically. He places the light source on the interior back wall of the train and shines it toward a mirror on the front wall. Patty on the platform will measure how much time it takes a pulse of light to travel one round trip, from the light source to the mirror and back. She will then use the speed of light, the speed of the train, and the time she measures to calculate the length of the train. Patty will take care of the math for us. We want to see how the length of the train compares when Troy measures it versus when Patty measures it. Patty will determine the length using her calculation, while Troy will simply measure the length using a tape measure across the floor.

Troy begins the experiment by turning on his light source. The light pulse does not gain any speed from the motion of the train (remember, the speed of light is always constant, no matter what). While the light pulse is in mid air, the train continues to move forward. The light pulse then has to catch up to the front wall of the train. Therefore it takes a *longer* time for the light pulse to make this trip, from the back to the front of the train, than it would if the train were stationary. Next, the light pulse bounces off the mirror and proceeds toward the back of the train. In this direction, it takes a *shorter* time for the pulse of light to make the trip than if the train were stationary (see figure 3.7).

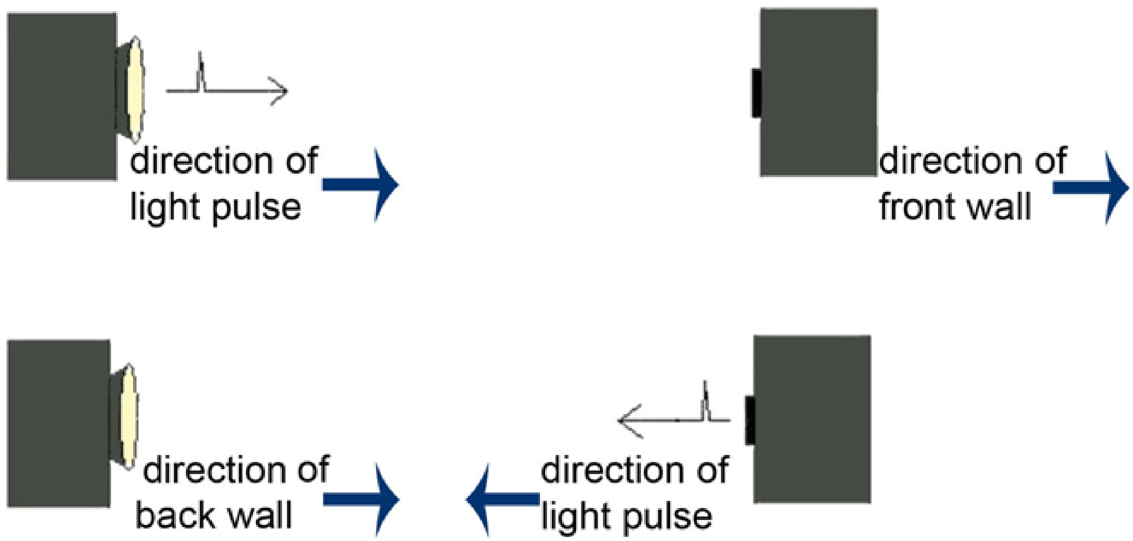

Figure 3.7. When the light pulse travels from the back to the front, it is moving in the same direction as the wall that it is approaching. When it travels from the front to the back, the light pulse and the back wall approach each other head-on. Of course, Troy would make these same observations if this set-up were on the platform.

But even though the trip appears to take longer from back to front and shorter from front to back, the two effects don't exactly cancel each other out. When the light travels from back to front, it is moving in the same direction as the wall that it is approaching. But when it makes the return trip, the light and the wall that it is approaching are moving *toward* each other (see figure 3.7). So the light reaches the wall in a very short time on the return trip. It reaches the wall in such a short time that it overcompensates for the longer time it took to travel from back to front.

When Patty does the math, she sees that the total amount of time for the round trip is *less* than it would be if the train were stationary. Therefore, when Patty calculates the length of the train, it is also *shorter* than it would be if the train were stationary. Thus the train appears to be shortened along the direction of motion, from Patty's point of view; see figure 3.8 [4].

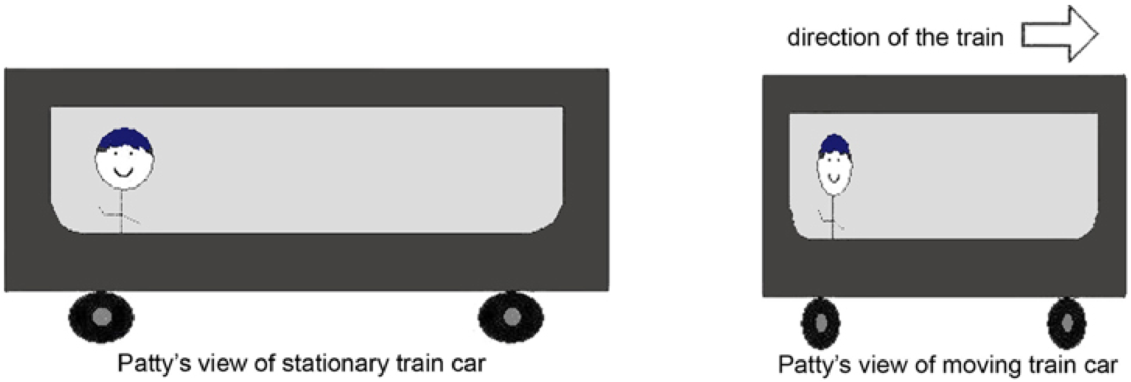

Figure 3.8. Patty's perspective of a stationary train car and a train car in motion. The train car in motion appears shortened (this effect is exaggerated in the image).

Length contraction is relative just like time dilation. Troy considers himself to be stationary while the platform speeds past him. So if Patty had a light on one end of her platform, and Troy made calculations of the length of the platform based on the time it takes a pulse of light to travel across the platform and back, then Troy would say the platform is shortened (see figure 3.9).

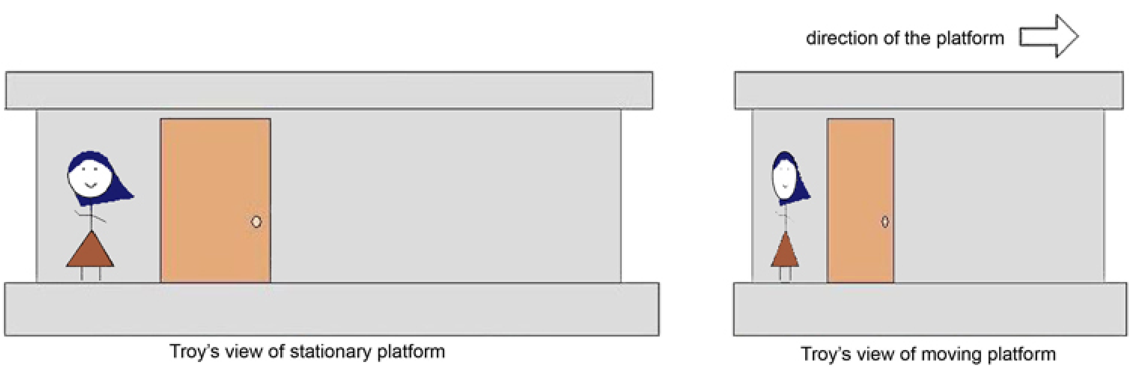

Figure 3.9. Troy's view of the platform when it appears stationary and when it appears in motion. The platform in motion appears shortened (this effect is exaggerated in the image).

Conclusively we can say that, *when an object appears to speed past you, it will be shortened along the direction of motion.* Again, this phenomenon is not noticeable in our everyday lives because the contraction is not significant unless the object in motion is traveling at an incredibly fast speed.

The concepts of time dilation and length contraction have fascinating implications. They are experimental verifications of the relative nature of time and space. If time and space are relative, then all objects hold equally important positions in the universe. In relation to time and space, we are all equally special. Thus we must throw out our concept of absolute time and space as an empty stage for the production of the evolution of the universe. Instead, time and space are relative phenomena, and they depend on the perspective of the observer.

We have discussed how an outside observer sees time slow down for a person in motion. We also mentioned how this is a relative effect, so that the person in motion could claim that he is stationary while the world around him slows down. The faster you move through space, the slower time appears to pass in the world around you. You might think, then, that you could stop time by traveling at the speed of light. Or maybe you'd assume you could even travel *backward* in time if you traveled *faster* than the speed of light. We could discuss why you can't travel into the past form a philosophical point of view, but physically you'll never be able to do it anyway! That's because *it is impossible for anything in the universe to travel faster than the speed of light.* Einstein provides an explanation for this, which he abbreviated into his famous simple equation $E = mc^2$.

Einstein's equation $E = mc^2$: what does it mean?

In Einstein's famous equation, E represents energy, m represents the 'rest mass' of an object, and c^2 represents the speed of light squared. Rest mass is the mass of an object when it is not moving, i.e. when it is in its inertial reference frame, 'at rest'. The equation indicates that there is a direct relationship between the rest mass of an object and the amount of energy you can get out of it. Essentially, it states that energy is equivalent to rest mass times a constant, and the constant happens to be the speed of light squared. So the greater amount of mass an object possesses, the more energy you can get out of it.

While objects possess rest mass, they also gain a 'relativistic mass' when they are in motion. This relativistic mass is the result of the kinetic energy associated with the moving object. Kinetic energy is simply the energy that the object has since it is in motion. As an object's speed increases, its kinetic energy increases. Einstein's original equation ($E = mc^2$) is slightly modified for objects in motion, but the relationship is similar in that greater kinetic energy means greater relativistic mass. So as kinetic energy increases to 'relativistic' speeds (i.e. speeds comparable to the speed of light), the object gains relativistic mass. At some point, the object could not gain any more kinetic energy because it would be too massive to move! There is a precise speed at which the kinetic energy can no longer be increased due to the massiveness of the object. That speed is the speed of light. This is why no object can move faster than the speed of light. Light has zero rest mass, so it can travel at the fastest possible speed: 671 million miles an hour. But nothing in the universe can travel faster than this speed [2].

Nature leaves us with a lot of questions. Why does it turn out that this particular number, 671 million miles per hour, is the speed limit for all objects in the universe? Why is light massless while other things have mass? There are some things we just have to take for granted. Science can provide us with answers to *how* the universe operates, but questions concerning *why* nature is the way it is often leave us wondering.

References

[1] Barr S M 2010 *Informational Meetings* (Newark, DE: University of Delaware, Department of Physics and Astronomy)

[2] Greene B 2003 *The Elegant University* (New York: Vintage Books)

[3] Will C M 2010 *Einstein's Relativity and Everyday Life.* [Online] Available at: http://physicscentral.com/explore/writers/will.cfm

[4] Griffiths D J 1999 *Introduction to Electrodynamics* 3rd edn (Upper Saddle River, NJ: Pearson Education)

[1]Contrary to popular belief, the person who operates the train is called the engineer, not the conductor. Don't ever call the engineer the conductor—especially if the engineer is your uncle Tony!

[2]If I say 'your clock is running slow', I mean that less time has passed on your clock as compared to my clock. So if we both start our clocks at 1 pm, and later my clock says 1:10 pm but yours only says 1:08 pm, then your clock is 'running slow' compared to mine. I've avoided using numbers in the train example to emphasize the concepts alone.

IOP Concise Physics

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Chapter 4

But that’s not all! … space and time in Einstein’s general theory of relativity

The preceding chapter may have sounded a lot like science fiction. We are not used to thinking of space and time as things that can change depending on a person’s state of motion. The nature of space and time, as described by Einstein’s *general* theory of relativity, is equally surprising.

We typically think of time as something uniform which passes by from day to day, and we imagine space as an emptiness which we fill with people and things. We would not expect that filling time and space with objects would somehow alter time and space themselves, because we never see that type of effect in our day-to-day lives. Well Einstein once again proved that there is so much more going on than we can perceive. His discoveries regarding the nature of time and space as a continuous fabric came about with his investigation of gravity.

Einstein was trying to fit the gravitational force within the framework of special relativity. He was troubled by Newton’s theory of gravity because it required that the gravitational attraction between two bodies must act *instantaneously*. That means, if Jupiter somehow flew out of its orbit around the sun and left the solar system, we would feel our gravitational attraction to Jupiter instantly weaken. But that would violate Einstein’s special theory of relativity, which claims that nothing can travel faster than the speed of light. The gravitational force would be traveling faster than the speed of light if its change were to be felt instantaneously. Confident in his own principle that nothing can travel faster than light, Einstein set out to revise Newton’s theory of gravity.

Einstein’s equivalence principle

In Newton's theory, there are two types of mass, known as inertial mass and gravitational mass. An object's inertial mass determines what its acceleration will be when an external force is applied to it. An object with large inertial mass is greatly resistant to an applied force. Objects with large inertial mass are more difficult to accelerate than less massive objects. Gravitational mass, on the other hand, is the mass that determines how strongly an object will gravitationally 'pull' on other objects. Something with large gravitational mass will pull strongly on other objects, e.g. the sun pulls strongly on the planets. In Newton's theory, the equations work out such that inertial mass is equal to gravitational mass. However, Newton did not have a physical explanation for why this was true. In his theory it just appears to be a coincidence.

Einstein considered the equivalence of inertial and gravitational mass in a thought experiment. He first considered what would happen if he were in an elevator somewhere deep in outer space, far from the gravitational attraction of other bodies. If the elevator were to accelerate upward, his inertial mass would resist this change. So he would be pressed against the floor of the elevator due to his inertial mass. Einstein compared this effect to his experience in an elevator on Earth. Even if the elevator did not move, he would be held to the floor due to the gravitational attraction of the earth. That attraction is actually a mutual attraction which results from their gravitational masses. We say that he is in the 'gravitational field' of Earth when he is held to the floor of the elevator and the elevator is not moving. Einstein declared that the laws of physics in these two situations are the same. The person's experiences in the accelerating elevator far off in space are physically equivalent to those in the stationary elevator within Earth's gravitational field. This *physical equivalence of an accelerating reference frame and a reference frame that is stationary within a gravitational field* is known as Einstein's *equivalence principle* (figure 4.1) [1].

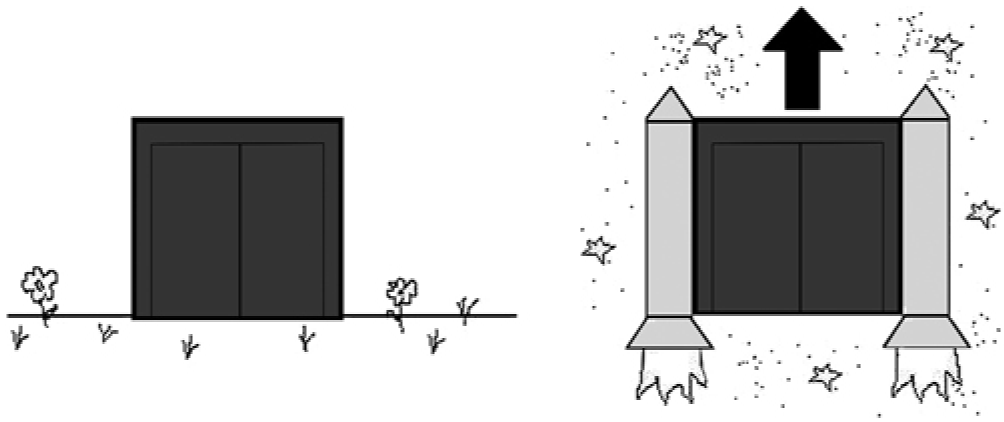

Figure 4.1. The equivalence principle states that the laws of physics are the same in these two situations. In the first, an elevator is at rest on Earth; in the second, an elevator is accelerating somewhere far off in outer space (away from any significant gravitational attraction to other bodies).

Space and time become spacetime

As a result of the equivalence principle, it is clear that fitting gravity into a theory of relativity amounts to incorporating accelerating (non-inertial) reference frames. Incorporating non-inertial reference frames requires new mathematical representations of space and time. After Einstein published his special theory of relativity, a mathematician named Hermann Minkowski came up with a geometry which describes the postulates of special relativity in a new framework. Minkowskian geometry unifies space and time into one four-dimensional manifold known as *spacetime*. When we look at the physical reality implied by this mathematical model, we realize that space and time can no longer be considered two separate things. Anything that affects space also affects time, and vice versa. It is as if space and time are threads woven together into one fabric.

Unifying space and time into a four-dimensional fabric was a step toward incorporating non-inertial reference frames into a complete theory of relativity. However, Minkowski spacetime could not adequately describe the effects we observe in accelerated motion. Minkowski spacetime is a flat geometry. Einstein had to work out his equations with the mathematics of a *curved* spacetime. This type of geometry is known as Riemann geometry, named after the mathematician George Bernhard Riemann who came up with it. Using curved geometry was not just some abstract mathematical trick for convenience. Rather, Einstein used the mathematics to describe the physical reality of spacetime. By using this type of mathematics to describe accelerated motion, Einstein was asserting that *spacetime itself must be able to curve and warp* [2]!

Spacetime curvature and gravity

As stated with Einstein's equivalence principle, experiencing accelerated motion is equivalent to sitting still within a gravitational field. After formulating this principle, Einstein realized that his understanding of accelerated motion could be applied to gravitation. If accelerated motion involves a warping of spacetime, then so does gravity. There must be some intimate relationship between gravity and the curvature of spacetime. There must also be an intimate relationship between gravity and massive objects, since all objects with mass exert the force of gravity on other objects. Einstein put together these concepts to redefine what we experience as the force of gravity. His revolutionary conclusion was that *massive bodies cause spacetime to curve, and it is this warping of spacetime which causes gravitational attraction.*

Typically when we refer to massive bodies, we mean extremely large ones. But in this chapter, a massive body is simply any object which possesses mass. Any object which has mass experiences gravitational attraction. That's because any object with mass causes spacetime to curve.

Let's examine more closely the relationship between gravity and the curvature of spacetime. Massive objects cause spacetime to curve in such a way that anything in the close vicinity of the object is inclined to fall freely toward the object's center of mass. For instance, if the ground were not holding us up, we would fall freely toward the center of Earth. We would each be following a 'geodesic' through spacetime toward Earth's center, with nothing stopping us. Our tendency to follow these geodesics results in our experience of the so-called ‘gravitational attraction’ to Earth. So to understand gravity, we really must understand geodesics in spacetime.

Straight lines in spacetime: not so straight

What is a geodesic? A geodesic is the ‘straightest’ path that can be drawn on a surface, in a space, or in spacetime. On the surface of a beach ball, for instance, a geodesic path between a point on the very bottom and a point on the very top goes along one of the seams of the beach ball. On a typical red and white beach ball, it would be a line drawn along a seam where a red section meets a white section (see figure 4.2). This line may not appear straight because it is drawn on a curved surface. But in several ways it is the straightest possible path between the two points. First of all, it is straightest in the sense that it is the *shortest* path between the two points on the ball’s surface. Second, it is a path that keeps going in the same direction on the ball’s surface (in this case, upward) without swerving to either side.

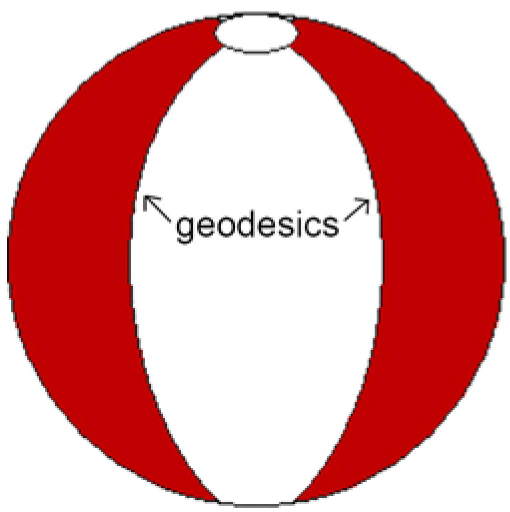

Figure 4.2. Two geodesics on a beach ball, from the bottom of the ball to the top of the ball, lie along the seams.

In this example, the curved surface of a beach ball may be taken as an analogy for the curvature of four-dimensional spacetime. We see how, on the surface of the beach ball, the straightest possible path may not appear straight because of the curvature of the ball. In general relativity, spacetime is similarly curved, so we talk about geodesics instead of straight lines. We say that an object moving freely moves along a geodesic in spacetime.

It is difficult to imagine what a geodesic in spacetime would look like, because spacetime is four-dimensional and curved. We aren't able to see the fourth dimension (time). A geodesic in spacetime is certainly different from a straight line through space alone. In fact, from our perspective, two objects that are moving along geodesics of four-dimensional spacetime sometimes appear to swerve towards each other as if they gravitate towards each other. Indeed, this is how gravitation is explained in general relativity: it is caused by the curvature of spacetime.

Spacetime curvature and the orbits of the planets

The shape of geodesics through spacetime depends on how spacetime is curved. In Einstein's theory, objects with mass change the curvature of spacetime, so they also change the paths of the geodesics. Let's take the orbits of the planets for example. The sun causes the spacetime in its vicinity to curve or warp, which changes the route of a planet's geodesic. When the solar system formed, the rock that would become Earth was flying through space past a star that we would later call the sun. As Earth passed the sun, its geodesic curved around the sun. Earth continues to follow this geodesic today. Earth's initial velocity keeps it moving in orbit and prevents it from simply falling toward the sun.

This all sure sounds different from what we learned in elementary school about gravity and the solar system! General relativity is not typically taught because Newtonian physics is assumed to be easier to grasp. Also, if we make approximations, in many cases Newtonian physics gives nearly the same results as general relativity. The force of gravity in Newtonian physics is stronger for more massive objects, and it weakens exponentially as you move farther away from those massive objects. These effects are the same in general relativity. In fact, for large distances, the strength of gravity depends on distance in almost exactly the same way in general relativity as in Newton's theory. But Newtonian physics does not provide physical reasons for why gravity is experienced the way it is. It gives us the results without the causes. General relativity shows that the cause of our gravitational experiences is the curvature of spacetime [3].

Observable effects of spacetime curvature

While Newtonian physics is approximately correct in most circumstances on Earth, its predictions are not accurate for some astronomical phenomena. Predictions made by general relativity are in better agreement with observations. The discrepancy between Newton's predictions and Einstein's predictions becomes significant in certain cases.

One notable prediction of general relativity is that the curvature of spacetime will cause light to follow a bent path around astronomical objects such as stars and planets. Since light going through empty space feels no real forces, it moves 'freely', i.e. follows geodesics in spacetime, according to Einstein's theory. These geodesics would appear bent due to the curvature of spacetime, which in turn is due to the sun's mass. Einstein predicted that the bending of light would cause stars behind the sun to appear to be in different positions than they actually exist. His hypothesis was confirmed during a solar eclipse in 1919, seven years after he published the general theory of relativity. Newtonian physics would not have been able to explain this observation.

The bending of light results in an effect known as gravitational lensing. When we look into outer space, the light from a distant galaxy must bend around large objects in the foreground of our line of sight. The background galaxy then appears distorted, similar to how an object would appear distorted when you look at it through a wineglass. The galaxy may appear in different shapes and in more than one place at once. Figure 4.3 displays a photograph of this effect.

Figure 4.3. The bright cluster in the center of the photograph is a galaxy cluster in the foreground. Many of the oval-shaped objects surrounding it are distorted images of one galaxy in the background. This background galaxy appears to be in multiple places at once due to gravitational lensing. Credit: NASA, ESA, H Lee & H Ford (Johns Hopkins University); see https://apod.nasa.gov/apod/ap090823.html.

An even more interesting effect of spacetime curvature is another type of time dilation known as *gravitational time dilation*. Similar to time dilation in special relativity, gravitational time dilation is a phenomenon which results in clocks running at different rates. If we place one clock on top of Mount Everest and another clock at sea level, the clock on top of Mount Everest will appear to run *fast* when compared to the clock at sea level. That's because the clock at sea level is closer to the center of Earth. The curvature of spacetime is most dramatic at short distances from the center of an object. Where curvature is more dramatic, the effects of gravity are felt more strongly. So the clock that is at sea level feels a stronger force of gravity than the clock on top of Mount Everest.

We can see how the stronger force of gravity will cause the clock at sea level to run slower than the clock on Mount Everest if we consider a new type of light-clock. If we shine a flashlight straight downward from a cliff on the mountain, the light wave that it emits will be affected by gravity. Let's say that each time the light wave wiggles left or right, we count this as one 'tick' of our light-clock. While the speed of light is not affected by gravity, the frequency of the light wave changes. The emitted light wave is stretched in such a way that it will wiggle left and right less frequently as it approaches the surface of the earth (see figure 4.4). That means it ticks less often as it approaches the ground. So, closer to the center of Earth, time runs slower. This is a measurable effect. In fact, GPS units must account for gravitational time dilation in addition to the time dilation of special relativity discussed in the previous chapter [3].

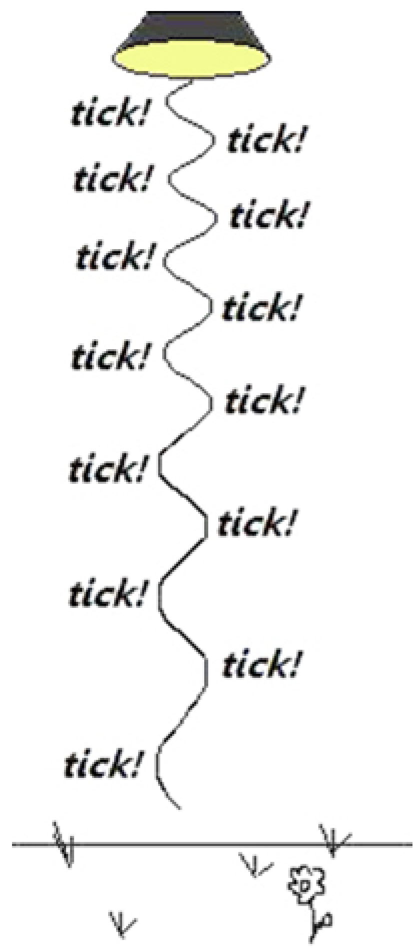

Figure 4.4. As the light wave approaches the surface of Earth, it is stretched due to the affects of gravity. The ticks of the clock become less frequent as the light gets closer to the earth. (In this light-clock, a full beam of light is shined, which is why it appears as a complete wave. In the previous chapter, the light source only emitted a short pulse of light and turned off, rather than emitting an entire beam of light.)

So far we have described light in a few different ways throughout this book. In our chapter on subatomic particles, light took the form of a particle known as a photon. In this chapter, light is described as a wave. So which is it? In the following chapters, we'll see just how peculiarly light behaves—sometimes as a particle, sometimes as a wave.

References

[1] Gardner M 1997 *Relativity Simply Explained* (Mineola, NY: Dover)
[2] Greene B 2003 *The Elegant Universe* (New York: Vintage Books)
[3] Wald R M 1992 *Space, Time, and Gravity* (Chicago, IL: University of Chicago Press)

IOP Concise Physics

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Chapter 5

Light and the concepts of quantum mechanics

Light is like a spork

Before we begin a discussion of light, I'd like to talk a little about sporks. No, I'm not about to introduce some new particle or force or anything called a spork… I really am talking about the plastic spoon–fork utensil that you often see at fast food restaurants. It's rounded like a spoon, but it has prongs on the end of it so it can also be used as a fork (see figure 5.1). Now, imagine you're eating some soup with your spork and a child comes over and asks you, 'Is that a spoon or a fork?'. You look at him with a smile and say, 'Well… it's both. It can be used as a spoon or a fork. It's a spork!'. The child replies, 'But you're eating soup with it. You need a spoon to eat soup, so it's got to be a spoon.'. You say, 'Well sure, I can use it as a spoon now, but when I'm finished with my soup I'll eat my tater tots, and for those I'll use it as a fork.'. The child looks at you, perplexed. 'How can it be a spoon one minute and a fork the next?' he asks. At this point you might be a little annoyed. You tell him, 'It's simply a spoon and a fork at the same time, and I can use it one way or the other depending on what I'm eating.'. The child finally accepts your explanation and runs off to dive into the ball pit.

Figure 5.1. Sketch of a spork, a common plastic utensil.

So what's the point of all this spork business? Well, it just so happens that light is a lot like a spork. Instead of being a spoon–fork combination, light is a wave–particle combination. Physicists like to use the phrase 'wave–particle duality' when discussing this nature of light. In the same way that a spork can sometimes be used as a spoon and sometimes be used as a fork, light can sometimes behave as a wave and sometimes behave as a particle. We may as well call it a 'warticle'.

The wave-like nature of light

Light is most familiar to us in its wave form. A wave of light has two characteristics—its wavelength and its frequency. A shorter wavelength corresponds to a higher frequency; a longer wavelength corresponds to a lower frequency. Figure 5.2 shows these two characteristics of light.

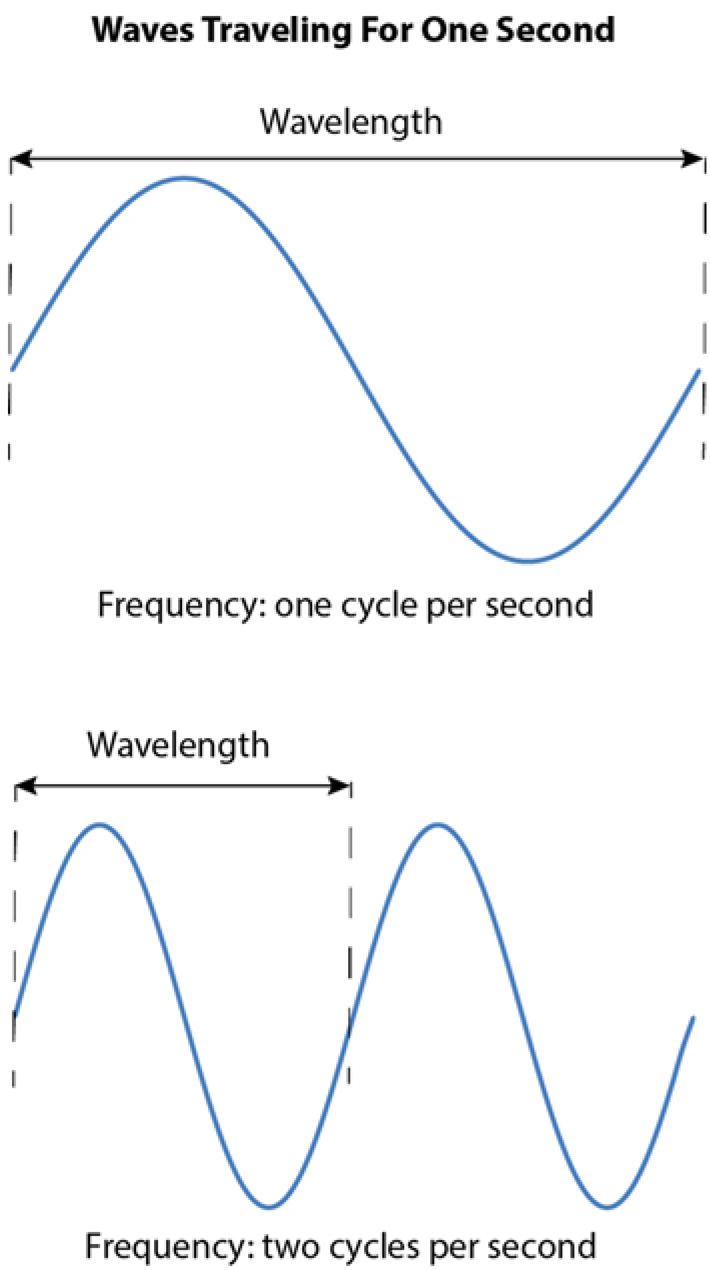

Figure 5.2. Wavelength and frequency of a wave. The wave with a shorter wavelength (below) has a greater number of cycles of the wave in the same amount of time. Therefore, the wave with a shorter wavelength has a higher frequency.

The colors we see in all things around us are determined by the wavelengths or frequencies of the light that are reflected from the surfaces of those objects. Light comes in more forms than just what is visible to our eyes. Radio waves and x-rays are also made of light, but those forms have different wavelengths and frequencies from visible light. Light waves with higher frequencies have higher energies. The spectrum which shows the different forms of light is shown in figure 5.3. This is known as the electromagnetic spectrum.

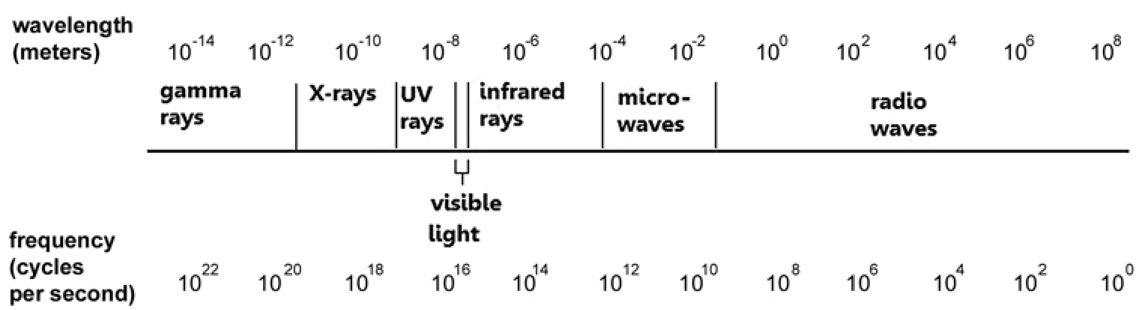

Figure 5.3. The electromagnetic spectrum.

What is a light wave made of? A wave of light is actually two fields—an electric field and a magnetic field—traveling at 90-degree angles with respect to each other. The combination of these two waves is the light wave. In the preceding chapters, when we've said that light travels at one constant speed, we've been talking about the speed given by the wavelength multiplied by the frequency of the light. The speed of a light wave will always be that same constant speed we mentioned in our discussions of relativity: 671 million miles per hour, the fastest speed in the universe.

The first double-slit experiment

The wave nature of light was confirmed in the early 1800s when Thomas Young performed his famous double-slit experiment. Young shined a beam of light toward a wall with two vertical slits in it. The slits allowed the light to pass through as water might pass through two openings between rocks in a river. Behind the wall he placed a screen. On the screen, he observed bands of light and dark areas, indicating that waves of light were interfering with one another, similar to the way that ripples of water might interfere with one another. The lighter bands indicated areas where the waves of light overlapped such that their peaks matched up, causing them to reinforce one another. The darker bands occurred in between the lighter bands, indicating places where the waves overlapped such that they canceled each other out (figure 5.4).

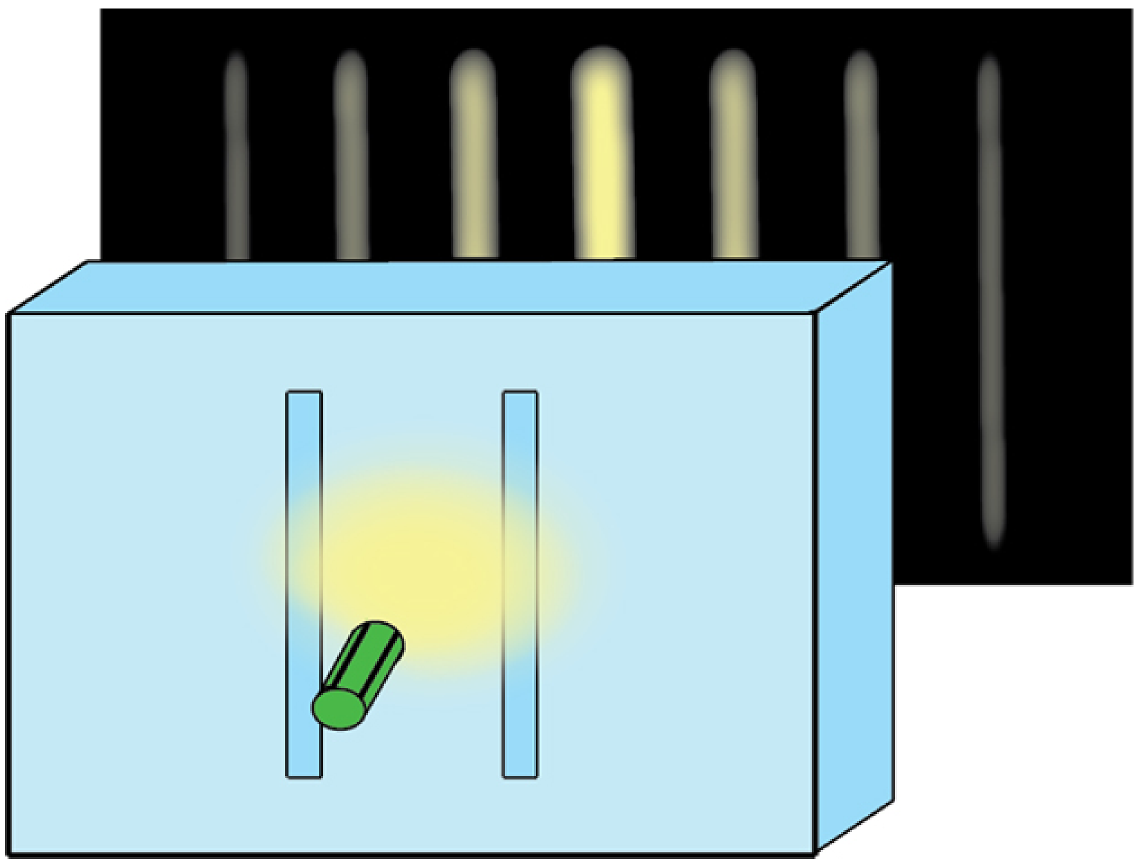

Figure 5.4. The original double-slit experiment. Concentrated light such as a laser beam shines toward two slits. The waves of light pass through the slits and interfere with one another on the other side, creating the pattern shown on the screen.

The photoelectric effect

The particle nature of light was not confirmed until a century later when Albert Einstein explained the *photoelectric effect*. When light shines on a metallic surface, electrons are knocked off that surface. If a brighter light of the same color is shined on the surface, more electrons are emitted from the surface. However, the energy of each electron does not change. Einstein's explanation was that the energy of the light is 'quantized' into packets, and those energy packets of light are what we call photons. The brighter light consists of more photons than the dim light. You cannot have half of a photon or one third of a photon; they are not divisible. That is the essence of quantization: the object which is quantized can only come in discrete values. Each photon can knock off one single electron. So increasing the brightness of your light will increase the number of photons being fired at the metallic surface, which will increase the number of electrons emitted.

The moral of the story is, the energy of light is quantized rather than continuous. Each photon of light behaves like a particle, a ball of energy knocking an electron from a surface. Interestingly the energy of each photon depends on the frequency of the light, which is one of its wave-like properties. To increase the energy of each electron that is emitted, you would have to increase the frequency of the light you're using. Light really is both a particle and a wave at the same time.

Wave–particle duality and the double-slit experiment

We see even more evidence of light being both particle and wave when we repeat the double-slit experiment with different modifications. First, we shine light toward the two slits by emitting only one photon at a time (see figure 5.5). We expect an individual photon to either hit the wall or pass through one of the slits. There would be no way for the photons to interfere with one another if we fire them one at a time, right? So we expect the pattern on the screen to be just two bands of light directly behind the two slits. To our surprise, *the light still makes an interference pattern on the screen!*

So how is it that we end up with an interference pattern? Each photon must somehow be interfering with itself! How can a particle interfere with itself? This seeming absurdity is the heart of quantum physics. It's not just that a whole light beam has wave-like properties. It is also that *each individual photon behaves as both a particle and a wave!*

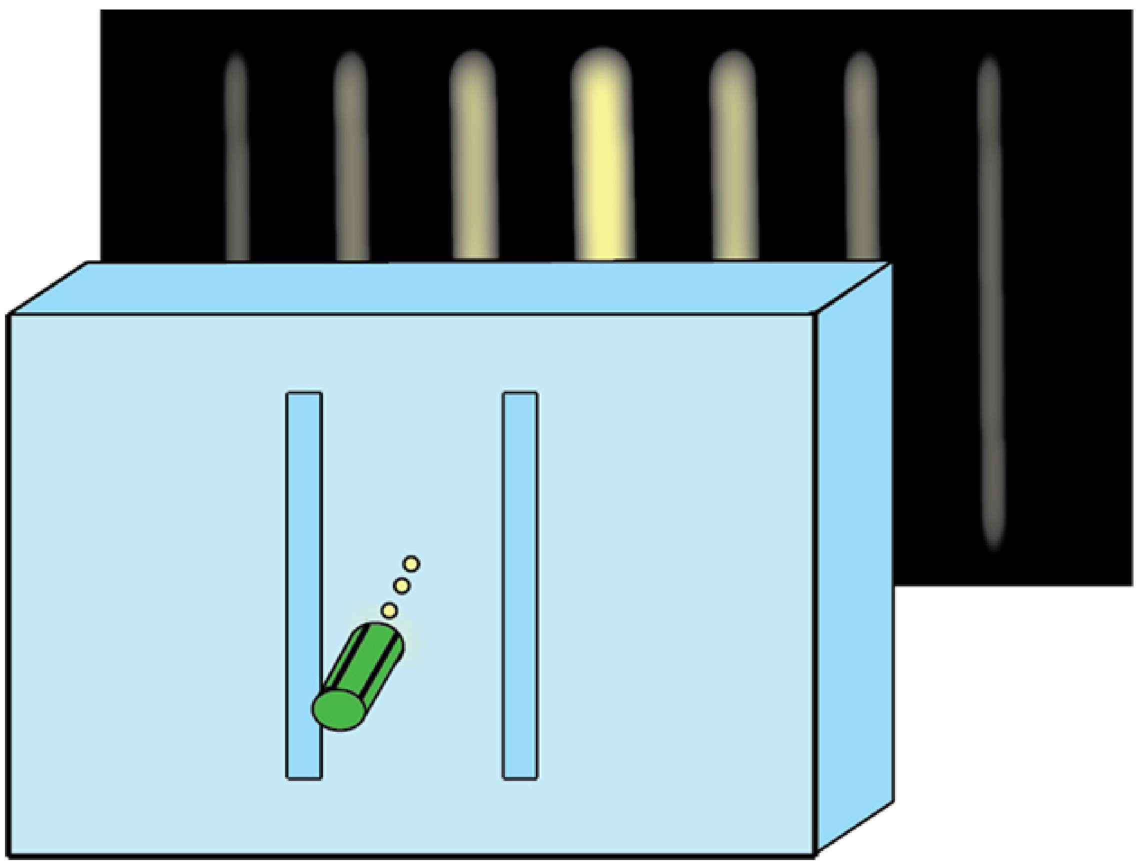

Figure 5.5. Double-slit experiment with photons fired one at a time.

Things get even weirder still. If we replace the photons with electrons instead, firing them one at a time, we still see an interference pattern! But electrons are particles with mass. How are they behaving like waves? It turns out that light is not the only thing with both wave-like and particle-like properties. In fact, all particles with momentum have a wave-like nature. For a particle that has mass, the wavelength associated with the particle is called the de Broglie wavelength. It was postulated by French physicist Louis de Broglie that all particles exist as a wave–particle duality, and the wavelength for their wave nature is named after him. The de Broglie wavelength is given by a constant (called Planck's constant, h) divided by the momentum of the particle. Planck's constant is a super small number, so the wavelength of most particles is so miniscule that their wave nature is unrecognizable. But for very tiny particles with small masses, the de Broglie wavelength becomes significant in determining the position of the particle [1].

What's going on? Interpretations of the double-slit experiment

In the double-slit experiment, for electrons, how do we obtain an interference pattern on the screen? Electrons are not electromagnetic waves like photons are. There must be some other kind of wave in the picture. There are a few different interpretations of what is happening as an electron moves from the source to the screen, and some of these interpretations rely on a mathematical entity of quantum mechanics called a *wavefunction*.[1] Each electron has its own wavefunction, and *the wavefunction determines the probability of the electron being in one place or another.* With a little math, we can use the wavefunction to get a probability distribution showing the probability of the electron being in each location in spacetime. I imagine this probability distribution as a ghost-like, fuzzy wave, kind of like an aura, around each particle. However, this image is not realistic, because the probability distribution is not something that can actually be seen. The probability distribution is described as a wave of probability, so that the peaks of the wave are the most probable locations where you would find the electron, and the low points of the wave are the least probable locations for the electron to exist (see figure 5.6).

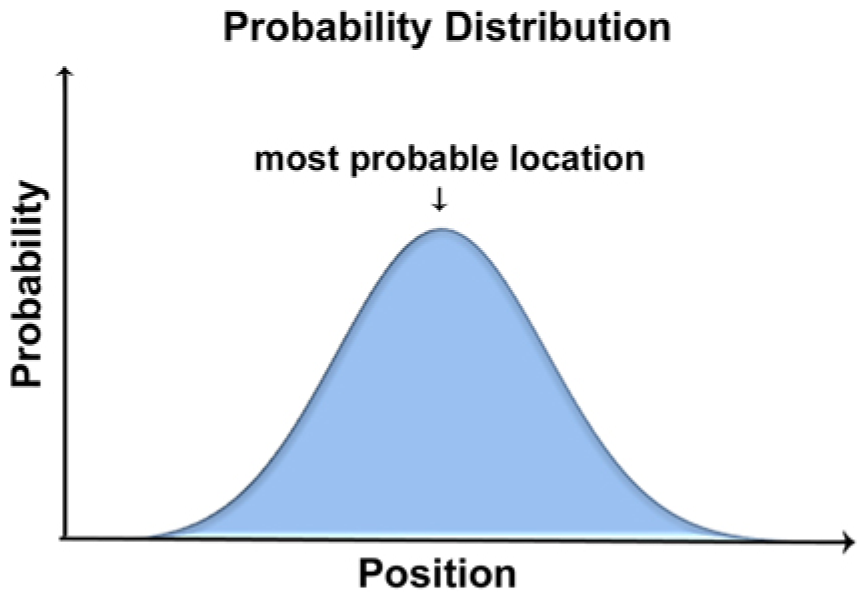

Figure 5.6. An image of a probability distribution, also called a probability wave, obtained from a wavefunction. Reproduced from http://abyss.uoregon.edu/~js/images/wave_function.gif.

de Broglie, along with theoretical physicist Erwin Schrödinger and physicist/mathematician Max Born, gave a well-accepted description of the probability wave and its role in the double-slit experiment. In their view, as each electron approaches the two slits, its probability wave passes through both slits and the ripples interfere the way water waves would. The electron itself does not have any well-defined position as it passes through the slits. The probability wave passing through the slits then determines the possible locations where the electron can land on the screen. The electron finally appears as one distinct point somewhere within the probability distribution. Since the electrons are identical and they are each emitted with the same speed, they have the same probability waves. So each electron takes a position on the screen in accordance with the same probability wave. Electrons are more likely to land where there are peaks of the probability distribution. Over time, as more and more electrons arrive at the screen, the interference pattern emerges. The brighter bands represent locations where electrons are most likely to land, as governed by their probability waves.

Another unique perspective was given by American physicist Richard Feynman. Feynman claimed that each electron actually goes through both slits at the same time. In fact, he claimed that each electron takes *every* possible path from the light source to the screen. For instance, each electron takes a path from the light source, to Argentina, through the left slit, and onto the screen. Each electron also goes from the light source to a distant star, then through the right slit and to the screen. Each electron takes every possible path from the light source to the screen simultaneously. This may be even more difficult to imagine than probability waves! But mathematically, it works. If Feynman takes the summation of all of these possibilities, he gets the same results as the physicists who use the wavefunction approach [1].

No peeking allowed!

So what's actually happening? The answer is we simply don't know, and when we try to find out, reality changes! What I mean is, if we try to use some type of device to observe which slit the electron passes through, the pattern on the screen is different from the usual interference pattern. In fact, we obtain two distinct bands, one directly behind the right slit and one directly behind the left slit (figure 5.7).

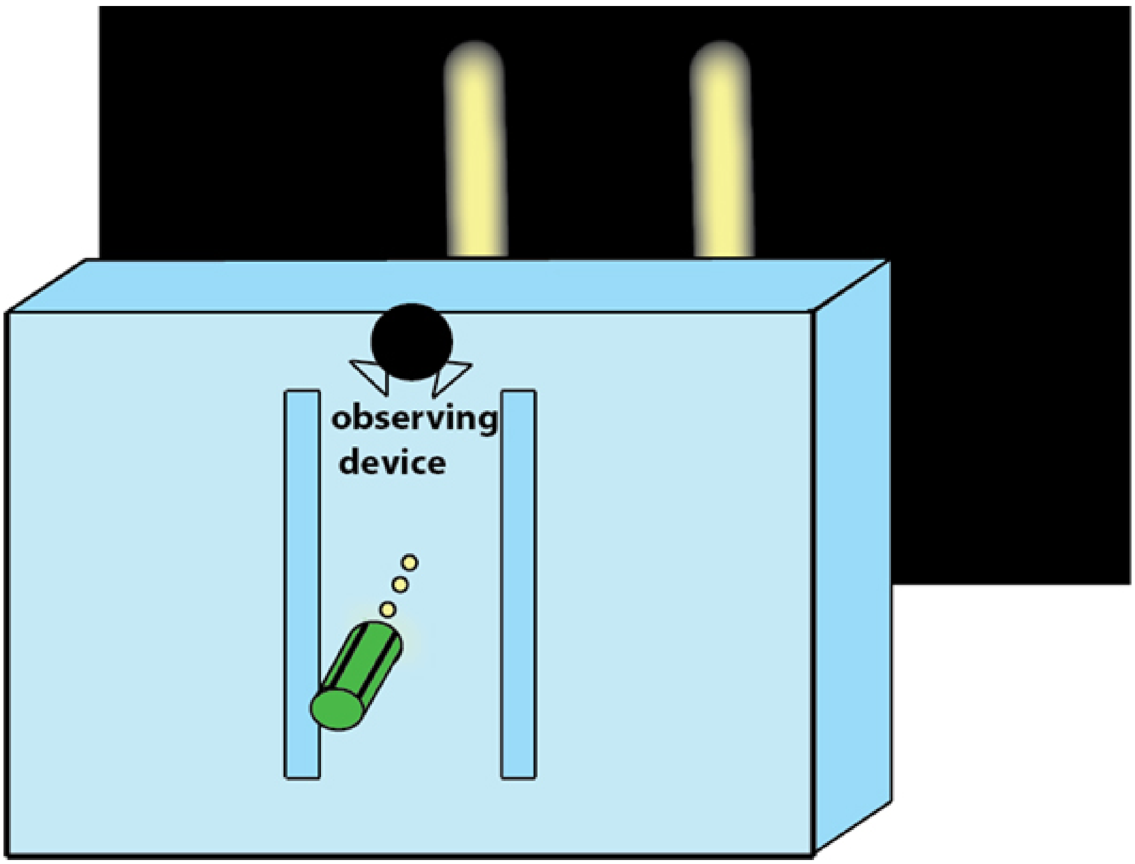

Figure 5.7. Double-slit experiment for electrons with an observing device present. When an observing device detects which slit the electron went through, we no longer see an interference pattern on the screen.

The act of *observing* somehow forces the particle to choose a particular slit. The detector does not observe waves or copies of an electron taking all different paths. It observes the electron either passing through one slit, passing through the other slit, or not passing through either slit (hitting the wall). From de Broglie's point of view, our observation 'collapsed the wavefunction' of the particle. This means that there is no longer a probability distribution and only one possibility remains for the path of the particle. From Feynman's perspective, the particle has to choose one of its many possible paths. In either case, the pattern that results is no longer the original interference pattern obtained when no observation is made. Is that weird, or what!?

Physicists then have another task at hand: to figure out what's happening when we make an observation. What exactly is an observation? In order to observe, say, an electron passing through a slit, we have to shine some type of light on it to detect where the electron is. (This will have to be light of very high frequency, such as gamma rays.) The light is then received by our observation device. But the light interacts with the electron when we make the measurement, so it alters what's physically happening to the electron. Some physicists regard the moment that the light interacts with the electron as the moment of observation. Others believe the moment the light is received by the detector to be the moment of observation. Still others think that there has to be a conscious being interacting with the system for an observation to be made, so the observation isn't made until the human observer looks at his detector and sees which slit the electron went through. In this view, *consciousness* plays a role in the observation [2].

The latter explanation implies that conscious observers are somehow special in nature; they can alter reality! Why would we have that privilege? And how, physically, does that work? It seems as though there would have to be some kind of mental energy that interacts with the experiment at hand. However, no such thing has been quantified in physics yet, and many physicists consider that possibility to be mumbo-jumbo. It is possible that consciousness has nothing to do with the physics of an observation. The jury is still out.

Schrödinger's cat

Let's consider the questions of who is an observer, and at what moment is the observation made. I'll paraphrase a classic thought experiment known as Schrödinger's cat. We take a soundproof box, and in this box we place a radioactive substance along with a kitty cat. Collectively the atoms of the substance have a 50% chance of decaying within a certain amount of time —say, one hour. If they do decay, they will turn into a poisonous gas that will kill the cat. If they don't decay, the cat survives. So the cat has a 50%

chance of surviving and a 50% chance of being killed. The question posed is this: after an hour is up, but before we open the box, is the cat dead or alive?

There are a number of things to consider in this thought experiment. Does the cat count as an observer? Some say that he does, and the wavefunction collapses when he observes the state of the radioactive sample after an hour. Other physicists say that the cat exists in an entangled state of dead and alive until a conscious person opens the box. But the cat is a large-scale object. Its de Broglie wavelength is so small that it shouldn't exhibit a wave-like nature with probabilistic states of aliveness. In other words, it shouldn't behave quantum mechanically like tiny particles do. It must be either dead or alive at any given moment. We wonder then, if the cat dies, is the cat dead at the moment the radioactive sample decays? Or is it only definitely dead once we open the box? We can speculate that it must be dead when the radioactive sample decays because we'll be able to investigate the cat in various ways. For instance, when we finally open the box, we can tell by the cat's temperature whether it has been dead for a while.

Well if we consider the wavefunction of the radioactive sample, when did the wavefunction collapse? Was it when the radioactive sample decayed, or was it when the observer opened the box? Some physicists say that the wavefunction collapsed when the observer opened the box, but the cat really did die when the sample decayed. From this point of view, the collapse of the wavefunction determines the observer's *state of knowledge* about the situation. The atoms of the radioactive sample obey a probability distribution which is determined by the wavefunction, but the wavefunction does not collapse until the observer opens the box [2].

Schrödinger's cat is left open to interpretation, and quantum theories are often judged by how well they deal with this thought experiment. A theory should be able to explain Schrödinger's cat in a logical and mathematically acceptable way. This experiment also illuminates the mysterious nature of quantum mechanics and its relationship to the large-scale world. It brings together a quantum mechanical atomic state with a macroscopic state of a cat's aliveness. This meshing of the physics of the very small with the physics of the very large often puzzles even the brightest scientists. Feynman himself said, 'If you think you understand quantum mechanics, then you don't understand quantum mechanics.' While physicists are able to work through the mathematics of quantum theory and use it to make experimentally verifiable predictions, the reality of what is actually happening is not yet 100% known.

The Heisenberg uncertainty principle

The Heisenberg uncertainty principle (named after the renowned theoretical physicist Werner Heisenberg) illustrates more of the mysteriousness of quantum physics. In quantum mechanics, the properties of a system, such as the momentum of a particle, its intrinsic spin, its location, etc, can only be measured with a finite degree of certainty. For example, the position and momentum of a particle cannot be known simultaneously. *The more you know about the position of a particle, the less you can know about its momentum, and vice versa*.

Here's an analogy to help explain. Let's say you use an old-fashioned camera to take a picture of a runner. First, you set the shutter speed on the camera to be very slow, so that the runner runs by and appears blurry in the picture. With a little math, we could determine the speed of the runner based on how blurry he appears. Qualitatively we know that, if he is very blurry, then he was moving quickly. If he is not so blurry, we know he was moving slowly. But if I asked you where exactly he was, you wouldn't be able to tell me his location. He appears smeared out across the image. Next, you set the shutter speed of the camera to be very quick so that the runner doesn't appear blurry at all. You get a very sharp image of the runner, but now you can't tell how fast he's moving! It's the same way with tiny particles: the faster a particle moves, the less we know about its location; the more we know about its location, the less we know about its momentum.[2] When we know one of these observables with certainty, we can only approximate the other.

Entanglement

Another interesting phenomenon within the quantum world is *entanglement*. An example of entangled particles is two electrons within the same orbital of an atom. Electrons tend to pair themselves so that an electron with spin-up is coupled with an electron with spin-down. (Spin-up and spin-down are the two different states that an electron's 'intrinsic spin' can be. Intrinsic spin is a characteristic of elementary particles which is related to their angular momentum.) Now, let's say we remove the electrons from the atom, but we don't measure their spins yet. So when we remove them, we don't know which one is spin-up and which is spin-down. Next, we send one of the electrons to a far away galaxy and keep the other one by our side. If we measure our nearby electron to be spin-up, we *instantaneously* know that the other electron is spin-down. However, from relativity, we know that no information can travel faster than the speed of light. So how is it that we can know the information about the other electron instantaneously?

The explanation is that the electrons are *entangled.* Their states are both part of the same wavefunction. So when we collapse the wavefunction of the spin-up electron by observing it, we simultaneously collapse the wavefunction of the other electron, since it has the same wavefunction. Any effects that the second electron's spin would have on its surroundings take place the moment the first electron is observed. In quantum mechanics, physicists also speak of entanglement between states of a single particle. For instance, the radioactive material in the Schrödinger's cat experiment exists in an entangled state of decayed and not decayed until an observation is made.

Entanglement is conceptually cool, and it's also practically useful. For instance, scientists have begun developing quantum computers that use entanglement to transmit information instantaneously. Using entanglement in this way allows computers to operate at much faster rates than computers which operate via classical physics.

Quantum physics and gravity at odds

Quantum physics is most accurate when it accounts for the relativistic effects described by special relativity such as time dilation. However, the large-scale effects of spacetime curvature and other predictions of *general* relativity are not compatible with quantum mechanical theories. While both quantum mechanics and general relativity have been used for practical purposes such as computation and GPS tracking, the two theories are not physically attuned. When we try to apply the equations of general relativity to extremely small particles, we obtain nonsensical answers. Similarly, we have yet to come up with equations in quantum field theory which can explain gravitation. While quantum physics can explain the strong, weak, and electromagnetic forces, it is unable to quantify gravity. This is a major problem to most theoretical physicists. The two theories are useful, but they are incomplete. In fact, neither quantum mechanics nor general relativity can explain what happens inside a black hole, which we will investigate in the following chapter [3].

References

[1] Greene B 2003 *The Elegant Universe* (New York: Vintage Books)

[2] Barr S M 2010 *Informational Meetings* (Newark, DE: University of Delaware, Department of Physics and Astronomy)

[3] Pittel S 2010 *Informational Meetings* (Newark, DE: University of Delaware, Department of Physics and Astronomy)

[1]Every particle, or physical system of particles, has an equation associated with it called the Schrödinger equation. The solutions of the Schrödinger equation make up the wavefunction of the particle or system. The wavefunction can be used to find the probability of the particle or particles in a system being in one physical state or another.

[2]Note that there is a difference between speed and momentum. The momentum of a massive particle is its mass times its velocity. When a particle moves quickly, it gains speed and also gains mass, as discussed in the chapter on relativity when we refer to relativistic mass. This photograph analogy leaves out the part about mass.

IOP Concise Physics

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Chapter 6

In the shadows of the cosmos: black holes, dark matter, and dark energy

What is a black hole?

You've probably heard of black holes from sci-fi movies, books, or maybe even a science class. You might have an image in your head of objects being sucked into a black hole, never to return. But what exactly are black holes, and how are they formed? Many people would be surprised to learn that a black hole is actually the remnant left over from a dead star.

How does a star 'die'? It's obviously nothing like human death at all. Stars go through many different events in their lifetimes. They are born from hydrogen in nebular clouds, and through various atomic processes they fuse hydrogen into larger elements such as helium and carbon. Low and intermediate mass stars (less than eight times as massive as the sun) have a different event at the end of their lives than do high mass stars (more than eight solar masses). When the core of a low mass star becomes dense with the elements it produces, its gravitational attraction becomes stronger. The core then begins to contract. While the core contracts, the outer layers expand and cool. Eventually the low mass star sheds its outer layers and the core that is left behind is called a white dwarf.

The dying process for high mass stars is more violent than this. After some processes of expansion and contraction, the star's core contracts and heavier elements are formed from nuclear fusion. Eventually the star's core contracts until it gets so dense that the electrons combine with the protons to make neutrons and neutrinos. The neutrinos escape and the core ends up being entirely made of neutrons. The star's outer layers bounce off of the core in a supernova explosion, blasting elements into the universe. Only the core remains as a super-dense neutron star. If the star is more than 10 times the mass of the sun, this neutron star is too massive for its small volume, and gravitation causes it to collapse to a tiny point. That point is known as a black hole singularity. It is effectively the center of a black hole [1].

A black hole is ‘black’ because its gravitational pull is so strong that even light cannot escape it. This means that the escape velocity of a black hole is greater than the speed of light. Anything which tries to run away from the black hole singularity would have to run faster than the speed of light. Since nothing can travel faster than light, nothing can escape a black hole!

There is a distance from any black hole that is called the event horizon. The event horizon is essentially an invisible edge of a black hole. Inside this edge, any object will be pulled toward the center of the black hole without any hope of coming back. The human body certainly couldn’t survive this experience. If you were to fall head-first toward the center of a black hole, the gravitational pull on your head would be stronger than the pull at your feet, and your body would be stretched until it could no longer hold itself together. But perhaps some type of subatomic particle would be able to survive the journey toward the center. Where would it go?

Scientists really don’t know the answer to this question. Often times, when scientists are unsure of what happens physically, they can work with mathematical equations and see what the results will be. They can then see what is physically implied by the mathematics since the mathematics deals with variables that represent physical quantities. But for black holes, we have problems. The equations of general relativity and the equations of quantum mechanics both break down very near the center of a black hole. That means both theories give us results which are mathematically impossible and physically meaningless. This has been a major problem for modern theoretical physicists, and it continues to be a subject of interest today.

Dark matter

Black holes aren’t the only invisible things leaving physicists puzzled. There is also a substance permeating space all around us known as *dark matter*. Current estimates predict that dark matter makes up about 23% of the mass in the universe, yet we don't know what it is! We believe that it exists based on its gravitational effects.

Several observations have been made which suggest the presence of dark matter. Stars in spiral galaxies appeared to be moving at velocities high enough to break free from the gravitational attraction to the other observable objects in their galaxies, yet for some reason the stars don't fly away from their galaxies. This indicates that some unseen type of matter or other force must be the source of the extra gravitation keeping stars bound to the galaxies. Galaxies within clusters would also require more mass than we can detect to remain as they have within their systems. Additionally, models of the universe based on modern inflationary cosmology suggest that there should be more mass in the universe than we actually observe.

The Bullet Cluster, examined in August, 2006 with NASA's Chandra x-ray Observatory, gives one direct piece of evidence of dark matter (figure 6.1). When two subclusters of galaxies collided, the stars in those subclusters passed between one another with ease. But the gas particles interfered with one another and got stuck in the middle of the collision. This event separated the stars from the plasma gas. Astronomers had a good idea about the masses of the stars and gas in the galaxy clusters before the collision. They believed that most of the mass was attributed to the plasma gas. After the collision, they had a chance to measure the mass in the star regions and the mass in the gaseous region separately using gravitational lensing techniques. The mass in the star regions was measured to be *greater* than that in the gaseous region, contrary to their expectations. Thus some unseen substance or force must have passed through the collision along with the stars. Astronomers conclude that the substance is dark matter [2].

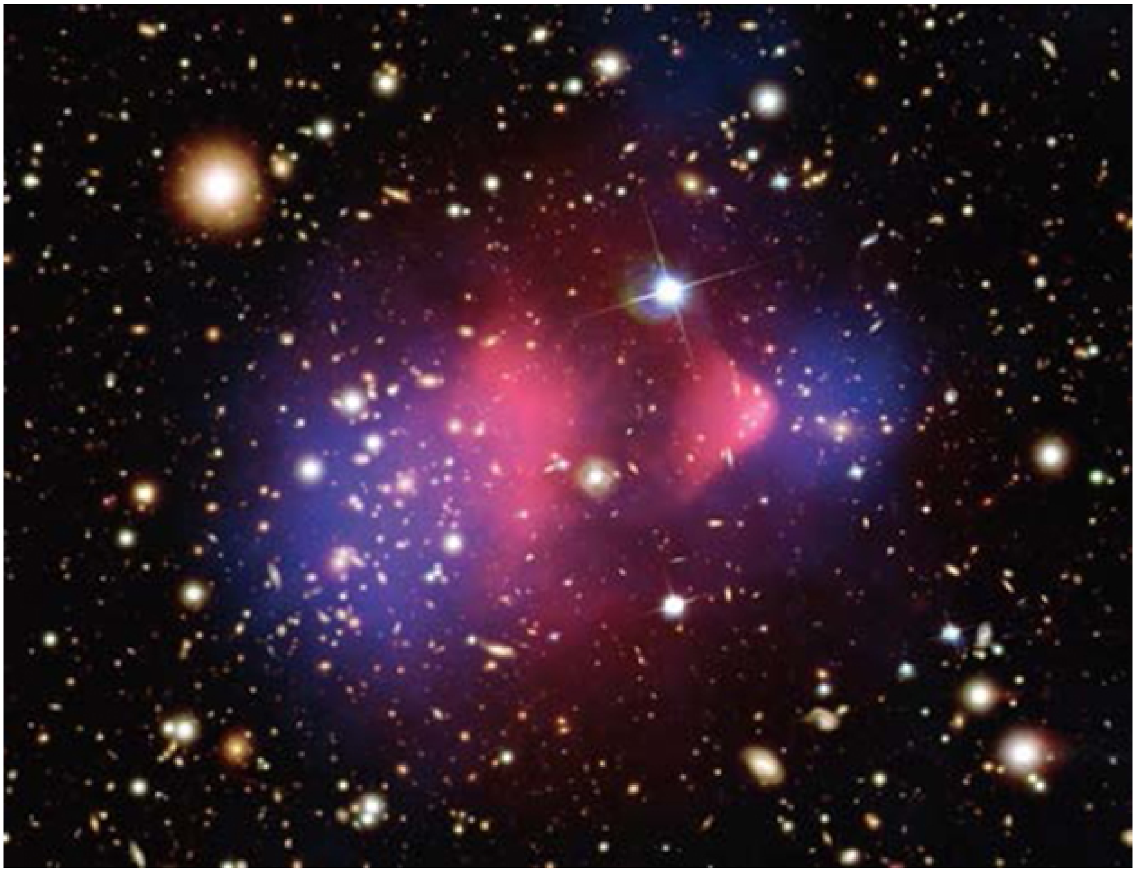

Figure 6.1. The Bullet Cluster after a collision of subgalaxies. The center region, with its characteristic bullet-shaped lump, consists of plasma gas. The regions to the left and right of center contain stars and dark matter. Composite Credit: x-ray: NASA/CXC/CfA/ M Markevitch *et al*; Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/ D Clowe *et al* Optical: NASA/STScI; Magellan/U.Arizona/D Clowe *et al* Reproduced from https://apod.nasa.gov/apod/ap060824.html.

While astronomers have not found a way to detect the particles of dark matter, they have some predictions about the type of particles they may be. Based on cosmological models of the evolution of the universe, dark matter should be non-baryonic, which means it is not made of quarks like protons and neutrons are. It is supposedly made of particles which do not experience the strong force or electromagnetic force. As far as we know, dark matter only interacts with other matter gravitationally and possibly by the weak force. Two candidates for dark matter are axions and weakly interacting massive particles (WIMPs). Theories in particle physics predict that axions were created during the Big Bang, and their existence would explain why the neutron has not been observed to have an electric dipole moment. One type of WIMP is called a neutralino. Neutralinos are predicted to exist according to supersymmetry theories. We will not go into further detail regarding these particles, because neither axions nor neutralinos have been detected [2].

Dark energy

While dark matter makes up 23% of the mass in the universe, about 72% is in the form of *dark energy* (remember, mass and energy are interchangeable). The remaining 5% is made up of the matter that we are able to detect, such as atoms, radiation, and neutrinos. It is pretty incredible that only 5% of the mass in the universe is well known to us. Dark energy remains mysterious because we don't know its source. We know that it exists because the universe is expanding at an increasing rate. If there were no dark energy, the gravitational attraction between all massive objects would cause the expansion of the universe to slow down and possibly even reverse. The universe would collapse upon itself, ending in what is called a 'Big Crunch.' But observations show us that the universe is expanding at an increasing rate rather than slowing down. All galaxies are moving apart faster and faster. Rather than a Big Crunch, physicists expect all objects to continually recede away from one another and dissipate throughout the universe.

The expansion of the universe is actually the expansion of space itself. Distant objects are receding from one another because the space between them is stretching. A common analogy for the expanding universe is the surface of a balloon. While space is three-dimensional, the surface of a balloon is only two-dimensional, so this is quite a simplification. In any case, we'll imagine using a marker to make several dots all over the surface of a balloon. When we blow up the balloon, the fabric of the balloon stretches so that the dots all move away from one another. The expanding fabric of the balloon is analogous to the expanding fabric of space.

Why does dark energy cause the expansion of the universe to speed up rather than slow down? It has to do with the fact that dark energy has a huge negative pressure. Matter, such as ordinary atoms or even dark matter particles, causes objects to move toward each other. We may say that matter results in positive pressure. Dark energy has the opposite effect. It must be some other kind of thing not made up of particles. Scientists have come up with a few explanations, but nothing has been proven yet. One explanation came from Einstein. In his theory of gravity, even empty space (vacuum) can have energy and negative pressure. Such vacuum energy appears as a term in Einstein's equations of gravity, called the 'cosmological constant' term. Other theorists believe that dark energy may be some kind of undiscovered dynamic fluid or field which they named 'quintessence'. Still others believe that perhaps our theory of gravity is incomplete. Perhaps dark energy is some manifestation of gravity that Einstein's general relativity does not account for. If they are right, Einstein's general relativity would have to be revised or even replaced with a completely new theory [3].

References

[1] Goldsmith D 1990 *Supernova: The violent death of a star* (New York: Oxford University Press)

[2] Lochner J 2009 Possibilities for dark matter. [Online] Available at: http://imagine.gsfc.nasa.gov/docs/teachers/galaxies/imagine/dark_matter.html [Accessed 5 April 2010]

[3] Netting R 2010 Dark energy and dark matter. [Online] Available at: http://nasascience.nasa.gov/astrophysics/what-is-dark-energy [Accessed 28 March 2010]

IOP Concise Physics

The Universe Untangled

Modern physics for everyone

Abigail Pillitteri

Chapter 7

A glimpse into the future

The preceding chapter presented some of the mysteries of the universe that remain under investigation. While it is sometimes disconcerting how much we have left to discover, is important to realize that scientists are persistent in their quests for knowledge. Uncertainty is the motivation for discovery. Experimentalists working in laboratories throughout the world are trying to find ways of detecting dark matter and other missing particles. Meanwhile, theoretical physicists are working hard to develop a grand unified theory that incorporates gravity and somehow stitches together the concepts of general relativity and quantum mechanics.

The Large Hadron Collider

One of the machines at work is the most powerful particle accelerator in the world, the Large Hadron Collider (LHC). It is called a hadron collider because its purpose is to collide hadrons, which are particles made of quarks. The collider consists of many magnets and a circular underground tunnel below parts of Switzerland and France. The tunnel of the collider is 17 miles (27 kilometers) around! It needs to be so large in order for the particles to accelerate to unthinkable velocities. Each time a particle completes a loop around the circular tunnel, the magnets act so as to accelerate the particle faster.

Scientists send particles in opposite directions and cause them to collide with enough energy to produce new particles. Those scientists are simulating energies of the early universe when particles were formed. With these high energies at the LHC, physicists were able to produce the Higgs boson. This particle is supposedly the source of the mass of all objects, as we discussed in chapter 2. With this powerful machine, who knows what particles we'll find next [1]!

Gravitational waves

An even more recent discovery was that of gravitational waves. These waves are basically ripples in spacetime that travel at the speed of light. The waves are produced by massive accelerating objects. For instance, when two neutron stars circle each other, they draw each other closer together and accelerate faster and faster. Gravitational waves travel away from the stars, similar to ripples traveling away from a rock tossed into a pond (figure 7.1).

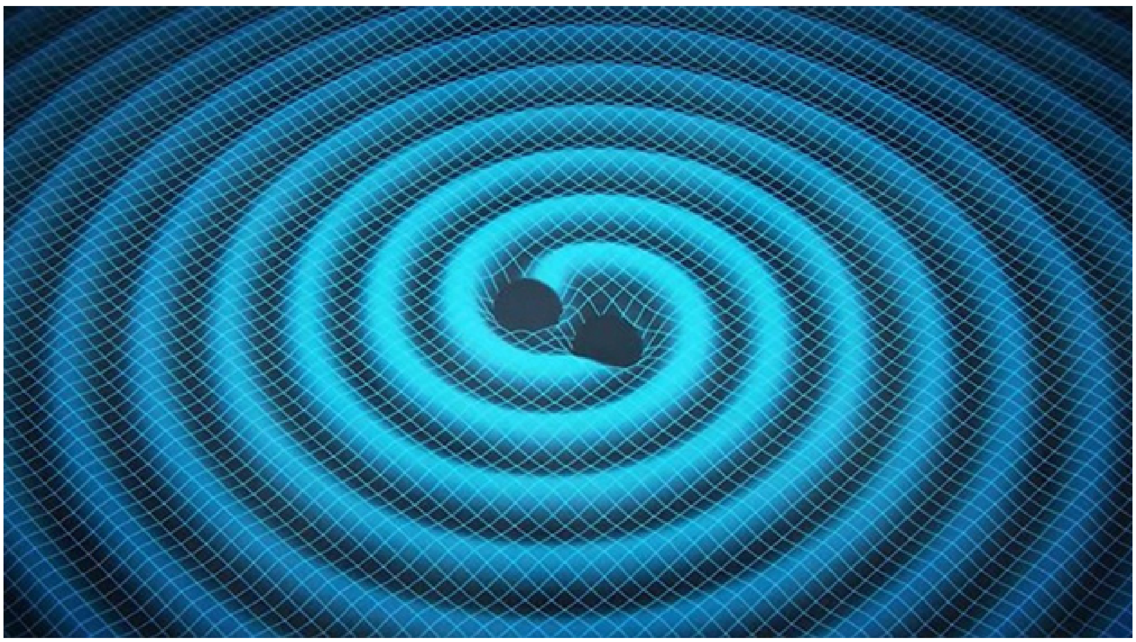

Figure 7.1. An artist's depiction of gravitational waves produced by two stars circling each other. Reproduced from NASA http://nasa.tumblr.com/post/139124558019/what-are-gravitational-waves.

Einstein predicted the existence of gravitational waves in his general theory of relativity, but they were not directly detected until September of 2015. They were detected by two Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors. The gravitational waves came from two colliding black holes far off in the universe. The waves passed by Earth, causing a change in the distance between the mirrors in the detectors. This distance changed because space itself was affected by the gravitational waves! The detectors were able to pick up on this change using lasers. Two laser beams were sent in different directions down identical paths. The beams bounced off mirrors and returned. But one laser light returned later than the other because it was traveling in the direction affected by the passing gravitational wave. Space itself was stretched by the wave, so the light traveling in that direction had to travel a further distance [2].

Scientists at LIGO look forward to detecting more gravitational waves with hopes of making new discoveries. The waves could tell us more about the accuracy of general relativity, the kinds of events that produce gravitational waves, and the nature of gravitational waves themselves.

Theoretical physics and string theory

Outside of laboratories, we have physicists who work on developing new theories to explain the universe. One major contemporary theory in the world of theoretical physics is known as string theory. The basic idea of string theory is that all objects are made up of tiny vibrating strands of energy called strings. These strings would be smaller than any particles that we know of, including quarks. They would be so small that no instrument of current technology would be able to detect them. Each string would vibrate with a particular frequency. The different frequencies of vibration would determine different types of particles. For instance, a red quark would be made of strings that vibrate at a different frequency from the strings inside a green quark [3].

Many theorists believe that string theory looks promising. It might be the theory needed to bring general relativity and quantum mechanics together. However, since the strings are too small to be detected, other physicists think that string theory is merely a philosophy. If it cannot be physically tested, then it cannot be proven. Until it is proven, or at least able to produce some type of evidence for its truth, many physicists will remain skeptical of its legitimacy as a physical theory.

While other theories also exist, none have proven to be a grand unified theory of everything. A grand unified theory has been deemed the Holy

Grail of modern physics. Scientists hope for a complete theory that can describe all of the physical workings of the universe. While some theorists believe we are on the verge of such a discovery, others think it might be unattainable. Perhaps the human mind is just not capable of understanding the entirety of the universe in which we live. Certainly, doubt will not stop the human race on our quest for knowledge. In fact, it will only push us onward.

References

[1] Griggs J 2012 Peter Higgs: Boson discovery like being hit by a wave. [Online] Available at: https://www.newscientist.com/article/dn22033-peter-higgs-boson-discovery-like-being-hit-by-a-wave/?full=true&print=true

[2] LIGO What are gravitational waves? [Online] Available at: https://www.ligo.caltech.edu/page/what-are-gw

[3] Greene B 2003 *The Elegant Universe* (New York: Vintage Books)