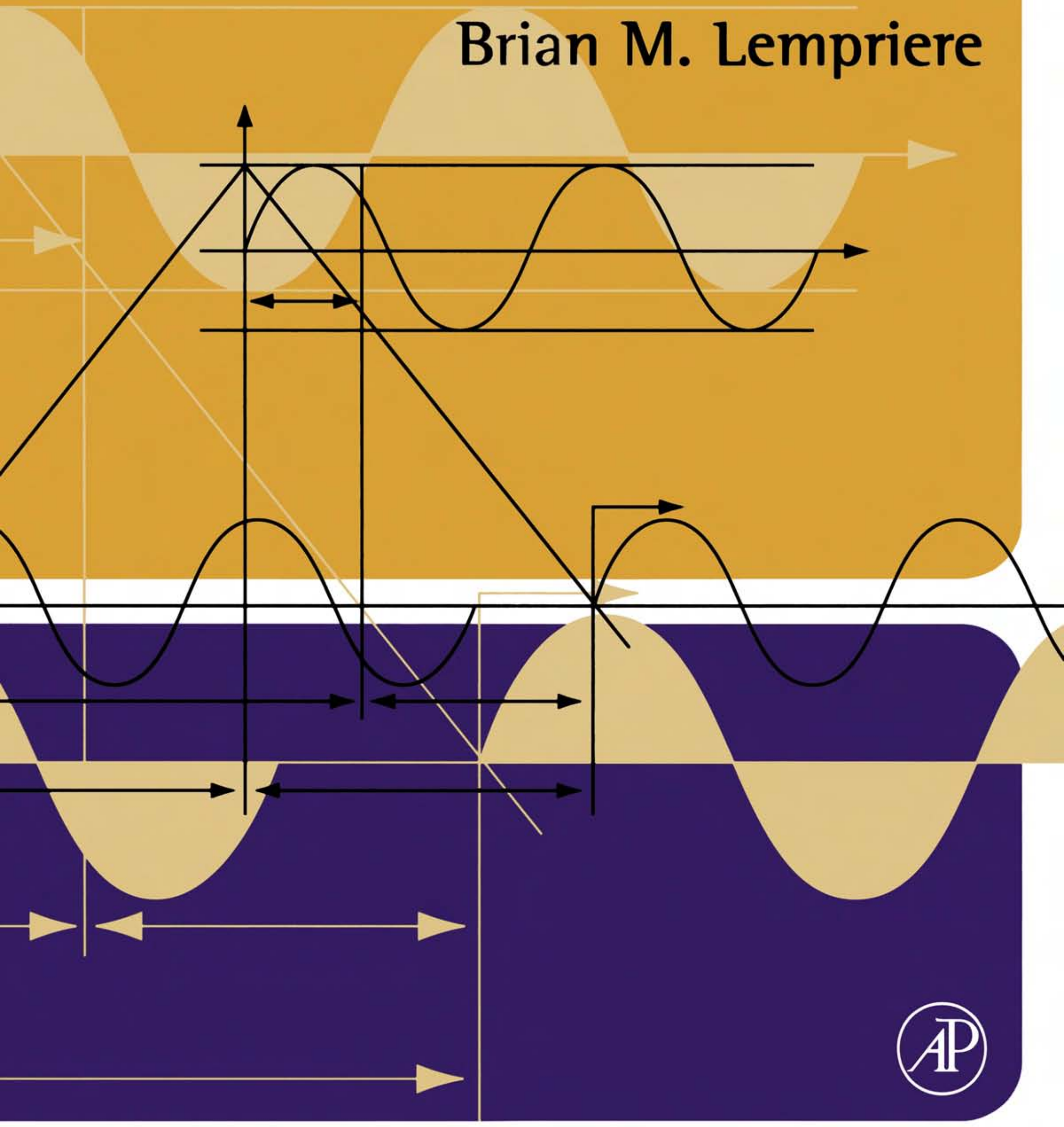
ULTRASOUND AND ELASTIC WAVES

Frequently Asked Questions



Brian M. Lempriere

Ultrasound and Elastic Waves

Frequently Asked Questions

This Page Intentionally Left Blank

ULTRASOUND AND ELASTIC WAVES

Frequently Asked Questions

Brian M. Lempriere

ACADEMIC PRESS

An imprint of Elsevier Science

Amsterdam Boston London New York Oxford Paris San Diego
San Francisco Singapore Sydney Tokyo

This book is printed on acid-free paper. ∞

ACADEMIC PRESS
An imprint of Elsevier Science
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA
http://www.academicpress.com

Academic Press
Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK

Library of Congress Catalog Number: 2001098014

ISBN: 0-12-443345-6

Printed in the United States of America
02 03 04 05 06 MB 9 8 7 6 5 4 3 2 1

I dedicate this book and its author to my wife, Cherie.

To my Readers, I reaffirm the timeless advice:
When all else fails, read the instructions.

This Page Intentionally Left Blank

CONTENTS

3 Wave Propagation Concepts

4 Wave Interactions

5 Hardware: Equipment Concepts

6 Software: Data Processing

APPENDIX 4
Balance of Forces and Newton's Law of Inertia 147

APPENDIX 5
Theory of Wave Propagation 151

APPENDIX 6
Solutions to the Wave Equations 159

APPENDIX 7
Dispersion, Group Velocity 169

APPENDIX 8
Transducer Beam Forming 175

APPENDIX 9
Solutions for Anisotropy 183

APPENDIX 10
Oblique Interactions between Waves and Boundaries 195

APPENDIX 11
Lateral Stress and Strain in Rods under Axial Loads 205

APPENDIX 12
Bending Waves in Beams and Plates 211

APPENDIX 13
Time-Domain Analysis 221

LIST OF FIGURES

SECTION 1

SECTION 2

SECTION 5

APPENDIX 7

APPENDIX 8

APPENDIX 9

APPENDIX 10

APPENDIX 11

APPENDIX 12

APPENDIX 13

PREFACE

WHAT DOES THIS BOOK DO?

This book provides a broad view of ultrasound or ultrasonics, or, in the medical community, sonography, commonly abbreviated as UT. It is a source book of facts and theory, presented in a user-friendly format: it is written, as far as possible, in simple language with intuitive illustrations.

The book is intended for use as an introduction or a reference by anyone with a need to understand what ultrasonics can and cannot do, what are its possibilities and perils, and how to gain a generalized insight. It is not an academic text. Some useful formulas are given and their derivations are given in appendices. These are not required reading, as a knowledge of differential calculus, vectors, and matrix algebra is assumed of the reader for these.

The book brings together the practical and theoretical aspects of UT, which is a field involving art and experience, as well as scientific knowledge. A few historic yet pertinent and current references are included but no extensive literature review is given—that is done in research publications.

Section 1 provides an overall introduction to the subject of UT. Section 2 provides definitions. Section 3 describes wave propagation concepts which are the fundamental processes of UT. Section 4 describes the interactions between types of waves, and between waves and surfaces. Section 5 describes the concepts of transducers and electronic hardware (without reference to specific proprietary information). Section 6 reviews data processing software concepts. Appendices 1–13 provide technical details of relevant theories.

The illustrations are sketched as schematic, each emphasizing a specific concept rather than reproducing actual measurements in which several processes may have occurred, obscuring their significance.

WHY IS THIS BOOK NEEDED?

Ultrasonics deals with propagating waves which depend on two very old principles: elasticity, defined in 1660; and inertia, defined in 1687. A thorough appreciation for these principles is sufficient to gain an understanding of UT. There have been many developments in the intervening years up to the present, though the most significant are quite old. Recent advances have come because of developments in electronic equipment and data processing, and in related fields such as medicine, seismology, acoustics, and engineering dynamics. The UT processes are, however, still often misunderstood.

UT applications began in the late 19th century, but significant progress was not made until the development of equipment for submarine detection in the two world wars of the 20th century. Before the digital age, many techniques used hardware approaches for controlling transducer excitation and for rudimentary signal processing (such as delay lines), together with analog circuitry for signal processing. Test results were displayed on oscilloscopes as analog depictions of waveforms. Most texts of a decade and more ago concentrated on these approaches, which are now obsolete. In these early applications the reflection of a simple wave propagating in a uniform medium was evaluated visually.

Modern applications of UT testing often lead to erroneous conclusions, because the waves exhibit complicated behaviors whose interpretation must be based on a proper understanding of the phenomena involved. This is the most significant aspect of UT at the start of the 21st century.

The field of UT is gaining more and more applications, and questions as to the meaning of test results are raised by people who have not had the benefit of a specialized education or experience in either the generalities, depths, or subtleties of the subject. These people—and the author found himself to be one of them—must ask questions about UT because it is not commonly taught and simple clear textbooks are few. Both practical laboratory experience and theoretical knowledge is necessary. Answers are given to frequently asked questions (FAQs) in the form of questions-and-answers based on theory and experience.

A thorough knowledge of the phenomena also provides the basis for developing new techniques.

WHAT ARE SOME USEFUL BOOKS?

The classical textbook on UT theory is by Krautkramer and Krautkramer (1990), now in its 4th edition, having been released first in the 1940s.

A complete review of UT suitable for a specialist or an advanced graduate student is provided by Ensminger (1998). He includes industrial and medical applications. However, extensive mathematical development is blended in with the text, rendering it hard to read for content. The book provides an interesting history of the subject and discusses natural ultrasonics in bats (echolocation), whales (communication), and dogs (which hear UT whistles), among many other animals and even insects.

Kutruff (1991) also provides a useful overall text for advanced students and researchers, covering most UT concepts and including detailed mathematics. He covers wave propagation in gases, liquids, and solids; the piezoelectric effect and transducers; radiation and beam forming; scattering and absorption; displays; cavitation and high-intensity effects; and descriptions but inadequate explanations of medical applications.

He describes many hardware and signal analysis techniques (such as delay lines, analog filter circuits, and displays) which are essentially out-of-date. The book is a translation from the German edition of 1988 and contains some translated terms which conflict with general English usage, such as the "impulse echo" technique, instead of the "pulse echo" technique.

Of the many research-oriented books, Thurston and Pierce (1990) is one of a series of advanced books on UT. It provides the history of UT and deals with UT devices, wave propagation, measurement techniques, advanced details of the radiation fields of transducers, measurement of velocity and attenuation, electromagnetic transducers (EMATs), optical detection, characteristics of piezoelectric devices, visualization of UT pulses by photoelasticity, etc.

Elastic wave theory is presented in the classical text on elasticity by Love (1944, originally published in 1892) in an old-fashioned format, and in a modern format in the seismology text by Aki and Richards (1980). Several texts on wave propagation, such as Mason (1958) and Achenbach (1973), came with the growth of technology after World War II. Waves in anisotropic (also called aeolotropic) materials were first analyzed by Musgrave (1954).

The theory of elasticity is presented in detail by Love (1944), and in the concise tensor notation by Sokolnikoff and Specht (1946), where the theory of vectors and tensors is also reviewed. A clear exposition of this field is given by Wills (1958). The theory of waves in rods and plates is described by Doyle (1989).

Many specialized research papers on all aspects of UT, elastic waves, and digital techniques have been published in such journals as *Ultrasound*, the *Journal of the Acoustical Society of America*, the *Journal of Applied Physics*, and the *IEEE Journal*, as well as various texts, conference proceedings, and dissertations, many available on the Internet. Some of the material in this book is from unpublished reports and conference proceedings written by the author.

This Page Intentionally Left Blank

ACKNOWLEDGMENTS

The author particularly thanks Ken Friddell, who opened many doors into the world of UT; Brian Coxon, who prodded him into writing this book; Gary Schwartz for discussions of medical applications; and Malcolm Povey, for help with scattering theory. He also thanks all those practitioners of UT he has worked with and who tolerated his insolent questions and provided the understanding which forms the basis of the book.

He also acknowledges the helpful comments given by the reviewers, and he thanks Joel Claypool, Executive Editor, and all the team at Academic Press who took a rough manuscript and turned it into this book, as well as Cordelia Sealy who convinced the Press to publish it.

This Page Intentionally Left Blank

1

INTRODUCTION

WHAT IS ULTRASONICS?

Ultrasonics is the science and exploitation of elastic waves in solids, liquids, and gases, which have frequencies above 20 khz (the nominal limit of human hearing). Waves of lower frequencies are called acoustic. Waves with a wide range of frequencies are called acousto-ultrasonic. An upper limit to the frequency in many UT applications is taken at about 10 MHz, but many applications use frequencies as high as 5 GHz. (Note that these frequencies can lead to propagating electromagnetic (radio) waves, and so they are referred to as radio frequencies, RF.)

UT tests often produce an audible buzz. This arises from repetition of each test many times, electronically at frequencies around 100 to 1000 Hz, which is audible.

HOW HAS UT TECHNOLOGY PROGRESSED?

The advent of the digital age led to a change from hardware to software for all aspects of UT testing, such as robotic scanning, controlled transducer excitation, and signal acquisition. The increase in computing power in the 1980s and 1990s, and the advances in data processing developed for space photography and radar, have led to new signal processing and display techniques, thereby improving the detail which can be extracted from a UT signal.

Ultrasonic methods have been used extensively in manufacturing since the 1940s as a means of detecting flaws in metals for quality control (QC). Those techniques mostly looked at the echoes from a propagating pulse. Advances in electronics in the 1960s and 1970s made possible applications to the precise measurement of wavespeed for a variety of purposes.

Advances made by the medical community in the 1970s and 1980s included the UT scanning of a fetus in a pregnant woman and the measurement of blood flow in arteries. UT is now used to explore pathologies in the human body, represented by variation in the shape, size, or motion of organs, or the presence of abnormal material such as tumors. Human organs have odd shapes, sliding interfaces, veins, fibers, and fluids.

In the same period, the silicon industry invented the ultrasonic microscope for examining transistor chips.

The extensive use in the 1980s and 1990s of plastics and thin fiber-reinforced composite materials in aerospace structures, and the push to improve products in many industries, such as lumber and plywood has spurred a variety of applications. These materials are usually nonhomogeneous (they vary from place to place) and are anisotropic, meaning that directionality has a strong effect. Modern metals, such as titanium, beryllium, and even stainless steel, are also anisotropic. The configurations of parts do not have simple shapes with flat surfaces, and they include thin sheets which exhibit bending.

Also in the 1980s and 1990s extensive use has been made of UT in processing and evaluating food (see for example Povey, 1997a and 1997b).

WHAT ARE SOME CURRENT INDUSTRIAL/MANUFACTURING APPLICATIONS OF UT?

The most common use of UT in industry is to inspect fabricated parts for defects, such as cracks in welds and at holes, porosity in castings, and irregularities in composite materials. Material properties and distances can also be measured.

UT was originally exploited as part of quality control (QC), being known as nondestructive testing (NDT) or nondestructive inspection (NDI). As the technology was developed into quantitative measurements of properties, it acquired the names nondestructive evaluation (NDE) or quantitative NDE (QNDE).

What Is Defect Detection?

In industrial applications, a defect is either a discrete feature of an object, such as a crack, or a region of faulty material. It is often sufficient to detect and indicate the presence of such defects with no quantitative measurement, but in some cases a quantitative assessment of the defect is provided. Several examples of such detection follows.

Two types of defect and their effects on wave propagation are illustrated in Fig. 1. An internal crack produces an early echo (a reflection) with less attenuation (i.e., a stronger signal) than that from the back face of an object. Low-density material attenuates and slows a wave so that it arrives late and with reduced amplitude.

The detection of a delamination in a composite is illustrated in Fig. 2: a cylinder made of an oblique layup of fabric which included a delamination (introduced by inserting a piece of mylar between two layers) was scanned by pulse echo. Typical echo waveforms are shown, with the positive phases of the waves darkened to improve interpretation. The variations in arrival times over a region of the surface were mapped by a contour plot, revealing the delamination. Note that a map of wave amplitudes would not reveal it.

Several composite panels fabricated with differing skin thicknesses and core stiffnesses were examined by through-transmission ultrasonics (TTU).

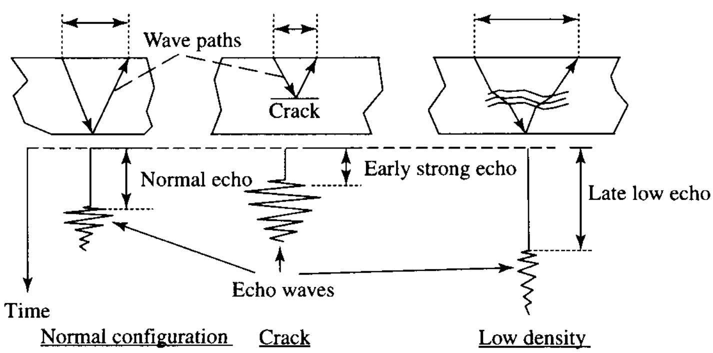


FIGURE 1 The effects of two types of defect

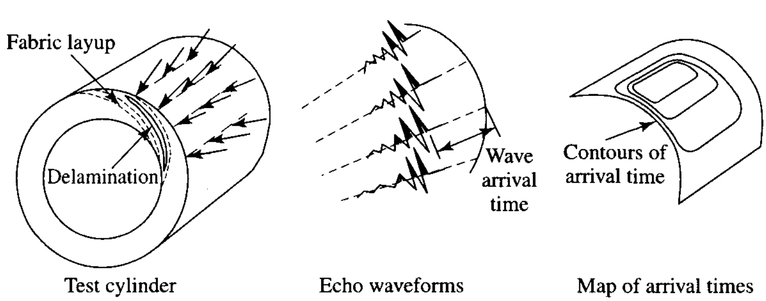


FIGURE 2 Detection of a delamination by pulse echo

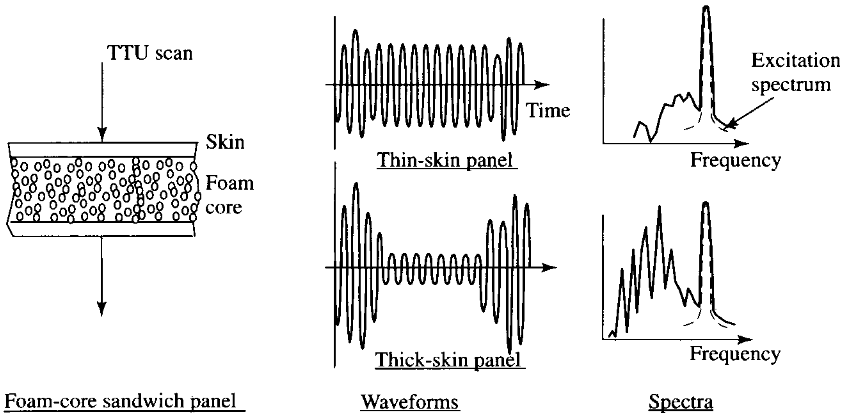


FIGURE 3 Responses of sandwich panels

The results of peak amplitude scans were inconsistent with panel configuration and test frequency, so individual waveforms were examined. As shown in Fig. 3, these consisted of long series of oscillations (tone-bursts) with short low-frequency transients superposed at the beginning and end. The thin-skinned panels showed strong signals at the excitation frequency, but the thick skins inhibited this somewhat. The transient oscillations were stronger in the thick case. Spectral analysis showed that the transients were at several low frequencies relative to the excitation, due to bending of the skin.

What Is Acoustic Emission (AE)?

When a material is stressed heavily, it emits waves from points of microscopic internal damage, such as initiating cracks, twinning and dislocation motion in metals, or breaking fibers in composites. These waves can be detected by transducers to interpret the onset of damage. They are emitted as sharp transient bursts, having broad frequency spectra. This topic is related to UT inasmuch as elastic waves are monitored by similar techniques, but it is generally treated as a separate field.

What Is a Tap Test?

A valuable and common technique for quality testing an object (from cracks in a structure to the ripeness of melons) is to tap the object sharply with a coin or other hard object, and to evaluate, aurally, the emitted sound. This technique is one of experience and skill, and is very successful when done right. Attempts at developing automated high-technology versions of this test have failed.

How Are Material Properties Measured?

By directly measuring the time-of-flight (time of wave propagation through an object) the wavespeed or distance of propagation can be measured. If the distance (size of the object) is known, then wavespeed follows. This provides a direct measurement of elastic properties of a material. In addition, correlations can sometimes be made between the wavespeed and certain properties, such as strength. If the properties of the material, in particular the wavespeed in the object, are known, then the distance of propagation follows, enabling dimensions to be verified.

WHAT ARE THE MAJOR MEDICAL APPLICATIONS OF UT?

Medical applications include diagnosis, commonly called sonography, and therapy.

Sonography is used for evaluating the condition of internal organs and tissues, most commonly for imaging neonatal fetuses and the beating heart, and for measuring blood flow. Therapy is provided by high intensity waves which heat tissues providing massage, or break stones. These techniques are reviewed by Wells (1969) and Ensminger (1998).

What Is Sonography?

Two distinct methods are used: pulse-echo (based on reflected and backscattered waves for mapping organs), and Doppler shift (which measures the flow rate of blood or the motion of an organ).

For organ mapping, a hand-held array of transducers or a single transducer is passed over the body, directing waves toward the internal organ of interest, as illustrated in Fig. 4. Extensive software is used to convert the received signals into a suitable visual presentation, based on depicting the echoes on a plan of the UT beam, as illustrated in Fig. 5. Essential to sonography is the display of images, which allows a medical practitioner to evaluate any pathology in the tissues. Generally, sonographers are trained to understand the physics behind the images, but not their medical significance.

The arrangement for measuring blood flow uses a transducer applied to the skin at an angle to propagate waves into a blood vessel, such that a component of the waves lies in the direction of flow, as illustrated in Fig. 6. The returning echo is modified in frequency by the Doppler effect of the blood flow, as discussed in Section 4 and Appendix 9.

What Is Lithotripsy?

Lithotripsy is the application of intense UT waves to breaking up kidney stones in human patients, avoiding surgery. Two or more high-intensity beams are

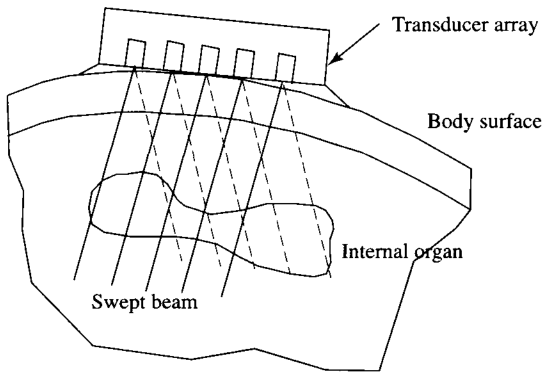


FIGURE 4 Sonographic scanning concept

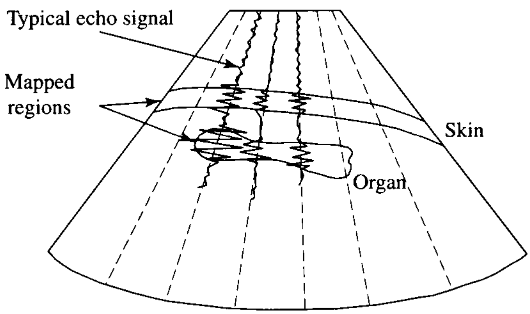


FIGURE 5 Sonographic image forming

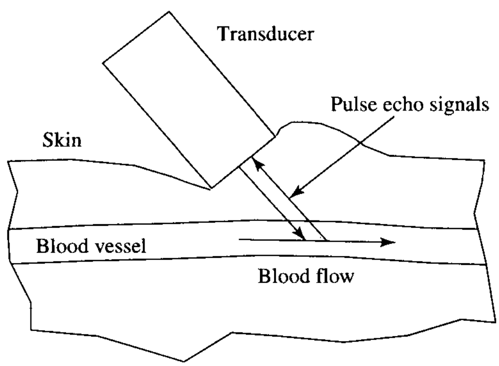


FIGURE 6 Doppler flow measurement

focused from carefully positioned sources onto the stone, to generate a very high local stress state.

WHAT ARE SOME UNIQUE APPLICATIONS OF UT?

The following are a few of many imaginative applications of UT. It can be expected that many more will be developed.

UT is used as a detector in automatic door openers, where the interruption of a beam signals the presence of a person, or in remote controls for TV, etc., where a signal is transmitted on a beam. (Note that many such devices are now based on infrared rays, IR.)

The Polaroid camera (and undoubtedly other modern cameras, too) uses UT as a range finder for automatic focusing. The camera emits a signal aimed at an object, and the echo is detected. The waves propagate through air, whose wavespeed is known. The propagation time, or delay between emission and return, is then a measure of the distance of the reflecting object. (It is said that technicians needing simple range-finding devices often bought a camera and extracted the device!)

Sonar and depth finders use sonic pulses in water to locate echoes from the sea bottom or from shoals of fish. Although these are of a lower frequency than traditional UT, the operation is essentially the same.

Ultrasonic waves in a water tank are used for cleaning jewelry and window blinds.

An advanced application determines the state of the huge foundation bolts (6–10 in. dia) of an oilrig through the effects of nonlinear elasticity: increasing load increases the modulus so that the wavespeed falls. Any change in the load in the bolt can be detected by continuously monitoring the wavespeed of waves propagated along the bolt.

The pressure of gas in a storage tank on a rocket was determined by measuring the wavespeed through the gas, which depends on the pressure.

What Quantities Can Be Measured by UT?

The propagation of some specific feature of a wave (such as its onset or the magnitude of its peak) can be measured with great accuracy, limited only by how precisely the feature can be identified. (Because the wave can be distorted during propagation, the feature may be hard to identify.) There are two classes of measurable features:

Time
Amplitude

The propagation time for a wave can be used to measure wavespeed when the path length is known, or the path length when the wavespeed is known.

Wavespeed can be used to calculate density if elasticity and wave type are known, or elastic properties if density is known. The path length allows accurate measurement of distances (e.g., thickness, internal dimensions).

The amplitude of a wave can be used to determine the damping (energy absorption) characteristics of the medium, thereby indicating the nature of the material. Energy is an important quantity that can be determined by the amplitude integrated over time.

These definitions can be refined by analyzing the spectrum of the waveform, to obtain its frequency content.

What Is the Difference between a Solid, a Liquid, and a Gas?

There can be several fundamental types of wave in a given medium: a solid can have two or three, a liquid or a gas can have one. A solid and a liquid can have surface waves, but a gas cannot.

The differences arise from the type of elastic response between force, or stress, and deformation, or strain, which the medium exhibits. A solid exhibits compressibility, shear, and Poisson directional effects, whereas a liquid or a gas exhibits only compressibility. Viscosity and gravity have a notable effect on liquids, but only to a small degree in solids and gases. Nonlinearity affects gases, but has less of an effect on solids and liquids.

Some of the theory of UT waves treats them as acoustic, propagating as if in a fluid, where only pressure forces arise, and there are no shear or directional effects. This is sometimes a gross and misleading simplification, since shear waves arise in many conditions and shear forces and deformations exist in all waves. These effects must be considered when interpreting wave arrivals, attenuation effects, etc. In this book, Poisson and shear effects are included wherever possible, although the associated mathematics becomes complicated and the approximation of neglecting them is sometimes justified.

2

DEFINITIONS AND BACKGROUND

WHAT IS A WAVE?

A wave is a moving transition between two states of a medium. The state in an elastic wave is the stress acting on the medium and its velocity. An elastic wave carries changes in stress and velocity. The state can take a jump, it can oscillate periodically in time and position, or it can vary in some arbitrary fashion. A step and an oscillating wave are illustrated in Fig. 7.

A wave of elastic deformation is created by a balance between the forces of inertia and of elastic deformation. A wave moves at a speed (the wavespeed) which is determined by the material properties, the body shape, and sometimes the frequency.

Wavespeeds are relatively fast and vary from about 0.33 mm/μsec (1100 fps, 760 mph) in air and 1.5 mm/μsec (5000 fps) in water to over 6 mm/msec (20,000 fps) in metals.

What Is Material (Particle) Velocity?

The wave imparts motion to the material as it propagates. This is referred to as particle (or material) motion, to distinguish it from the wave motion, and is usually specified as a velocity (magnitude and direction).

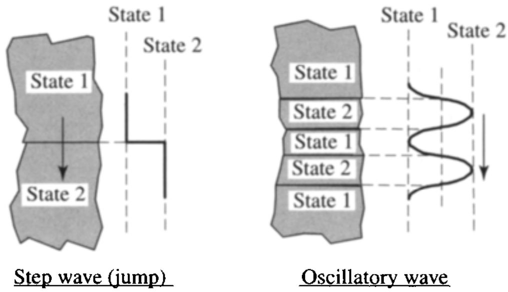


FIGURE 7 A wave is a moving transition between two states of a medium

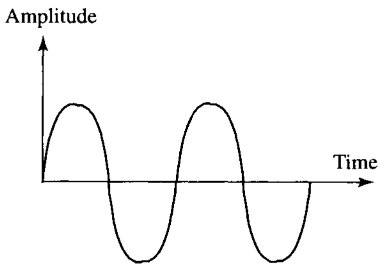


FIGURE 8 Sine waveform

In UT the material moves much more slowly than the wave, i.e., particle velocity is much smaller than wavespeed. The velocities induced in the material are usually very small, typically 10^{-9} in./sec, determined by the severity of the excitation which produces the wave. Using fluid dynamics terminology, ultrasonics is subsonic! (Ultrasonic refers to frequency, subsonic to velocity.)

How Big Are Typical Stress Oscillations?

Stresses are typically around 10^{-6} psi. These levels, which approach the magnitudes of molecular motions, are much smaller than experienced in engineering situations so that UT instruments are much more sensitive than those used in engineering.

What Is a Waveform ?

The sequence in time of the motions in a wave is called the waveform. The simplest mathematically is the sine wave illustrated in Fig. 8.

A typical signal includes several waveforms: the initial excitation (main bang), though this is frequently suppressed; a front-face echo (in a system with a stand-off transducer, such as in a tank or a squirter, see Section 5); and one or more back-face echoes, as sketched in Fig. 9.

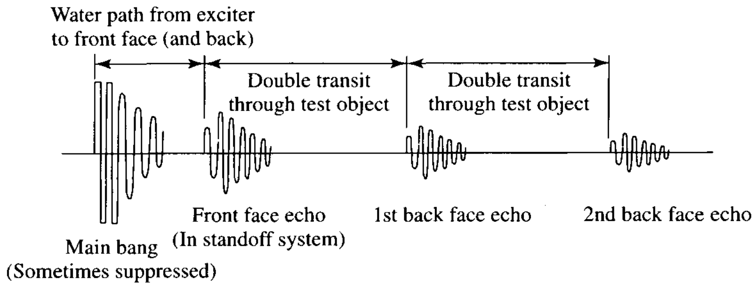


FIGURE 9 Typical signal from a pulse-echo system

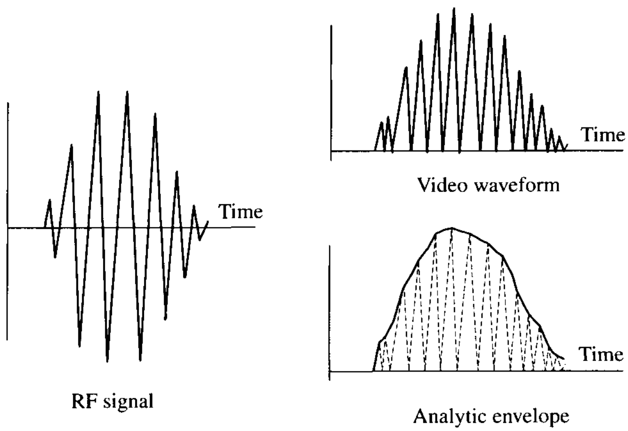


FIGURE 10 Types of waveform display

A through-transmission system (TTU) produces a direct transmission pulse, without the main bang or front face echo, and possibly one or more reverberations.

What Are the Video and RF Displays and the Analytic Envelope ?

A waveform can be displayed in three ways as illustrated in Fig. 10 (and described in Section 6A):

- The RF, or radio-frequency, waveform as measured (i.e., the electrical analog of the oscillation)
- The video waveform after rectification (inverting the negative swings) and sometimes, smoothing to display essentially the amplitude envelope
- The analytic envelope, constructed through a data processing algorithm

What Is a Wavefront, a Wave Surface ?

Propagating waves form a surface whose shape depends on the material type, and the excitation, i.e., the distribution and timing of the source. This shape is the wavefront.

For a point source in a uniform isotropic medium, the wave propagates as growing or shrinking spheres with the same speed in all directions, as sketched in Fig. 11.

In an anisotropic material, the wave surface is distorted, since the wave propagates at different speeds in different directions. This requires numerical analysis, as discussed in Appendix 9.

A force distributed over a flat surface excites waves which propagate as planes, called plane waves, as sketched in Fig. 12.

The shape of a wavefront generally depends on the shape of the surface where excitation is applied, on the direction of the surface force, as well as on the wavespeed in the direction of propagation.

The wavefront may reflect the shape of the surface if the material is uniform and the excitation is normal to the surface at all points, but if the properties vary so that the wavespeed varies, then the wavefront may take a different shape.

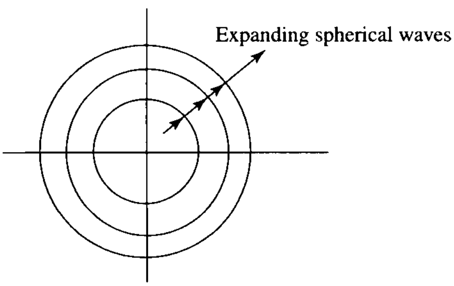


FIGURE 11 Typical wavefronts in isotropic material

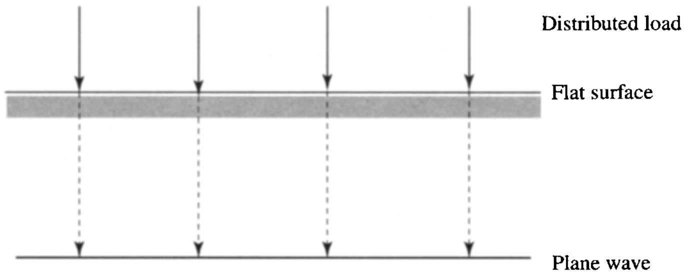


FIGURE 12 Formation of a plane wave

What Are the Propagation and Polarization Vectors?

The direction of propagation of a wave is the propagation vector. The direction of motion of the material is the polarization vector illustrated in Fig. 13.

What Is a Ray, a Wave Pattern?

A ray is a straight or curved line which follows the normal to the wavefront (i.e., the propagation vector is tangent to the ray) and represents the two- or three-dimensional path of the wave.

One must know the ray patterns of a wave system in order to understand any UT measurement based on waves arriving at a point on the surface. A typical ray pattern in two space dimensions as sketched in Fig. 14 would be created by multiple reflections in two or more layers. Evidently the echoes arrive at various points along the surface.

What Is an $x-t$ Wave Diagram ?

The propagation of a wave can be illustrated in a time–distance wave diagram, called an x–t diagram, as shown in Fig. 15. This diagram shows a distance–time

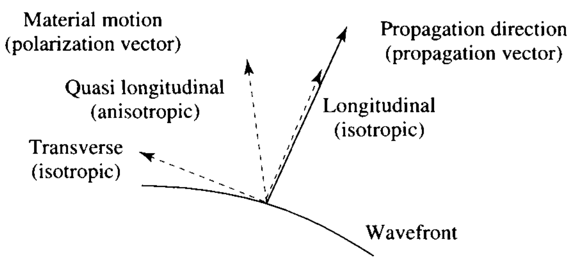


FIGURE 13 Definition of propagation and polarization vectors

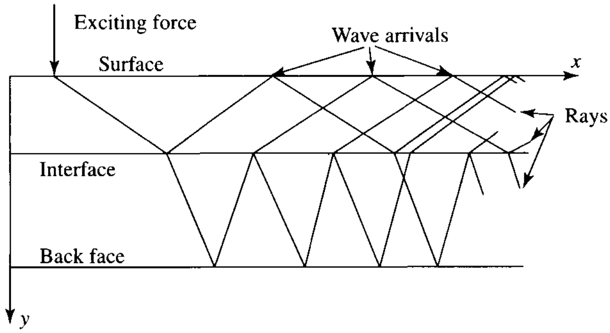


FIGURE 14 Example wave ray pattern in a two-layer system

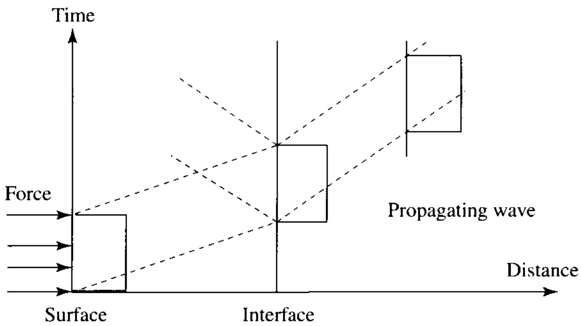


FIGURE 15 Illustration of a wave diagram

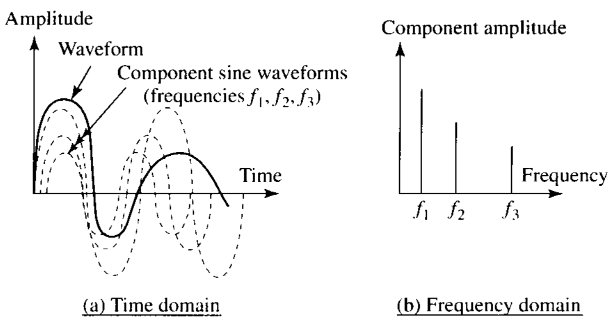


FIGURE 16 Time and frequency domain representations of a waveform

graph in which a wave system is represented by one or more lines which show the position of the wave as time passes.

The x–t diagram is derived and described in detail through the concept of Riemann invariants in Appendix 5.

WHAT ARE THE QUANTITATIVE MEASURES OF A WAVE?

A wave is usually not a pure single-frequency oscillation, but is a combination of sinusoidal oscillations with different frequencies, amplitudes, and phases as illustrated in Fig. 16. Nevertheless, most UT systems use waves in which one frequency dominates.

The sequence of amplitudes of those sine components is the spectrum of the waveform, also illustrated in Fig. 16. The several quantitative measures required to describe a wave are:

- Phase, frequency, period
- Wavespeed, wavelength, wave number
- Amplitude of particle velocity and stress

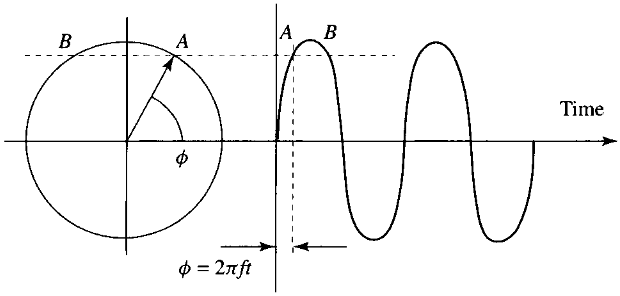


FIGURE 17 Rotating pointer representation of steady oscillation

What Are Phase, Frequency (Hz, kHz, MHz, GHz) and Period?

Frequency, f, is the rate of oscillation, i.e., number of oscillations in unit time. Conversely, the time taken for one oscillation is the time period, t_p, so that $t_p = 1/f$.

A steady sinusoidal oscillation with magnitude $a(t)$ as a function of time t, with an amplitude A, and circular frequency ω, written as $a(t) = A \sin \omega t$, can be represented by a pointer which rotates at a constant angular rate as sketched in Fig. 17. Such oscillations are called simple harmonic. The angle of rotation at a certain time is called the phase, ϕ, of the wave at the time t:

$$\phi = 2\pi f t = \omega t$$

where f is the angular frequency so that the circular frequency is $\omega = 2\pi f$.

The units of frequency are multiples of hertz (abbreviation: Hz). One hertz is the number of oscillations (cycles) per second.

k stands for kilo or 10^3 (1000)
M is for mega or 10^6 (1,000,000)
G is for giga or 10^9 (1,000,000,000).

What Is Wavelength and Wave Number ?

Wavelength, λ, is the distance occupied by one spatial cycle of the wave at an instant of time. Conversely, the number of cycles in a unit distance is $1/\lambda$, and the wave number is the number of radians in one cycle, $\kappa = 2\pi/\lambda$. Again using the rotating pointer representation, the phase shift over a unit distance is $\phi = \kappa x = 2\pi x/\lambda$, as illustrated in Fig. 18. Waves can propagate away from their source in either direction, so that the wave number κ can be positive or negative.

Just as a wave usually does not have a single frequency, it may not have a single wavelength, but it may have a dominant one.

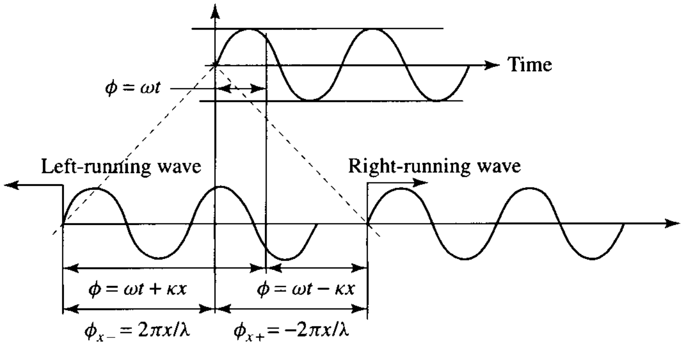


FIGURE 18 Phase in left- and right-running waves

What Is Wavespeed or Phase Velocity ?

The progress of a wave is described by its wavespeed or phase velocity. No material moves at this speed. Wavespeed depends on the material, and, except for isotropic materials, it depends also on the direction of propagation.

The total phase, temporal and spatial, of a propagating wave is

$$\phi = \kappa x - \omega t = \kappa(x - ct)$$

where $c = \omega/\kappa = f\lambda$ is the wavespeed. The temporal phase is generally, for a right-running wave, taken as negative to represent an increasing x-position of the wave as time increases. The opposite holds for a left-running wave. The wavespeed is determined by the forces of elasticity and inertia. The alternate name, phase velocity, is derived from the fact that it describes the rate of change of the phase for a fixed wave number or wavelength.

What Is the Amplitude of Particle Velocity and Stress?

The amplitude of a steady sine wave is the peak magnitude reached in each cycle. In a waveform comprising several superimposed component waves at different frequencies, the amplitude can be defined as the peak, but the wave is better characterized through the amplitudes of the individual components. This collection of the amplitudes is the spectrum.

The amplitudes of stress and velocity in a wave are related by the acoustic impedance:

$$\sigma = zv,$$

where $z = \rho c$ is the acoustic impedance, with ρ the density and c the wavespeed. When wavespeed depends on frequency, this relationship can only be applied to individual frequency components.

What Is Inertia ?

According to Newton's Law, a body at rest or in steady motion resists any change in its motion. When a force is suddenly applied to an elastic medium, deformation is resisted by inertia. Inertia introduces the time factor into deformation. A force is required to induce acceleration, with mass as a proportionality factor. (Newton's Law of Inertia is discussed in Appendix 4.)

WHAT IS ELASTICITY?

A material medium which deforms reversibly under load is called elastic: there is a unique correspondence between force and deformation through loading and unloading, as shown in Fig. 19. In a simple material such as most metals, the correspondence is a linear proportionality. Other materials such as rubber exhibit nonlinear elasticity.

The relationship between deformation and the force applied to an object depends not only on the nature of the material but also on the shape of the object, the distribution of the load, and the way deformation is measured. Deformation is defined as the displacement of one point of a body relative to another and depends on the distance and direction between the points.

Force and deformation are extrinsic quantities (determined by factors outside the object) and must be described through stress and strain (intrinsic quantities determined locally in the object), defined below. These are related by the elastic modulus, a purely intrinsic material property.

What Are Stress and Strain ?

A force which is applied to a surface can sometimes be treated as a point force, but if the details of its interaction with the material must be known, it must be

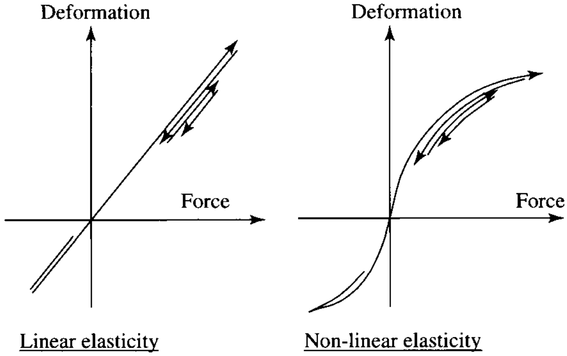


FIGURE 19 Illustration of elasticity: reversibility in deformation and force

treated as being distributed over a segment of the surface, as shown in Fig. 20. It is then described by the local force per unit area, which is stress.

When the direction of the force is along the normal to the surface, it produces a normal stress, and when it is tangential to the surface it produces a transverse or shear stress. Stress must therefore be described by the direction of both the force and of the normal to the surface, so that there can be six independent components of stress. The three directions for each of force and area lead to nine components, but because of symmetry required by rotational equilibrium, only six are independent. Stress has the dimensions of force divided by area (length squared).

Deformation of a body usually varies over a region, as illustrated in Fig. 21. In some cases it is sufficient to consider the overall deformation, but to describe the material behavior, it is necessary to consider local deformation. It is then described by the local deformation per unit length (strain).

Strain also has six independent components because deformation is directional and is defined over a small distance which is also directional. Again, strain is symmetric because of energy conservation. Strain is dimensionless and has

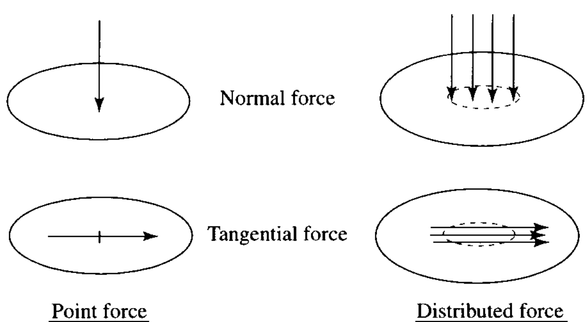


FIGURE 20 Point and distributed forces

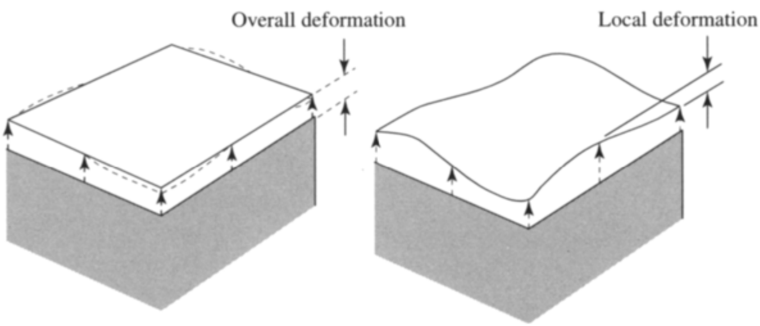


FIGURE 21 Overall and variable deformation

no units, being the ratio of two distances. Nevertheless, strain is often quoted with units such as inch per inch.

Detailed definitions of stress, strain, and elasticity are given in Appendix 1.

What Are Isotropy and Anisotropy (Aeolotropy)?

If the deformation of an elastic medium is independent of the direction of an applied force, it is called isotropic. A material in which the deformation depends on the direction in which the force is applied is anisotropic or aeolotropic. (See Appendices 2 and 9 for isotropic and anisotropic elasticity.) These types of response are illustrated in Fig. 22. A force applied in either of two directions to an isotropic object results in the same shape relative to the force direction, but the shape is different for an anisotropic object.

Examples of anisotropic materials are crystals; metals such as titanium, beryllium, and stainless steel; biological materials such as wood and bone; and man-made composites such as plywood and fiber-reinforced plastics.

What Is the Elastic Modulus ?

The ratio of stress to strain in the linear region is called the elastic modulus or elastic constant. The most direct way to measure elasticity is to measure the extension of a portion of a rod under axial tension. The ratio of applied stress (force/area) to axial strain (extension/length) produces the modulus called Young's modulus.

A simple alternative is to measure the UT wavespeed for one or more configurations and use the relationships given in Section 3.

Generally, a different modulus is obtained for each measurement of some component of strain caused by some component of stress. Since there are

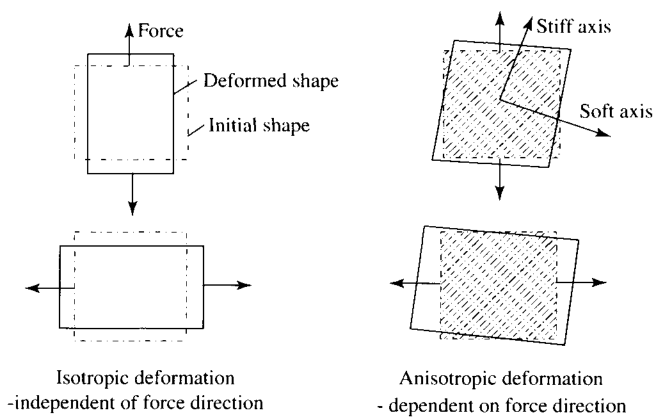


FIGURE 22 Isotropic and anisotropic response

6 independent components of strain and 6 of stress, it would appear that there could be 36 different moduli. However, energy conservation requires that there is symmetry in all interactions between components. Thus there are 6 direct moduli and 15 [i.e., (36 − 6)/2] interaction moduli, so that there are up to 21 independent moduli to describe the measurements, depending on the symmetry of the material structure.

The most common material has no directional dependence and is called isotropic. It has two independent elastic constants, commonly Young's modulus and Poisson's ratio, from which all others can be determined (see Appendix 2). Other constants such as the shear and bulk moduli or Lame's constants are also used.

A material with 1 axis of symmetry (as in the direction transverse to the thickness of plywood or in a hexagonal crystal) is called transversely isotropic and has 5 independent elastic constants. A material with 3 orthogonal axes of symmetry (as in beryllium or a laminated fiber reinforced polymer) is called orthotropic and has 9 independent elastic constants.

The units of modulus are those of stress, since strain has no units.

What Is Poisson's Ratio ?

A force applied to an elastic body produces a deformation in the direction of the force as well as transverse deformations in other directions. In an isotropic material, the ratio of transverse to in-line deformation is determined by Poisson's ratio, as illustrated in Fig. 23.

What Is Compressibility ?

The volume of a fluid (i.e., a liquid or a gas) or of a solid will change under a pressure (which is a force uniform in all directions), as illustrated in Fig. 24.

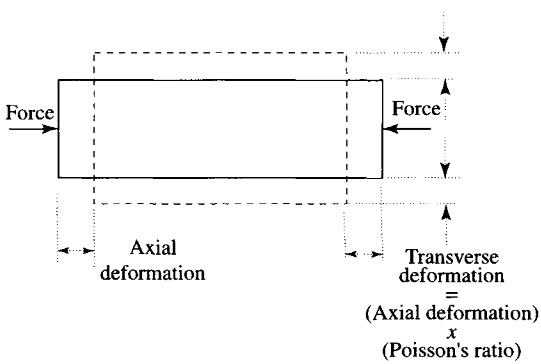


FIGURE 23 Illustration of the Poisson effect

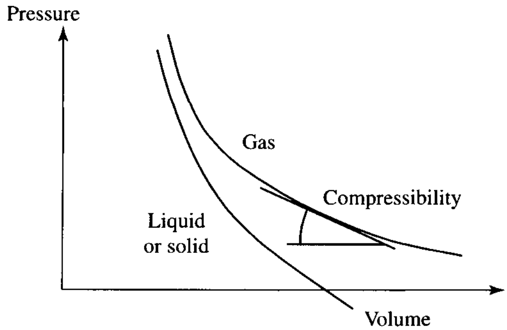


FIGURE 24 Pressure–volume relationship for liquids and gases

The proportionality between fractional volume change and pressure is the compressibility. It is a myth that water is incompressible! Since fractional volume change is a form of strain, and pressure is a form of stress, compressibility is an elastic behavior.

Gases are quite nonlinear but essentially reversible under modest pressures.

What Is Viscoelasticity?

In fluids (e.g., water, air, molten metals, uncured polymers), the stress which produces motion is dependent on the rate of strain, not the strain itself as in elasticity. This behavior is called viscosity. Most materials which appear to be solids (e.g., greases, wax, high-temperature metals, cured polymers) exhibit a combination of elasticity and viscosity: the stress is dependent on both the strain and the strain rate. This is viscoelasticity.

Viscoelasticity affects the wavespeed and the decay of a propagating wave.

What Is Acoustic Impedance?

The balance between elasticity and inertia develops into a linear relationship between stress, σ, and particle velocity, v, in a wave, $\sigma = zv$. The proportionality factor, z, is the acoustic impedance:

$$z = \rho c = (\rho C)^{1/2}$$

where ρ is the density, $c = (C/\rho)^{1/2}$ is the wavespeed, and C is the elastic constant appropriate to the wave type (discussed in Section 3 and Appendix 2).

The units of impedance are those of the product of density and wavespeed, e.g., g/cm^2-sec, or psi-sec.

What Is the Energy of a Wave?

A wave has mechanical energy which is the sum of elastic (or potential) energy and kinetic (or inertial) energy. At the peak stress, when there is no motion, the energy is entirely potential, and at the peak velocity when the stress is zero, the energy is entirely kinetic.

The energy is supplied by the source of the wave and is carried with it. It is reduced by processes which dissipate it as heat, through compression, viscosity, plasticity, etc. When these processes are small, the wave propagates a long way, but generally the wave decays in a moderate distance.

WHAT ARE THE UNITS OF ELASTICITY, DENSITY, AND WAVESPEED?

If length units are represented by [L], time units by [T], and mass units by [M], Newton's law relating force to acceleration of mass (see Appendix 4), shows that force has units $[F] = [M][L][T]^{-2}$. Stress and modulus have units of force/area, $[F][L]^{-2}$, which is then $[M][L]^{-1}[T]^{-2}$. Density is mass/volume, $[M][L]^{-3}$, so that wavespeed, which has units of $(\text{modulus/density})^{1/2}$, is $[L][T]^{-1}$, which are the units of velocity.

When making calculations it is essential that the units of force and mass be distinguished as in the following table, as required by Newton's Law (Appendix 4).

What Are the Systems of Units ?

There are three major systems of units, the English system (commonly used in U.S. engineering), the cgs system, and the Rational or MKS system. The units used for the relevant quantities in UT, and the conversions among them can be made using the following factors (e.g., distance in inches in the English system = 2.54 . . . × distance in cm. in the cgs system):

System of Units

	English	**cgs/dyne**	**MKS**
Distance	**in.**	**cm**	**meters, m**
English:	—	0.39376996	39.376696
cgs	2.5400051	—	100
MKS	0.0254. . .0.01		—
Area	in.^{2}	cm^2	m^2
English	—	0.1549997	1549.99. . .
cgs	6.541626	—	1.0E4
MKS	6.542. . .E-4	1.0E-4	—

Volume	in.^{3}	cm^3	m^3
English	—	6.1023376E-2	6.102...E4
cgs	16.3871628	—	1.0E6
MKS	1.639...E-5	1.0E-6	—

Velocity/speed	in./sec	cm/sec	km/sec
English	—	0.3938...	39.38....
cgs	2.540...	—	100
MKS	0.02540...	0.01	—
Common UT practice uses speed in/μs, cm/μs, or km/s, giving the following factors:	1.0E-6	1.0E-6	1.0E-3

Acceleration	in./sec^2	cm/sec^2	m sec^2
English	—	0.3938...	39.38...
cgs	2.540...	—	100
MKS	0.02540...	1.0E4	—
Gravity (g)	386.088 = 32.174 ft/sec^2	980.665	9.80665

Force	Pound (lb_F)	Dyne	Newton (N)
English	—	2.24808914 E-6	0.2248...
cgs/dyne	4.4482218 E5	—	1.0E5
MKS	4.448...	1.0E-5	—

Two types of force can be defined through Newton's Law.

1. Absolute force which gives a unit acceleration to a unit mass:

$$1 \text{ poundal} = 1 lb_M \times 1 \text{ fps}^2 \quad 1 \text{ dyne} = 1 \text{ g} \times 1 \text{ cm/sec}^2 \quad 1 \text{ newton} = 1 \text{ kg} \times 1 \text{ m/sec}^2$$

2. Gravitational force (expressed in mass units) which balances the gravitational attraction on a unit mass:

$$1 \ lb_F = 1 \ lb_M \times g = 32.174 \text{ poundals} \qquad 1 \ gr_F = 1 \text{ g} \times g = 981 \text{ dynes} \qquad 1 \ kg_F = 1 \text{ kg} \times g = 9.81 \text{ N}$$

An additional unit of mass, the slug, is defined in the English system such that 1 lb_F accelerates 1 slug by 1 fps^2.

Stress, modulus	lb_F/in.^{2} (psi)	Dynes/cm^2	N/m^2 (pascal, Pa)
English	—	1.4503807 E-5	1.450...E-4
cgs	0.68947161E5	—	10
MKS	0.68947161E4	0.1	—

Additional stress and modulus units are:

1 bar = 1E6 dynes/cm^2 = 1E5 Pa
1 GPa (GigaPascal) = 1E9 Pascal = 0.145 Msi = 0.145 E6 psi
1 atm = 14.7 psi = 1.014 bar.

Mass	lb_M	g	kg
English	—	0.002204768	2.204...
cgs	453.5924277	—	1.0E-3
MKS	0.4536...	1000	—

Density	$lb_M/in.^3$	g/cm^3	kg/m^3
English	—	3.6129892E-2	3.613...E-5
cgs	2.76779128	—	10
MKS	2.768...E4	1.0E3	—

3

WAVE PROPAGATION CONCEPTS

WHAT ARE THE TYPES OF ELASTIC WAVES?

There are two types of elastic waves: fundamental body waves which propagate inside an object, and surface waves which propagate near to, and are influenced by, the surfaces of an object.

What Are the Fundamental Body Waves?

Waves which propagate entirely inside an object, independent of its boundaries or shape, are called the fundamental waves. They are planar and, in some cases, spherical waves, governed only by elasticity and inertia, and are represented by the fundamental (simplest and most general) solutions of the wave equations.

In an isotropic material, the speed of a plane wave is the same in any direction, so a point source excites spherical waves, which can be regarded as the envelope of all plane waves emitted at the same time, as illustrated in Fig. 25. Similarly, the plane wave can be represented by an envelope of spherical waves emitted from points distributed over a plane, as illustrated in Fig. 26.

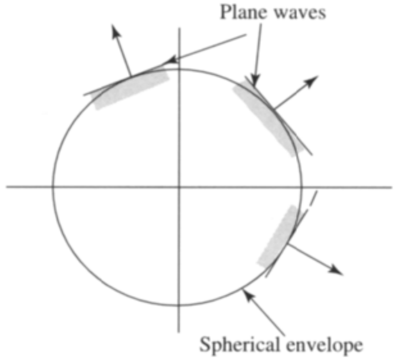


FIGURE 25 Spherical wavefronts in isotropic material

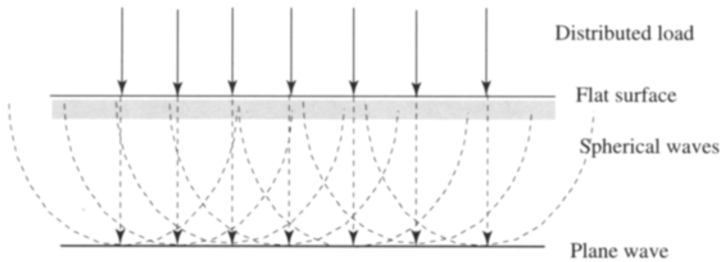


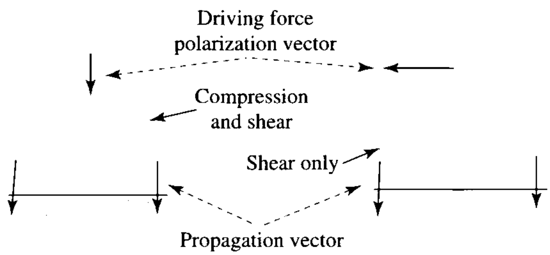
FIGURE 26 A plane wave represented by superposition of spherical waves

What Are Longitudinal (P-) and Transverse (S-) Waves?

Two types of fundamental waves can arise in an isotropic material: longitudinal (also called dilatational, primary, or P-), and transverse (or shear, secondary, or S-) waves as illustrated schematically in Fig. 27.

The deformation in a P-wave is parallel to the direction of propagation. It may be either expansion or compression, or an oscillation between the two. The driving stress component acts along the normal to the wavefront. In a solid, there are also stress components transverse to the wave front (i.e., in the plane of the wave) which are not equal to the normal stress. This stress state, which is not a pure pressure, includes a shear component. The deformation is uniaxial, i.e., only in the direction of propagation, and is a combination of compression and shear (see Appendix 2). In a fluid the P-wave is a pure pressure equal in all directions, hence the designation "P-wave," although the motions are the same as in a solid.

The driving stress in an S-wave is transverse to the direction of propagation and lies in the plane of the wave, perpendicular to the direction of propagation.



Longitudinal P-wave
(or primary, pressure)

Transverse S-wave
(or secondary, shear)

FIGURE 27 Schematic representation of the fundamental waves

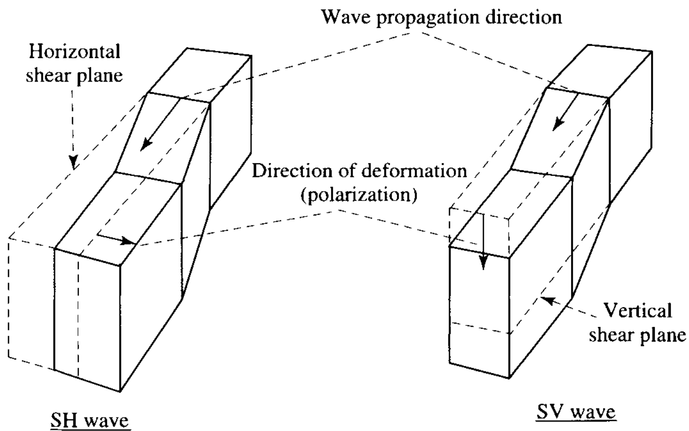


FIGURE 28 Transverse wave types distinguished by directions of deformation

The stress state is shear alone. The motions include both shear and rotation. There can be no S-waves in a fluid.

What Are SV- and SH-Waves?

Shear waves with shear planes perpendicular to one another can be considered to be independent types and can be combined to represent shear waves in other planes, excited by different shear stresses. These wave types are illustrated schematically in Fig. 28.

In geophysics/seismology, the shear planes are described in terms of three earth-surface coordinates: distance, horizontal, and vertical. Two shear planes of interest are then distance–horizontal and distance–vertical, so the waves are called SH- or SV-waves. There is nothing generically appropriate to this definition, but it has gained some popularity for laboratory descriptions. The shear planes should be defined for each configuration.

What Are the Fundamental Waves in an Anisotropic Material?

In an anisotropic material the only fundamental solution is the plane wave, and not the spherical wave. There are generally three fundamental plane waves, in which the direction of deformation, the polarization, is neither parallel nor perpendicular to the propagation, but is generally at an angle to the propagation direction, as sketched in Fig. 13. The wavespeeds of these plane waves vary with direction and must, generally, be analyzed numerically as discussed in Appendices 5 and 9.

In general, the motion may be nearly normal or nearly perpendicular to the propagation direction, depending on the degree of anisotropy and the direction of propagation with respect to the symmetry axes of the material. The waves are then called quasi-longitudinal or quasi-shear waves. If a wave propagates along a material axis of symmetry, the waves become P- or S-waves. At some angles the waves are quite dissimilar to P- or S-waves.

The directional dependence of wavespeeds for a typical fiber-reinforced polymer composite, a graphite–epoxy unidirectional fiber-reinforced laminate, are shown in Fig. 29. The properties of the material were given by Sahay and Kline (1991) as follows:

$$C_{11} = 161 \text{ GPa}, C_{22} = 14.5 \text{ GPa}, C_{12} = 6.5 \text{ Gpa},$$
$$C_{23} = 7.24 \text{ Gpa}, C_{44} = 3.61 \text{ Gpa}, C_{55} = 7.1 \text{ Gpa},$$
$$\rho = 1.61 \text{ g/cc}.$$

The symmetry axis for these values is the 1-axis, and the 2–3 plane is the symmetry plane. The propagation angle was chosen to be from the 2-axis, so that 0° is in the symmetry plane and 90° is along the 1-axis of symmetry.

The three wavespeeds in this graphite epoxy are distinct and vary considerably with direction, although the two quasi-transverse wavespeeds (modes 2 and 3) are similar and are the same along the symmetry axis. The speed of the quasi-longitudinal wave, mode 1, is very high.

It is very important to note that, in general, measurable wavespeeds do not propagate at these speeds, which are for simple nondispersive plane waves. In real situations where a UT beam diverges and contains a range of frequencies, waves are excited over a range of directions and speeds, and these combine to form group waves (discussed in Section 4).

The variation of polarization angle with angle of propagation in the graphite epoxy composite is shown in Fig. 30. The quasi-longitudinal wave (mode 1) and the first quasi-transverse wave (mode 2) have angles which are mirror images

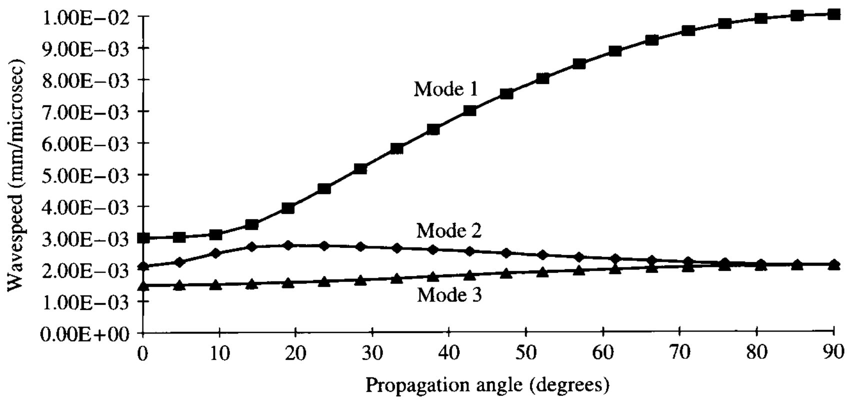


FIGURE 29 Wavespeed in an anisotropic fiber-reinforced plastic medium

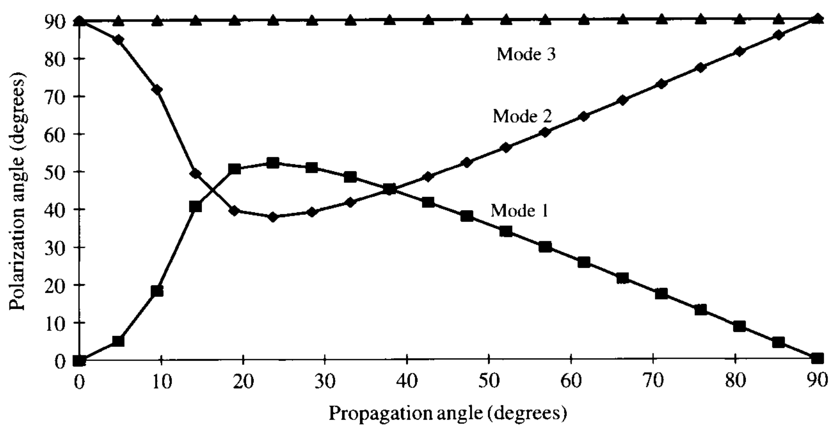


FIGURE 30 Polarization direction in a fiber-reinforced plastic

of each other. The second quasi-transverse wave is always at 90°, being in the plane of symmetry, and is therefore a true transverse wave.

What Are the Boundary-Dependent Waves?

Inevitably, all waves sooner or later reach the exterior boundaries of an object. The waves are then modified by the freedoms of the unconfined surface, resulting in the creation of new forms of steadily propagating waves (surface waves, plate waves, etc.).

There are three common types of geometry in which a free surface relieves the stresses in a fundamental wave to produce various forms of boundary waves:

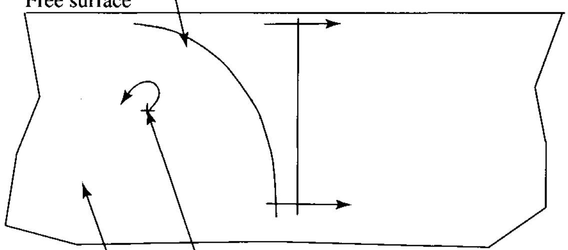


FIGURE 31 Schematic representation of motions in a Rayleigh surface wave

- Surface waves: Rayleigh waves, head waves, and interface waves
- Bar/rod waves
- Thin plate Lamb waves: Extensional and bending waves

What Are Rayleigh Surface Waves and Head Waves?

A Rayleigh wave propagates parallel to a free surface of an object at a unique speed, usually slightly slower than the shear wavespeed. The wavefront extends inward perpendicularly to the surface and can be planar, if excited by a line of force, or cylindrical, if excited by a point force. The motions combine compression and shear and follow elliptical paths, decreasing exponentially with distance into the object, as illustrated in Fig. 31.

A Rayleigh wave cannot be constructed simply from fundamental planar or spherical waves. When a force is applied to a surface, spherical P- and S-waves are excited and propagate into the object and along the surface away from the force. Normal and shear stresses are not supported along the surface, so additional waves arise to decrease the surface stresses to zero, thereby imparting additional motion. The P-wave runs ahead, followed by the S-wave. An oblique wave, called a head wave, runs from the P-wave to the S-wave, as sketched in Fig. 32, and a fan of waves runs from the S-wave, creating the Rayleigh wave.

The motions of the Rayleigh wave are the largest among all these waves. Analysis of the stresses and velocities, and of the wavespeed, in a Rayleigh wave is given in Appendix 10. Analysis of the system of waves excited by a point force on the surface of an isotropic half-space is called Lamb's problem (Aki and Richards, 1980).

The Rayleigh wave cannot exist on a liquid surface where there is no shear stiffness or Poisson effect. Other types of surface wave arise, controlled by gravity, but these are not generally of interest to UT.

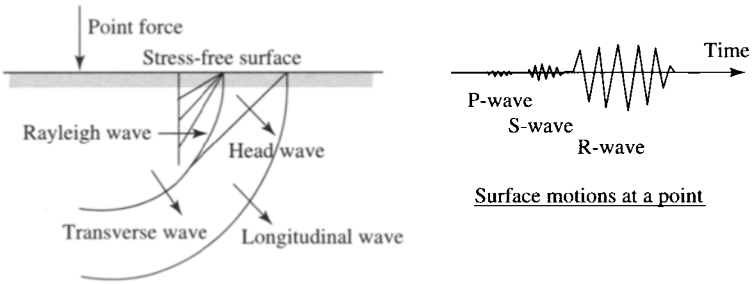


FIGURE 32 Excitation of Rayleigh wave by point force

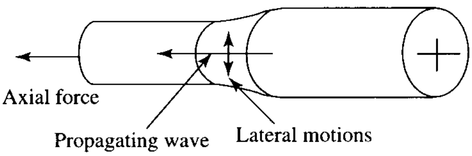


FIGURE 33 Longitudinal waves in a rod

What Are the Leaky Lamb and Stonely interface Waves?

A wave can propagate along an interface between two materials as a Leaky Lamb wave, or as a Stonely wave. It is excited by a wave which approaches the interface at a shallow grazing angle smaller than a critical angle which is determined by the wavespeeds in the two materials (see Section 4).

The Lamb wave is similar to a Rayleigh wave on the interface, penetrating into the material with the higher wavespeed, and radiating waves into the slower material at the critical angle.

The Stonely interface wave is a special class of wave which can arise only if the Poisson's ratios in the materials on the two sides of the interface are close in value. It is then similar to two interacting back-to-back Rayleigh waves.

These waves are common in seismology but less so in UT.

What Are Rod/Bar Waves?

When a force is applied to the end of a slender rod, a wave propagates axially along it, as sketched in Fig. 33. Since the free surface does not allow lateral stress, lateral motion develops and only axial stress can develop. The lateral motion introduces lateral inertia which becomes significant at high frequency.

The P- and S-waves which are excited at the loaded end of the rod propagate only a short axial distance as they also propagate radially over the end. They are relieved by the free surface where they excite surface waves.

The rod waves can thus be regarded as surface waves propagating along the free surface, extending radially into the rod and imparting motion in one

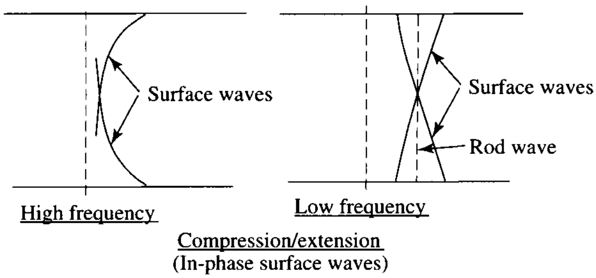


FIGURE 34 A rod wave as a combination of surface waves

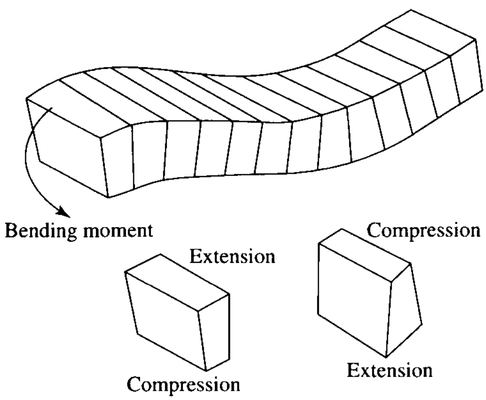


FIGURE 35 Bending deformation in a bar

direction over all the surface. If the wavelength is long compared to the cross-sectional dimensions of the rod, the surface waves extend completely across the diameter of the rod to form a plane wave of uniaxial stress propagating down the rod, as shown in Fig. 34. At shorter wavelengths the surface waves remain close to the surface, becoming independent surface waves at very short wavelengths. The speed of these waves depends on the wavelength as discussed in Appendix 11.

A torque applied at one end of a rod or a tube excites a torsional wave of shear which propagates along the axis.

What Are Bending Waves?

Bending of a bar creates lateral deformation, accompanied by a curvature of the bar, as illustrated in Fig. 35. According to the Engineering Theory of Bending the deformation is restricted to rotation of cross-sections of the bar: plane sections remain plane. This induces axial strain which varies linearly across the

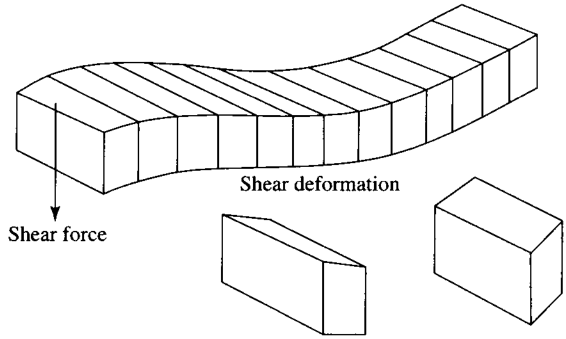


FIGURE 36 Shear deformation in a bar

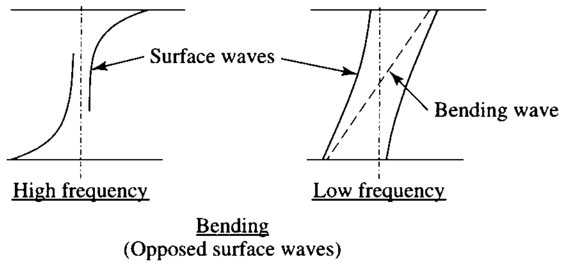


FIGURE 37 A bending wave as a combination of surface waves

cross-section. This state can be accompanied by shear deformation, as illustrated in Fig. 36.

The stresses on the ends of the bar are described by their resultants: a bending moment and a transverse shear force, as sketched in Figs. 35 and 36. This is discussed in Appendix 12.

Inertia of the lateral motion balances the gradient of the bending moment and the shear forces, leading to a dispersive wave system in which speed depends on frequency.

The bending waves can be regarded as surface waves with motions of opposite sense propagating along the opposite faces of the bar, as shown in Fig. 37. When the frequency is low, the wavelength of the surface waves is long, and the exponential depthwise variation of the waves is approximately linear. The resultant of the waves is then linear across the bar, as in classical bending. The resulting bending wavespeed is strongly dependent on frequency (see Appendix 12), falling to zero at zero frequency (i.e., static loading).

At high frequency, the surface waves on the two faces do not fully penetrate to the center of the sheet and propagate as individual surface waves at the wavespeed of surface waves.

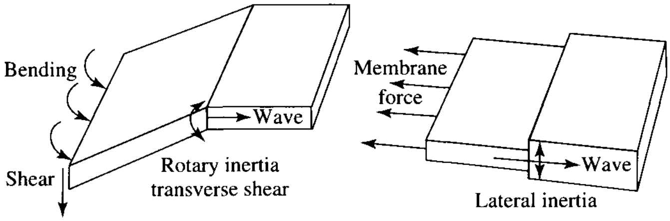


FIGURE 38 Illustration of lamb plate waves in bending and extension

What Are Lamb Plate Waves — Extensional and Bending?

There are two types of waves which propagate along thin plates: extensional and bending, as illustrated in Fig. 38. These waves are essentially the same as the rod and bending waves described above, except that the stress states are biaxial instead of uniaxial.

What Waves Propagate in Engineering Structures?

In traditional engineering, structural designs are complex assemblies of various components, including panels (which behave as plates in bending), webs (which behave as plates in shear), and spar flanges (which behave as rods), as well as large structures of concrete, wood, etc. In modern engineering, such as space vehicles, complex arrangements of plates (sometimes with sandwich construction), rods, and tubes (which behave as rods) are used.

Waves which propagate in these elements can be idealized as one or more of the types discussed above. Waves propagate across, or reflect from junctions between members according to the conditions for continuity of motion and balance of forces.

Many modern structures are made of fiber-reinforced composite (FRC) panels, sometimes as part of a sandwich construction with honeycomb, foam, or integrally molded web cores. FRC panels are laminated with many layers of plastic-impregnated woven fabric or linear fibers, usually in the plane of the panel, but sometimes laid up at an angle.

Laminates generally propagate bending and extensional Lamb waves at long wavelengths, but longitudinal or transverse waves propagate through the layers at short wavelengths (high frequencies). The long wave response allows bending stiffness to be evaluated, and hence damage to the outer layers can be detected. The short waves allow delaminations or porosity to be detected.

These materials are highly attenuative, mainly through scattering (see Section 4) by the nonhomogeneity of the fibers.

Sandwich panels respond with bending in the skin panels, coupled to shear or compression in the core, while very long waves produce bending or extension of the entire sandwich.

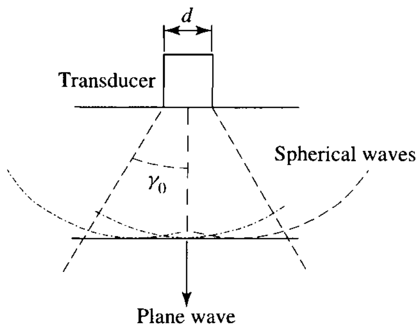


FIGURE 39 Illustration of a transducer beam

What Waves Propagate in the Human Body?

The human body is comprised of layers (skin, fat), cavities and tubes (the pulmonary system, the gastrointestinal system, and the circulatory system), irregularly shaped organs, fibrous muscles and ligaments, and bones (the skeleton). Generally, waves do not propagate through cavities or bones, so that these objects create strong reflections, concealing others beneath.

Wavelengths are never large enough to excite long-wave response such as overall bending or extension, and testing is usually made with penetrating longitudinal waves, which are essentially pressure waves, since shear is usually inhibited by the fluids and the sliding interfaces between organs.

WHAT IS A UT BEAM?

A transducer (see Section 5) which has a finite size excites waves from all parts of its contact face. These waves can be thought of as a sequence of spherical waves which coalesce where their phases match, forming a beam as discussed in Section 4 and Appendix 8.

This beam is conical, similar to a light beam, with a radial variation across it. Its strength varies with distance from the source. Outside the beam, the waves interfere and disappear, as illustrated in Fig. 39.

The beam angle and the radial variation are determined by the size of the source in relation to the wavelength, so that the beam will differ when a transducer is used on different materials. These aspects are discussed in Appendix 8.

The cone half-angle of the beam is approximately

$$\gamma_0 = \text{asin } (1.2\lambda/d) = \text{asin } (1.2c/fd)$$

where d is the transducer diameter, $\lambda = c/f$ is the wavelength, f is the frequency, and c is the wavespeed.

In the limit of a very small diameter transducer, a low frequency, or a high wavespeed, the beam diverges widely. Conversely, a large transducer, a high

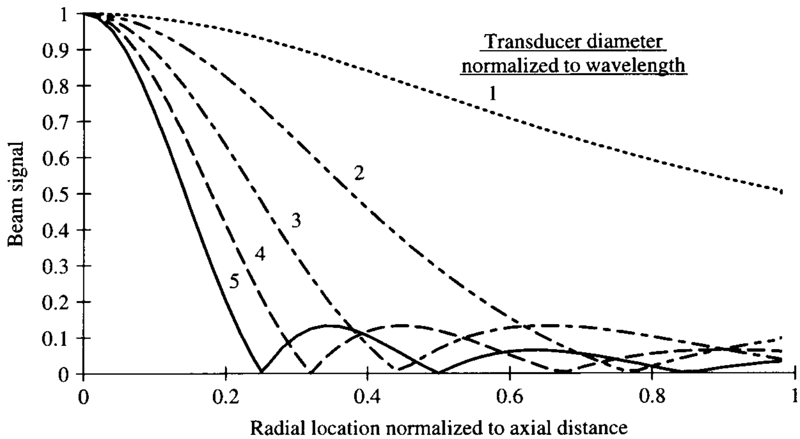


FIGURE 40 Radial variation of beam intensity

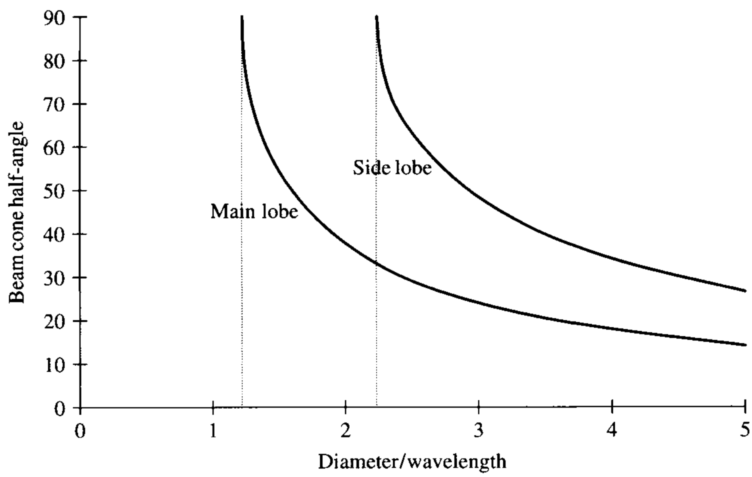


FIGURE 41 Half angles of beam

frequency, or a low wavespeed produces a narrow beam. The intensity of the plane wave varies with radius across the cone, as shown in Fig. 40.

These factors affecting the beam formation will influence test results achieved with different transducers or test set-up.

The variation is given by a Bessel function and has, beside the main lobe, several side lobes of reduced intensity, derived in Appendix 8. The half-angles of the beam and its first side lobe are shown in Fig. 41 as functions of the transducer diameter.

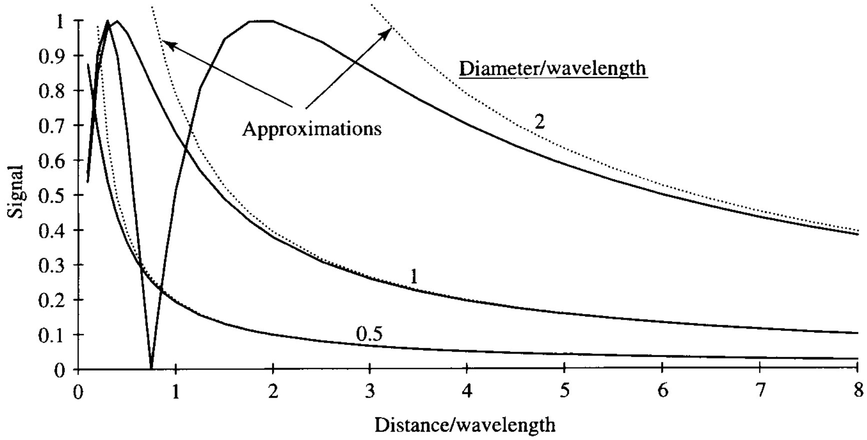


FIGURE 42 Axial variation of wave strength from a transducer

Note that there can be no conical beam for a transducer of diameter less than 1.21... wavelengths, and no side lobe for less than 2.23... wavelengths.

What Are the Near Field and the Far Field?

The amplitude or intensity of the beam varies along its axis as shown in Fig. 42. A factor $(1/2\pi)/(d/\lambda)^2$ is omitted from the signal strength to make it easier to visualize the differences in axial variations between the different transducer sizes. This factor changes the signal by approximately 2/3, 1/6, and 1/24, respectively, for the three cases of $d/\lambda = 2$, 1, and $^1/_2$.

At short distances from the transducer, the wave field is garbled by the interaction of multiple waves from all parts of the transducer surface. This is confined to a region called the near field, which is limited to the near-field distance, N, defined by the formula (see Appendix 8)

$$N = d^2/4\lambda = d^2 f/4c.$$

The far field is the region outside the near field where the transducer waves coalesce to produce a plane wave whose on-axis intensity decreases inversely with distance. For distances $L > 4N$ the intensity is approximately

$$S(L) \simeq S_0[1/(2\pi)](\lambda/L)$$

where $S(L)$ is the signal intensity at a distance L from the transducer and S_0 is the nominal intensity at the surface.

A large transducer produces a large near field and is thus not suitable for use on a thin object. A high-frequency and a low-wavespeed material have the same effect. Conversely, a small transducer produces a broad beam which attenuates rapidly.

What Are the Various Isotropic Elastic Wavespeeds?

The speed of a wave propagating in an unbounded elastic medium is constant—it depends only on elasticity and density, both of which are constant for the small range of stresses used in UT. There are two types of such waves, the longitudinal and the transverse. There are also two boundary configurations which produce unique easily measured wavespeeds with simple formulas. These are the axial wave in a slender rod and the Rayleigh wave on a free surface.

The simplest wavespeeds to measure, and those having the simplest formulas, are for axial and torsional (shear) propagation along a bar:

$$c_{\mathrm{b}} = (E/\rho)^{1/2}$$
$$c_{\mathrm{s}} = (G/\rho)^{1/2}$$

where c_{b} is the bar wavespeed, c_{s} is the shear wavespeed, $E = \rho c_{\mathrm{b}}^2$ is the Young's modulus of elasticity, $G = \rho c_{\mathrm{s}}^2$ is the shear modulus, and ρ is the density.

The moduli E and G, and also Poisson's ratio, $\nu = 2(c_{\mathrm{b}}/c_{\mathrm{s}})^2 - 1$, can be determined from these two speeds.

For the longitudinal and Rayleigh waves the speeds are

$$c_l = [(1-\nu)/2(1+\nu)(1-2\nu)]^{1/2}(E/\rho)^{1/2}$$
$$c_{\mathrm{R}} \sim [(0.87 + 1.12\nu)/(1+\nu)]^{1/2}(G/\rho)^{1/2}$$

where c_l is the longitudinal wavespeed and c_{R} is the Rayleigh wavespeed. An analysis and exact formula for c_{R} is given in Appendix 10. The ratios between various wavespeeds in an isotropic material depend only on Poisson's ratio:

Wavespeed ratio		$\nu = 0.3$
c_l/c_{b}	$[(1-\nu)/(1+\nu)(1-2\nu)]^{1/2}$	1.160
$c_{\mathrm{s}}/c_{\mathrm{b}}$	$[1/2(1+\nu)]^{1/2}$	0.620
c_{s}/c_l	$[(1-2\nu)/2(1-\nu)]^{1/2}$	0.655
$c_{\mathrm{R}}/c_{\mathrm{s}}$ (approx.)	$[0.87 + 1.12\nu)/(1+\nu)]^{1/2}$	0.873

These ratios show that the longitudinal wave is usually the fastest, and that the shear wavespeed is about half that. The Rayleigh wavespeed is a little slower than the shear wavespeed.

Similar formulas can be written for some waves in an anisotropic material (e.g., waves propagating along an axis of symmetry) as described in Appendix 3. For arbitrary directions in an anisotropic material, a numerical method must be used to calculate the wavespeed, as described in Appendix 9.

For nonelastic, e.g., viscoelastic or fluid media, or for high-amplitude stress waves or waves near a boundary, the wavespeed may depend on the frequency and other parameters such as dimensions. Waves whose speed depends on frequency are called dispersive (see Appendix 7).

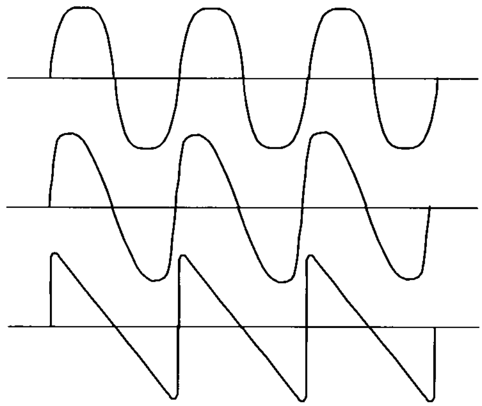


FIGURE 43 Nonlinear shock-up effect in a gas or liquid

What Is Shock-Up, the Nonlinear Effect on Wavespeed?

All materials have some degree of nonlinearity between force and deformation. The pressure–volume relationship in a gas is nonlinear with compressibility decreasing as pressure is increased. Liquids and solids are compressible to some degree, with compressibility decreasing with pressure. This effect allows the wavespeed to increase with pressure.

The result of this increase is that the high-pressure parts of a wave propagate faster than the low-pressure parts, gradually overtaking them. This results in a steepening of the wavefront, as illustrated in Fig. 43. Eventually, the wavefront becomes vertical, forming a jump. This effect is responsible for the breaking of ocean surface waves, but of course interior waves and waves in strong materials cannot break.

The wavespeed at some pressure in an ideal gas depends on the ratio of the pressure to that in the quiescent (no wave) state, as shown in the following analysis. However, this reduces to a dependence on only the absolute temperature, so that the wavespeed is a measure of temperature, and conversely the temperature determines the wavespeed.

The wavespeed in a fluid is

$$c = [(dp/d\rho)_{\mathrm{A}}]^{1/2}$$

where p is pressure, ρ is density, and the subscript A denotes adiabatic change, so that no heat flows into or out of the wave by conduction or radiation. This condition is violated at very high or low pressures. An ideal gas obeys the equation of state $PV = RT$ or $p = \rho RT$, where $V = 1/\rho$ is the specific volume, T is the absolute temperature, and R is the universal gas constant. An adiabatic process is governed by the formula $pV^{\gamma} = \text{Const}$, or $p/p_0 = (\rho/\rho_0)^{\gamma}$, where $\gamma = c_{\mathrm{p}}/c_{\mathrm{V}}$ is the ratio of specific heats, and the subscript 0 denotes the state at the start of the process. Then

$$c = c_0[(p/p_0)^{(\gamma-1)/\gamma}]^{1/2} = [\gamma p/\rho]^{1/2} = [\gamma RT]^{1/2}$$

with $c_0 = (\gamma p_0/\rho_0)^{1/2} = [\gamma RT_0]^{1/2}$, the initial wavespeed.

What Are Wavespeeds in Some Typical Materials?

Longitudinal wavespeed in some common materials is typically as follows:

Transducer crystal	3.3–7.3 mm/μsec (0.13–0.29 in./μsec)
Aluminum	6.3 mm/μsec (0.25 in./μsec)
Steel	5.6–5.9 mm/μsec (0.22–0.23 in./μsec)
Beryllium	12.9 mm/μsec (0.51 in./μsec)
Lead	2.2 mm/μsec (0.087 in./μsec)
Glass	5.3–6.8 mm/μsec (0.21–0.28 in./μsec)
Marble	6.2 mm/μsec (0.24 in./μsec)
Acrylic	2.7 mm/μsec (0.11 in./μsec)
Body tissue	1.5 mm/μsec (0.06 in./μsec)
Wood	1.4–4.8 mm/μsec (0.055–0.19 in./μsec)
Fiberglass	
Along thin sheet	1.6 mm/μsec (0.063 in./μsec)
Across honeycomb	2.7 mm/μsec (0.11 in./μsec)
Carbon phenolic	3.5 mm/μsec (0.14 in./μsec)
carbon–carbon	1.6–3.1 mm/μsec (0.06–0.12 in./μsec)
water	1.5 mm/μsec (0.059 in./μsec)
water vapor	0.4 mm/μsec (0.016 in./μsec)
alcohol	1.0–1.4 mm/μsec (0.04–0.06 in./μsec)
air (at standard temperature and pressure)	0.33 mm/μsec (0.013 in./μsec)

HOW DOES A PROPAGATING WAVE CHANGE?

A uniform plane wave in a uniform elastic material propagates indefinitely with no change. However, most materials are not purely elastic, having properties which induce changes through a variety of processes. Furthermore, in most configurations propagating waves diverge and are neither uniform nor planar, which leads to change.

A change in amplitude is called attenuation or gain (it may be a decrease or an increase). A change in the waveform results in distortion which reflects a change in the spectral content (discussed in Section 6).

What Is Attenuation, Gain (Decibels, dB, Nepers)?

A wave with reducing amplitude is illustrated in Fig. 44.

The reduction (or increase) in the amplitude of a waveform is expressed as attenuation (or gain), by the logarithm of the ratio of the magnitudes of the original to the attenuated amplitudes, a and a_0. This is commonly measured in

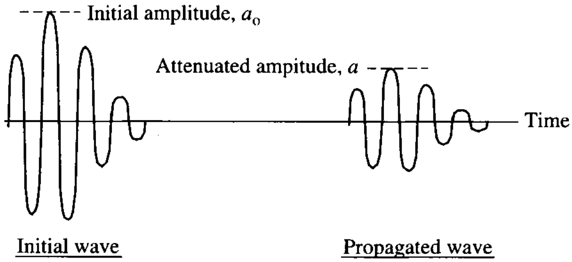


FIGURE 44 Decay of a propagating pulse

decibels (dB, tenths of bels, named after Alexander Graham Bell, inventor of the telephone):

$$A(\mathrm{dB}) = 20 \log_{10}(a/a_0).$$

The factor 20 comprises a factor of 10 for conversion from bels to decibels, and a factor of 2 introduced because in acoustics, a signal is usually characterized by its power, which is the square of its magnitude, i.e., $\log(a^2) = 2\log(a)$.

Attenuation is a negative quantity, but if the signal is amplified, the logarithm is positive and is then called gain. Since $\log(1/2) = -\log(2) = -0.3010\ldots$, an attenuation ratio of 1/2 is close to −6 dB (often referred to as 6 dB down), and since $\log(3) = 0.4771\ldots$, a ratio of 1/3 is roughly 10 dB down. Conversely, a gain by a factor of 2 is 6 dB up.

Alternatively, the natural logarithm ($\log_e$ or ln) of the ratio between the original and the attenuated amplitude is sometimes used, defining the unit of attenuation as the neper:

$$\mathrm{A(nepers)} = \log_e(a/a_0).$$

Since $\log_e(10) = 2.30\ldots$, then $A(\mathrm{Nepers}) = A(\mathrm{dB})/8.685$.

Because attenuation is measured in propagation over distance, it is usually measured as an attenuation coefficient, having units of dB/distance.

What Are the Processes in Attenuation and Waveform Distortion?

Two types of process influence propagating waves, affecting amplitude and/or waveform: those which involve material responses, and those which involve interactions between waves:

Materials response processes

- Beam spreading or focusing, called geometric attenuation
- Energy absorption, called material (or intrinsic) attenuation
- Dispersion.
- Nonlinearity

Wave interaction processes (discussed in Section 4)

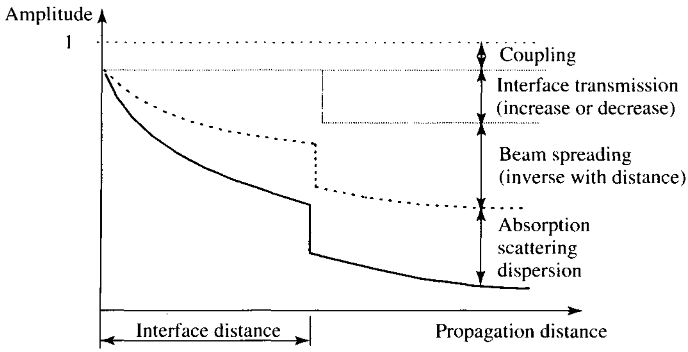


FIGURE 45 Factors in attenuation during pulse propagation

- Transmission across interfaces, called configurational attenuation
- Scattering by material variations, inhomogeneities, and defects
- The Doppler effect

In many cases, attenuation is measured in relation to the level of excitation, since the incident wave amplitude may not be known or easily measurable. This does not account for any loss between the wave exciter (the transducer, discussed in Section 5) and the surface. This loss arises through incomplete coupling and should be considered when results of attenuation as measured in different test configurations are compared.

Each attenuation process exhibits a unique behavior with regard to propagation distance, as illustrated in Fig. 45, and to frequency, which allows them to be distinguished.

What Is Geometric Attenuation?

Geometric attenuation is caused by the spreading of the wave because of its finite sized source, as discussed above (see Figs. 39–41). A source typically produces a conical beam, whose angle of divergence is determined by the size of the transducer and the wavelength. The amplitude of the beam decreases in inverse proportion to distance, with a frequency factor for the near-field distance. In a focused beam, where the source is designed accordingly, the amplitude can grow.

Geometric attenuation changes the energy content of a wave, but not the wavespeed or waveform (except for nonlinear effects).

What Is Material (Intrinsic) Attenuation?

Material attenuation occurs through internal friction, which converts kinetic energy into heat by active nonelastic responses of the material. The heat then conducts away from the wave region. These are molecular processes such as viscosity and plasticity, as well as response of the material structure such as slip at boundaries and defects. These effects are dependent on the history of stress, mainly shear (its rate of change and previous high), in contrast to elastic response which depends only on the current stress.

Intrinsic attenuation does not include scattering effects discussed later, in which the behavior of the waves is influenced by the material structure. Similarly, the effect of nonlinearity at high pressure, which generates heat, is included in a later category.

Internal friction is characterized by a dimensionless attenuation parameter Q, defined as the inverse ratio of energy lost in a cycle to the energy at the start of the cycle, or equivalently, the ratio of change in amplitude to the initial amplitude in a cycle:

$$Q = -1/\pi(\Delta A/A) = -1/2\pi(\Delta E/E).$$

The factor 2 is introduced into the energy expression since energy is the square of the amplitude, so that the change of energy is twice that of amplitude. The inverse definition results in a high Q for low energy loss, so that Q has the properties of a quality factor (a definition originally used for electrical circuits).

Evidently, if Q is independent of frequency this definition implies that the amplitude varies exponentially in distance propagated:

$$A(x) = A_0 e^{-\omega x/2Qc}.$$

In most cases this is an oversimplification so that real materials must be modeled by a frequency-dependent Q. An attenuation coefficient, $\alpha = \omega/2Qc$, gives the amplitude as

$$A = A_0 e^{-\alpha x}.$$

The attenuation in propagation over a distance, x, can then be expressed as the decibel loss:

$$\text{dB} = -8.686\ \omega x/2Qc = -8.686\ \alpha x.$$

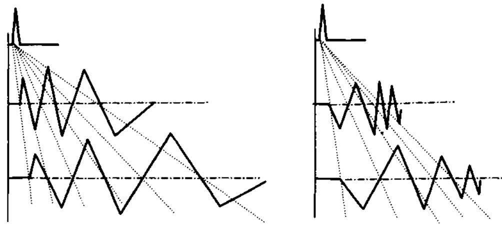
An equivalent parameter is the logarithmic decrement, δ, which is the natural logarithm of the ratio of two successive peaks:

$$\delta = \log_e(A_n/A_{n-1}) \sim \pi/Q.$$

It has been shown for certain rocks and certain classes of composite that Q is correlated to the square of the wavespeed, $Q \sim c^2$, so that higher velocity materials have a higher quality factor (i.e., less attenuation). In some cases, however, attenuation is desirable!

What Is Dispersion?

The speed of a wave propagating in a nonelastic (e.g., viscoelastic) medium or near the boundary of a body may depend on the frequency of the wave, as discussed in Appendices 11 and 12. Increasing frequency may decrease wavespeed



High wavespeed at high frequencies

High wavespeed at low frequencies

FIGURE 46 Typical cases of waveform distortion by dispersion

(as in rod waves), or increase wavespeed (as in bending). Waves whose speed depends on frequency are called dispersive.

The frequency dependence of wavespeed allows the high-frequency parts of a waveform (narrow peaks and sharp rises or drops) to move at a different speed from the shallow, smooth parts.

When high-frequency components propagate faster, they arrive earlier than the low frequencies. The late phases of the waveform become progressively more stretched out, smoothing and spreading the waveform, as illustrated in Fig. 46. Conversely, when low frequencies propagate faster, the later phases become compacted. Dispersion thus changes the waveform, including the amplitude of successive peaks, but does not generally change the energy content. Dispersion is discussed further in Appendix 7.

What Are the Effects of Nonlinearity?

Nonlinearity changes a waveform because the wavespeed can depend on the stress level, so that the high-stress parts at the peaks of a wave move at a different speed than those at low stress (some effects make the speed higher and some lower). This is mostly a factor in gases and fluids, as discussed earlier and illustrated in Fig. 43.

The effects of nonlinearity cause continuing changes in the waveform of a propagating wave. This can also increase or decrease the amplitude.

4

WAVE INTERACTIONS

HOW DO TWO OR MORE WAVES INTERACT?

A propagating wave imparts a change in state (stress and particle velocity) to the material it passes through. When two UT waves meet they interact by simply adding algebraically the changes in stress and velocity, because of elastic linearity. The waves continue to propagate with their respective state increments unchanged. Additional waves can be excited to satisfy continuity of motion and force. Although the physical principles are straightforward, the numerical method can be complicated, particularly for waves in anisotropic materials.

What Happens to Two Intersecting Waves?

As an example, the waves from a tensile load suddenly applied to each end of a rod are shown in the x–t diagram illustrated in Fig. 47. The load introduces a stress σ to each end, exciting waves which propagate into the rod which is initially in the stress-free and stationary state $A : (\sigma, v) = (0, 0)$. The waves carry jumps in stress–velocity state of $(\delta\sigma, \delta v) = (\sigma, \pm\sigma/z)$.

The right end of the rod moves to the right with a velocity $v = \sigma/z$, where z is the impedance of the rod, and the left end moves to the left at the same magnitude of velocity, but in the opposite direction. The stress–velocity state

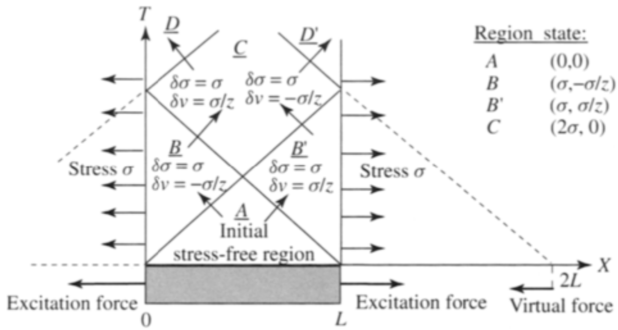


FIGURE 47 *x–t* diagram for intersection of two waves

produced by the waves in the end regions B and B' of the rod are $(\sigma, v) = (0 + \delta\sigma, 0 + \delta v) = (\sigma, \pm\sigma/z)$.

When the right-running wave from the left end and the left-running wave from the right end intersect, the stresses add, becoming 2σ, and the velocities cancel, becoming 0, i.e., the state in region C is $(2\sigma, 0)$.

These waves propagate until they reach the opposite loaded end, where they are reflected as oppositely running waves, ensuring that the boundary force is correctly represented. Note that the reflected waves can be regarded as imaginary, or virtual, waves emitted by virtual forces at virtual surfaces located outside the rod.

What Are Constructive and Destructive Wave Interference and Stationary Phase?

When waves of different waveforms meet, the parts which are in phase at one instant (i.e., positive or negative peaks all coinciding in time) add to produce large peaks, as illustrated in Fig. 48. This is called constructive interference, or the Principle of Stationary Phase. If there are as many positive as negative peaks at one instant (e.g., a random distribution of peaks) then they cancel, leaving no amplitude. This is called destructive interference.

A transducer emits expanding spherical waves from all points of its face. Near the transducer, these waves interfere destructively, forming a complex wave pattern of peaks and nulls called the near field, but at greater distances they interfere constructively, forming coherent plane waves in a conical region called the far field, which is the UT beam. This is described in Section 3 and analyzed in Appendix 8.

What Is a Group Wave?

Waves which have differing speeds can interfere constructively to form a group wave. Such waves arise in a material where the wavespeed depends on the

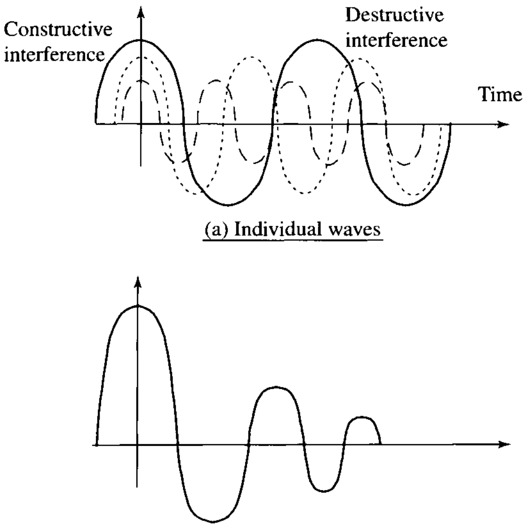


FIGURE 48 Illustration of constructive and destructive interference

frequency (called dispersion), on the amplitude (nonlinearity), or on the propagation direction (anisotropy). This results in a strong wave which propagates at a unique speed—the group velocity.

How Does a Group Wave Form?

The formation of a group wave through constructive interference of waves of various frequencies traveling in the same direction at different speeds is illustrated in the x–t diagram of Fig. 49. Lines of constant phase (multiples of π) in two waves traveling at different speeds are shown. The waves superpose to form a group wave of increased amplitude where the phases of the two waves are the same.

One of the best-known examples of group waves is the bow-wave of a moving boat, where sequences of small waves can be seen to coalesce into a large wave. In an anisotropic material, waves propagating in different directions at different speeds group together to form a wave surface as described in Appendix 9.

What Is the Dispersion Curve?

The relationship between wave speed and frequency is called the dispersion curve. Typical curves for a rod wave are given in Fig. 157 of Appendix 7, and for a bending wave in Fig. 159 of Appendix 7.

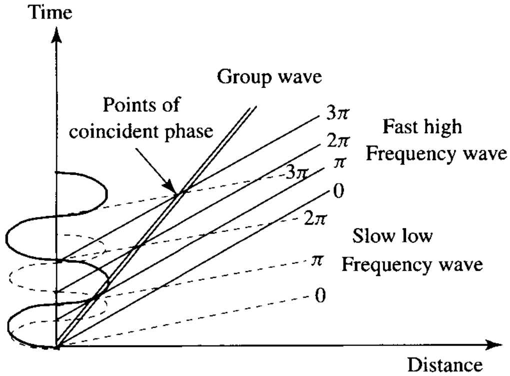


FIGURE 49 Group wave formation by constructive interference

The dispersion curve, $c = c(\omega)$, determines the group velocity through the condition of constant phase (discussed in Appendix 7) as

$$c_g = c/[1 - (\omega/c)(dc/d\omega)].$$

When the wavespeed is independent of frequency, i.e., when $dc/d\omega = 0$, the group velocity is the same as the wavespeed. When the wavespeed decreases with frequency so that $dc/d\omega < 0$, the group velocity is less than the wavespeed, and conversely. Evidently the group velocity in a rod wave is less than the wavespeed, but in bending it is faster.

Group velocities for a rod are given in Fig. 158 of Appendix 7, and for a beam in Fig. 160 of Appendix 7.

Waves in an anisotropic material travel at different speeds in different directions as illustrated in Fig. 165 of Appendix 9. At any time two neighboring wavefronts meet at a common point where they interact constructively, creating a group wave. Analysis of this situation is given in Appendix 9. These waves are very important to wavespeed measurements in anisotropic materials as they are in fact the observed waves.

An example of angular group velocities is shown in Fig. 50 for the laminated composite material considered in the plane wave analysis shown in Fig. 29 in Section 3.

How Does a Wave Interact with a Surface?

The surface of an object may be a free surface with nothing behind it, or an interface between two different materials. The surface may be flat or curved, and it may be inclined relative to the incident wave.

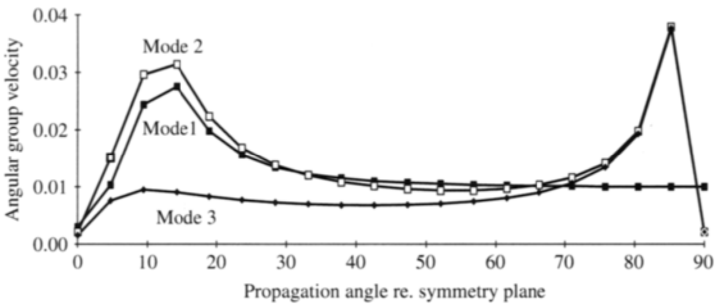


FIGURE 50 Angular group velocities for fiber reinforced plastic of Figs. 29 and 30.

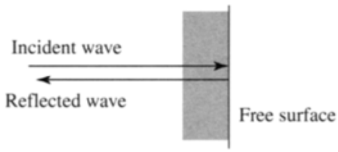


FIGURE 51 Normal reflection at a free surface

How Does a Wave Reflect from a Flat Free Surface at Normal Incidence?

The simplest surface interaction occurs when a plane wave meets a flat free surface at normal incidence, i.e., parallel to the wavefront as shown in Fig. 51.

The x–t diagram for a right-running wave normally incident on a free surface is shown in Fig. 52. The wave induces a stress-velocity state $(\sigma, v) = (\sigma_I, \sigma_I/z)$ into the medium, where σ_I is the incident stress and z is the impedance of the material. The stress cannot be supported when the wave reaches the free boundary, so a new left-running wave is induced with a stress jump $\delta\sigma = -\sigma_I$ to drop the stress to 0. This induces a velocity jump $\delta v = \sigma_I/z$ (the sign of the impedance, $z = \rho c$, changes with the direction of the wave, i.e., the sign of c). The velocity of the material behind the reflected wave, and of the free surface, is $v + \delta v = 2\sigma_I/z$.

The reflected wave can be thought of as a virtual wave which initiates at a point outside the surface, but propagates as though it were in the material.

Direct backward reflection off a plane reflector (e.g., a metal plate) is often used in water tank testing (see Section 5), so that rays traverse the specimen twice, and the received wave is delayed from the transmitted wave by the additional water path.

How Does a Wave Transmit Across and Reflect from an Interface at Normal Incidence?

A plane wave incident normally onto a flat interface between two materials is shown in Fig. 53. The wave creates a reflection from, and a transmission across

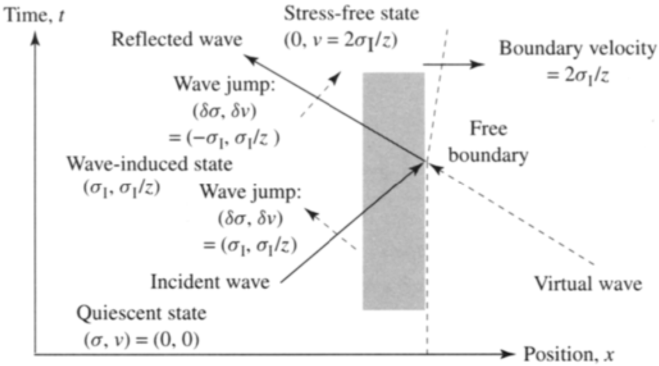


FIGURE 52 x–t diagram for normal reflection at a free surface

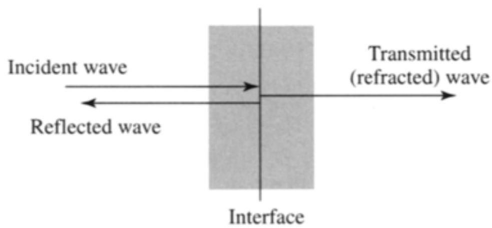


FIGURE 53 Normal reflection and transmission at an interface

the interface. The x–t diagram for a right-running wave normally incident on an interface between two materials of differing properties is shown in Fig. 54. The wave in material 1 induces a state $(\sigma_I, \sigma_I/z_1)$, where z_1 is the impedance of material 1. A wave is transmitted into material 2, developing the state $(\sigma_T, \sigma_T/z_2)$, where z_2 is the impedance of material 2. The stresses and velocities on either side of the interface must balance, but this cannot be if there is only the incident wave in material 1, because the impedances differ. This creates a new wave, reflected back into the first material, with stress jump $\delta\sigma_R = (\sigma_T - \sigma_I)$, and velocity jump $\delta v_R = -\sigma_R/z_1$. A balance of stresses on both sides of the interface requires that $\sigma_I + (\sigma_I - \sigma_R) = \sigma_T$, i.e., $\sigma_T = 2\sigma_I - \sigma_R$. The continuity of velocity requires that the velocities on both sides of the interface must be the same:

$$\sigma_I/z_1 - (\sigma_T - \sigma_I)/z_1 = \sigma_T/z_2$$

so that

$$\sigma_T = [2z_2/(z_2 + z_1)]\sigma_I,$$

$$v_T = [2z_1/(z_2 + z_1)]v_I,$$

and

$$\sigma_R = [(z_2 - z_1)/(z_2 + z_1)]\sigma_I,$$

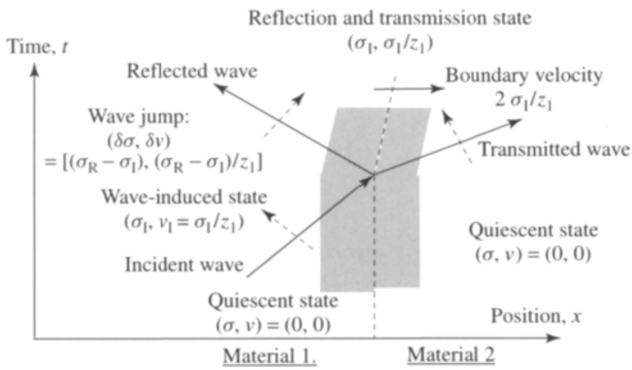


FIGURE 54 x–t diagram for normal reflection and transmission at an interface

$$v_R = -[(z_2 - z_1)/(z_2 + z_1)]v_I.$$

These expressions define stress coefficients of transmission and reflection, respectively,

$$C_{\sigma T} = 2r/(r+1), C_{\sigma R} = (r-1)/(r+1)$$

and velocity coefficients

$$C_{vT} = 2/(r+1), C_{vR} = -(r-1)/(r+1)$$

where $r = z_2/z_1$ is the impedance ratio.

When the impedance of material 2 is much less or much greater than that of material 1:

$$z_2 \gg z_1, \text{ then } \sigma_T/\sigma_I \Rightarrow 0, \text{ and } \sigma_R/\sigma_I \Rightarrow -1,$$

$$z_2 \ll z_1, \text{ then } \sigma_T/\sigma_I \Rightarrow 2, \text{ and } \sigma_R/\sigma_I \Rightarrow 1.$$

Also, when $z_2 = z_1$, there is no reflection and the transmitted wave is of the same strength as the incident, as it should be since the interface is not a mechanical one (i.e., there is no change in impedance or wavespeed).

The first of the foregoing results is the same as that given in the paragraph above, while the second shows that the wave doubles on transmitting into a very stiff or dense material, and the reflection off such an interface is of the same sign as the incident wave.

What Is Impedance Matching?

When a wave is propagated into a material from another of different properties, the transmission ratio can reduce the transmitted wave amplitude. For example, transducer crystals typically have much higher impedance than the materials they are used to test. To minimize this effect, a layer of intermediate impedance is imposed between the two materials as illustrated in Fig. 55.

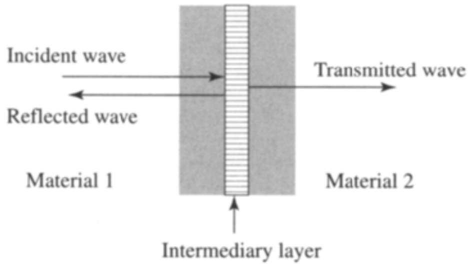


FIGURE 55 Impedance matching layer

The stress transmission coefficient between the two layers with no intermediary layer is

$$T_{12} = 2z_2/(z_1 + z_2)$$

With the intermediary layer, the overall transmission coefficient is the product of the transmission coefficients for both interfaces:

$$T_{12} = [2z_i/(z_1 + z_i)]\ [2z_2/(z_i + z_2)]$$

where $z_i =$ is the impedance of the intermediary layer. For optimum transmission, the impedance of this layer is found (by differentiation) to be the geometric mean of the other two (see Appendix 10):

$$z_i = (z_1 z_2)^{1/2}.$$

This improves the transmission by a factor

$$T_{max}/T_{nom} = 1 + [(z_1^{1/2} - z_2^{1/2})/(z_1^{1/2} + z_2^{1/2})]^2$$

which approaches a maximum of 2 when $z_1 \gg z_2$, where $T_{nom} = 2(z_1/z_2)$ is the nominal transmission without the intermediary layer.

What Is a Quarter-Wave and a Half-Wave Layer?

The thickness of an intermediary layer must be considered because the reflected wave within the layer interferes with the incident wave. The wave number in the intermediary layer is $k_i = 2\pi/\lambda_i = 2\pi f/c_i$, where λ_i is the wavelength in the layer at the frequency f, and c_i is the wavespeed. The phase of the reflected wave relative to the incident at the first interface, when it returns to the first interface, is $\delta\phi = 2kh$ as illustrated in Fig. 56, having traveled a distance of twice the thickness h.

Destructive interference occurs when the reflected wave is out of phase with the incident, i.e., when $\delta\phi = \pi$. This occurs when the thickness is a quarter wavelength, $h = \lambda/4$, or the frequency is $f = c/4h$. This configuration is called the quarter-wave filter, as it blocks waves at that frequency with that wavelength.

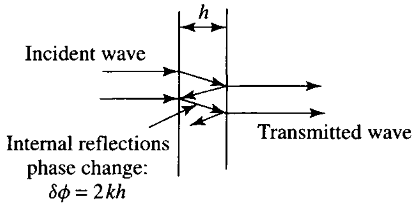


FIGURE 56 Interference between incident and internal reflections in intermediary layer

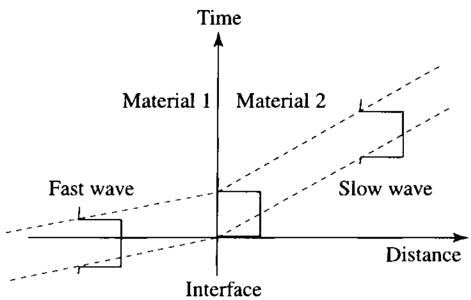


FIGURE 57 Waveform is independent of wavespeed

Conversely, constructive interference occurs when the phase difference is $\delta\phi = 2\pi$, so that maximum transmission occurs with a half wavelength, at the optimum frequency $f = c_i/2h$.

What Happens to a Waveform at an Interface?

As a wave propagates across an interface between two media, it is changed in both amplitude and speed. The speed is determined by the properties of the medium and the direction and type of wave, as well as by the characteristics of the incident wave.

A change of wavespeed does not change the shape of the wave, as illustrated for a square wave in Fig. 57, nor does it change its energy—it can continue to propagate indefinitely. This implies that there would be no change in any other waveform. The amplitude is determined by the ratio between the impedances of the two materials, as discussed below.

What Is Oblique Reflection and Refraction?

A wave meeting an interface at some angle other than normal results in oblique reflections and transmissions. Oblique transmission is also called refraction, in which a transmitted wave has a different angle than the incident. This refraction

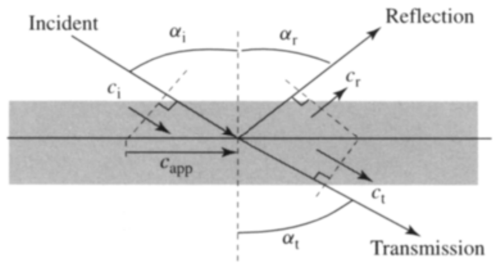


FIGURE 58 Reflection and refraction at an oblique interface

process, illustrated in Fig. 58, bends the rays and allows a variety of applications such as focusing, which are discussed later.

What Is Snell's Law?

As in normal reflection, the stress and particle velocity states of the waves must balance at the interface, but in addition, continuity of motion along the interface must be assured. This requires that the apparent wavespeed along the interface created by the incident wave spreading over it, $c_{app} = c_i / \sin \alpha_i$, where i denotes the incident wave, c_i is its wavespeed, and α_i is its angle to the interface normal, must be the same for all the waves. Hence the angles of the waves are related by Snell's law (see Appendix 10):

$$\sin \alpha_i / c_i = \sin \alpha_r / c_r = \sin \alpha_t / c_t$$

where r denotes the reflection, and t the transmission (or refraction). The angles of incidence, reflection, and transmission are measured between the normal to the surface and the direction of propagation.

Thus for transverse reflection of an incident longitudinal wave:

$$\sin \alpha_s = (c_s / c_l) \sin \alpha_l$$

where the subscript s denotes a shear wave and l a longitudinal wave. For the converse, a longitudinal reflection of an incident shear wave,

$$\sin \alpha_l = (c_l / c_s) \sin \alpha_s.$$

The angle of reflection of a wave of a like kind is the same as that of the incident wave, since they are in the same material and have the same wavespeed. However, a P- or an S-wave reflected from its opposite, an S- or a P-wave, has a different angle because the wavespeeds are different. Transmitted waves are generally at different angles also.

Reflection of a P-wave into an S-wave results in the reflection angle being closer to the normal than the incident wave, and vice-versa.

The relationship between the angles of P- to S-wave reflection in an isotropic material is given in Fig. 59 and analyzed in Appendix 10. The relationship for S- to P-wave reflection is illustrated in Fig. 60.

The concepts of Snell's Law used in optics to design curved mirrors and lenses can be applied to UT. A curved mirror or lens could be designed to collimate or focus a divergent beam, but problems arise because of the size of a typical transducer source in relation to wavelength except at very high frequency.

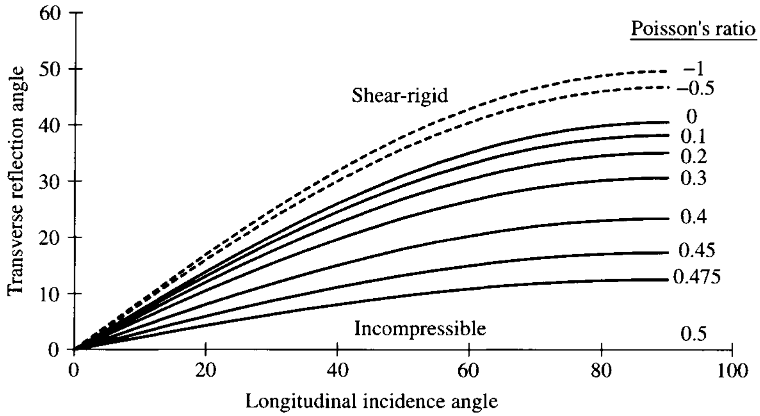


FIGURE 59 Angles of transverse reflection from a longitudinal wave in an isotropic material

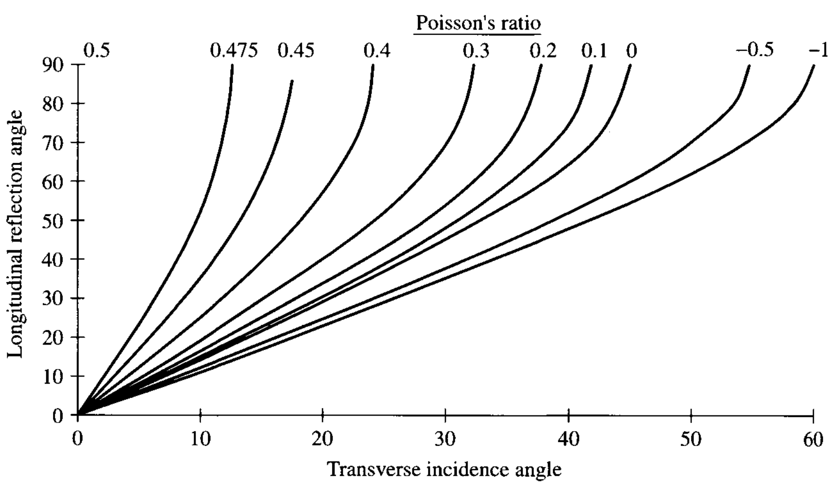


FIGURE 60 Angles of longitudinal reflection from a transverse wave in an isotropic material

What Is the Critical Angle for Reflection?

A feature of Snell's law is that, since the sine of the reflection angle must be less than 1, a wave incident at a large angle cannot be reflected as a faster wave, i.e.,

$$\sin\alpha_r \leq 1 \text{ only if } \sin\alpha_i \leq (c_i/c_r)$$

There is a limiting or critical incidence angle, $\alpha_{i\text{-}cr} = \operatorname{asin}(c_i/c_S)$, for longitudinal reflection of a transverse wave, such that

$$c_i/c_r = c_S/c_l = [(1-2\nu)/2(1-\nu)]^{1/2}.$$

The angle of incidence beyond which transverse-to-longitudinal reflection is not possible is then

$$\alpha_{i\text{-crit}} = \operatorname{asin}[(1-2\nu)/2(1-\nu)]^{1/2}.$$

This is plotted in Fig. 61. Evidently for a liquid, which is equivalent to a Poisson's ratio of 0.5, there is no critical reflection.

What Is the Refraction Angle of Transmitted Waves?

The angle of a transmitted wave is determined by the wavespeeds in the materials on both sides of the interface and by the types of the incident and transmitted waves. From Snell's law given above,

$$\sin\alpha_t = \sin\alpha_i(c_t/c_i).$$

When the transmitted wavespeed is smaller than the incident, the refracted angle is less than the incident, i.e., the transmitted wave is closer to the normal than the incident, and conversely.

These results could be used to design a UT lens which would collimate a divergent beam or focus a beam using a procedure similar to that in optics.

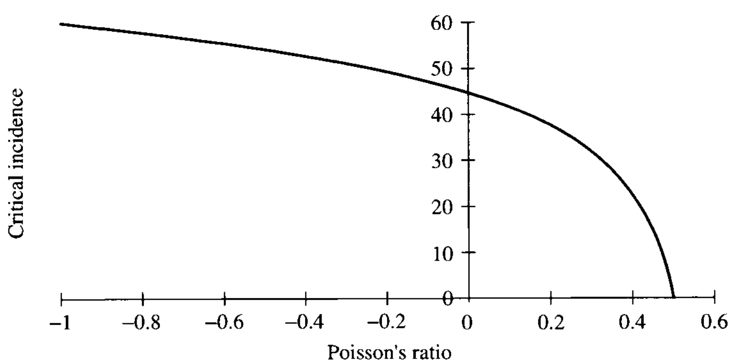


FIGURE 61 Critical angle for longitudinal reflection of a transverse wave

What Is a Grazing Incidence (Leaky) Lamb Wave?

For a transmission to be possible, so that $\sin\alpha_t \leq 1$, the incident angle must be less than $\alpha_i \leq \alpha_{i\text{-cr}} = \text{asin}(c_i/c_t)$. The transmission wavespeed is not related to the incident, falling in a different material. When this condition is violated because of a large angle of incidence, no transmission is possible. At the critical angle, the wave in the second material runs along the interface similar to a surface wave. The condition is called total internal reflection at grazing incidence. The interface wave emits waves similar to a reflection back into the first material at the critical angle, called leaky reflections. This wave system, called a leaky Lamb wave, is illustrated in Fig. 62. The leaking reflections can be monitored at various distances to detect the Lamb wave. Again, there can be no such wave in a liquid.

What Are the Reflection Coefficients at an Oblique Free Surface?

At a stress-free surface oriented obliquely to an approaching wave, all stresses with components normal to the surface must cancel. The incident and reflected waves each carry a normal and a tangential stress with components normal to the surface. A single reflection cannot satisfy this requirement, so that two reflections, one longitudinal and one transverse, must be created, as shown in Fig. 63.

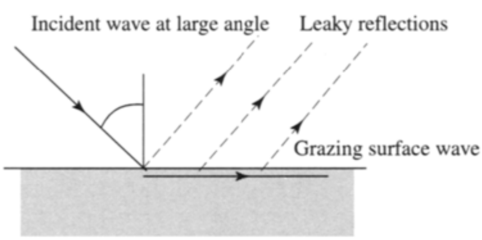


FIGURE 62 Leaky Lamb waves: critical refraction at grazing incidence

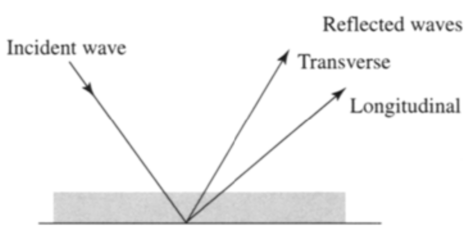


FIGURE 63 Oblique reflection at a free surface

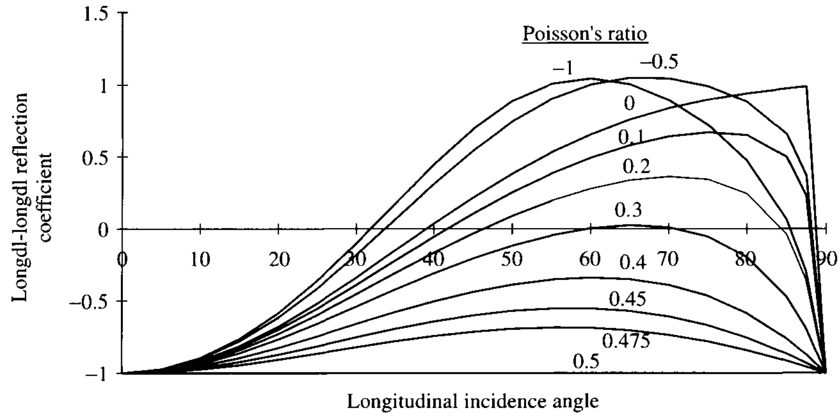


FIGURE 64 Reflection coefficient for a longitudinal wave from a longitudinal wave

There are then two equations, one for the normal and one for the shear stress, in the two unknown amplitudes of the longitudinal and transverse reflections:

$$\sigma_{yy\cdot i} + \sigma_{yy\cdot rl} + \sigma_{yy\cdot rx} = 0$$

$$\sigma_{xy\cdot i} + \sigma_{xy\cdot rl} + \sigma_{xy\cdot rx} = 0$$

where r refers to reflection, l to longitudinal, and x to transverse. The magnitudes of these reflections from a longitudinal or a transverse wave are analyzed in Appendix 10.

In an anisotropic material, there can be two quasi-transverse reflections which do not lie in the plane of the incident wave. In that case, there are three equations for stress balance which generally require numerical solution.

The results for the ratio of the longitudinal reflection stress to the incident stress is given as functions of the incident angle and Poisson's ratio in Fig. 64.

What Is Mode Conversion?

The generation of one type of wave (longitudinal or transverse) from another in reflection or transmission is called mode conversion. The stress ratios from conversion of a longitudinal wave into a transverse one is shown in Fig. 65. Similar results for an incident transverse wave are shown in Figs. 66 and 67. The coefficients shown are all strongly dependent on Poisson's ratio.

What Is the Interaction at an Oblique Interface?

Interaction of a wave at an oblique interface excites several waves to ensure the balance of stresses and velocities across the interface: an incident P- or S-wave in an isotropic material excites both P- and S-waves in reflection and in transmission into an isotropic material, as illustrated in Fig. 68.

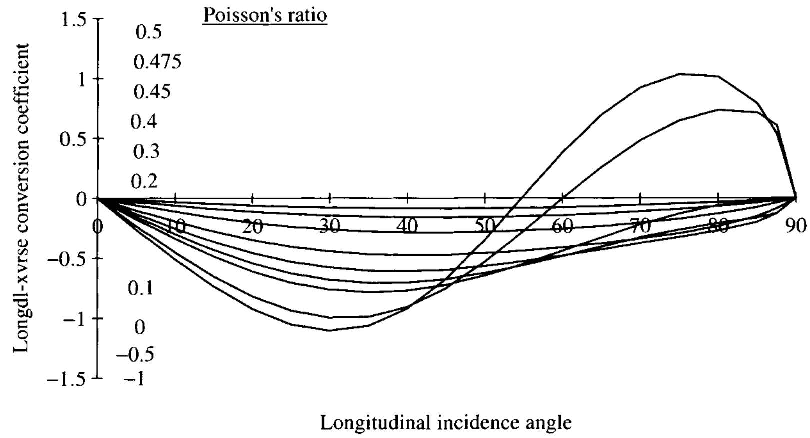


FIGURE 65 Conversion coefficient for a transverse wave from a longitudinal wave

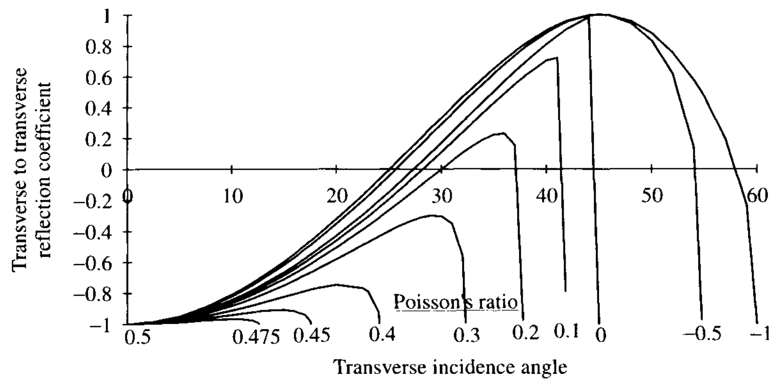


FIGURE 66 Reflection coefficient for a transverse wave from a transverse wave

The magnitudes of the four or six waves can be determined by requiring that the components of stress in the normal and one or two transverse directions, and of velocity normal and in one or two tangential directions to the interface, must balance on both sides. This leads to four (or six) equations of the type

$$\sigma_i + \sigma_{rl} + \sigma_{rx} = \sigma_{tl} + \sigma_{tx}$$

where r refers to reflection, t refers to transmission, l to longitudinal, and x to transverse. These equations require numerical solution as discussed in Appendix 10.

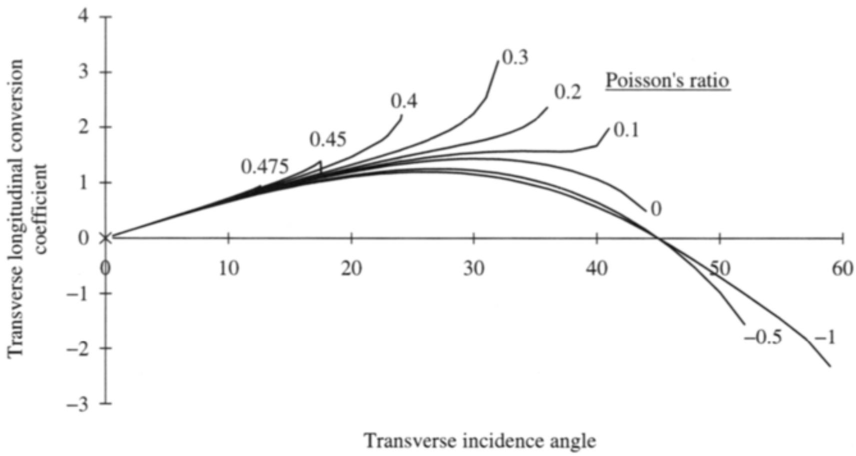


FIGURE 67 Conversion coefficient for a longitudinal wave from a transverse wave

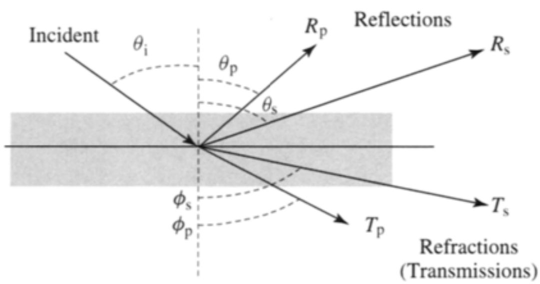


FIGURE 68 Mode conversions in oblique incidence

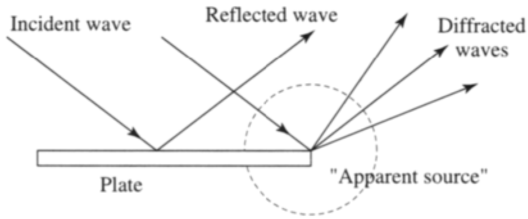


FIGURE 69 Scatter by diffraction at the edge of a plate

What Is Diffraction?

When a wave is incident on the edge of a surface, as illustrated in Fig. 69, it excites waves in all directions called diffracted waves. The edge of the plate serves as a source of diffraction.

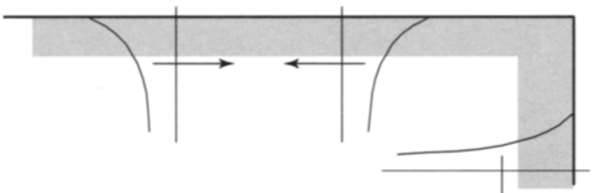


FIGURE 70 A Rayleigh wave at a corner

What Happens to a Rayleigh Wave at a Corner?

When a Rayleigh wave propagating along a surface reaches a corner, as in Fig. 70, it is partly reflected and partly transmitted around the corner.

HOW DOES A WAVE INTERACT WITH A NONUNIFORM OBJECT?

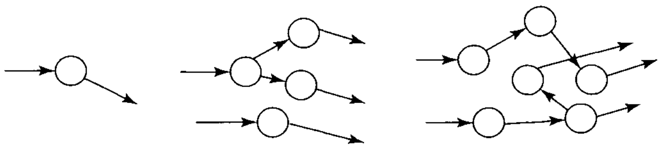
All real materials have nonuniformities (nonhomogeneities) on various scales. Real materials can be called perturbed media, in contrast to the idealized homogeneous media of analysis, which are called unperturbed. A nonuniformity is called a scatterer or a perturbation.

A nonuniformity is a region with properties which differ from those of the surrounding material. Its surface may be curved, or it may have corners or edges, and it may be moving. A nonuniformity may be an individual inclusion (a small closed region of different properties, such as a bone or a hole), or it may be an aggregated assembly such as the fibers in a reinforced plastic.

Metals and geologic materials have grain structure, inclusions, and porosity on a submillimeter scale and crystal imperfections on a micron scale. Fiber reinforced composites have fiber bundles on a submillimeter scale and laminations on a millimeter scale. Biological materials (flesh and bone, wood, etc.) are porous, fibrous, and layered (e.g., growth rings) on millimeter scales. Large structures, both natural (trees, rocks) and man-made (concrete), have discontinuities and gradations on scales from centimeters to meters. Fluids contain microscopic bubbles and particulates, which are in random Brownian thermal motion.

What Is Scattering?

The interaction of a wave with a nonuniformity produces reflections and transmissions in multiple directions, called scattered waves, in a pattern called the



(a) Single scatterer (b) Multiple scatterer (c) Random scatterer

FIGURE 71 Varieties of scattering processes

scattering pattern. A detailed but complicated description of scattering is given by Ishimaru (1978). A simplified view follows.

A wave can experience a variety of scattering types, depending on the distribution of scatterers, as illustrated in Fig. 71.

After scattering, constructive interference creates a wave which propagates in the same direction as the incident wave, called the forward scattered wave. The wave pattern propagating backward toward the source is called the backscatter.

The response of a scatterer to an incident wave can arise through deformation producing an omnidirectional scatter, or by motion as a whole producing a directed scatter. These effects can be regarded as created by sources of motion which are proportional to the differences in elastic properties and density between the scatterer and the unperturbed medium.

When the wavelength is large compared to the nonuniformity, and the nonuniformity is distributed over a large region, the effect is a small change in amplitude. At shorter wavelengths the effect of scattering is to produce a change in the waveform by the generation of new waves, as well as a change in amplitude. Wavelengths short in relation to the size of the nonuniformity respond to the discrete interfaces.

UT frequencies are chosen such that the size and distribution of nonuniformity represent typically no more than 0.1 to 1 wavelengths.

Typical scatter patterns for a single sphere, excluding the incident wave, are illustrated in Fig. 72. When the wavelength is small compared to the sphere diameter, most of the incident wave is reflected as backscatter, with little refraction in the direction of propagation. When the wavelength is large, the reflection is diminished by interference between the waves, but the transmission is enhanced into a considerable forward scattered beam.

The ratio of total scattered energy to the incident energy, called the scattering cross-section, is used to characterize the scatterer. It depends on the size and shape of the scatterer, and on its properties.

What Is Cavitation?

An oscillatory wave interacts with dissolved gases in a liquid, which exist as microscopic bubbles (normally in equilibrium between their internal vapor pressure, the external fluid pressure, and surface tension) by causing the bubbles to

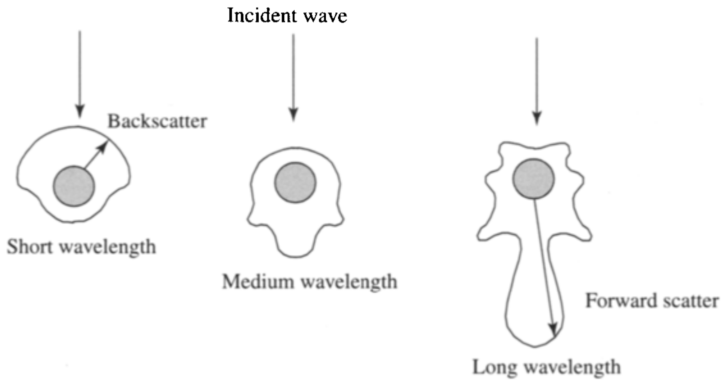


FIGURE 72 Typical scatter patterns for a sphere

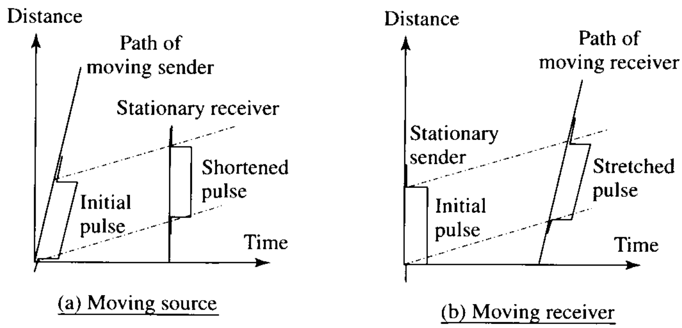


FIGURE 73 Illustration of the Doppler effect

grow or shrink. This process, called cavitation, is dynamic because of inertia of the liquid. The motions are unstable when the bubble is near a solid surface so that the bubble collapses inward in a jet, inducing high local pressure on the surface and emitting waves and sometimes light. The jet impacts on the surface, causing damage.

Cavitation can be a peril as it destroys the surface of propellers, pipes, etc but it can be beneficial when used for cleaning and as a stimulus for chemical processes. It can initiate dangerous conditions in medical applications.

What Is the Doppler Effect?

When a source of waves and a receiver move relative to each other, the waveform is distorted, as illustrated in Fig. 73. When a wave originates on a moving object as at (a) in the figure, the apparent wave seen by a stationary receiver is shortened because of the difference in path length between the early and late parts of the wave. The converse, shown at (b), arises for a moving receiver of waves from a stationary source.

The phase of a wave is $\phi = kx - \omega t$. When the distance is reduced because the source or receiver moves at a velocity v, the phase becomes $\phi = k(x - vt) - \omega t = kx - (kv + \omega)t$. Then an effective frequency is $\omega' = kv + \omega$, so that $\delta\omega = kv$. Writing $\omega = 2\pi f$ and $k = 2\pi/\lambda = 2\pi f/c$ leads to the result

$$\delta f/f = v/c$$

where f is the frequency, v is the relative velocity, and c is the wavespeed.

The change in frequency in the oscillating components of a wave is reflected in a change of pulse length and increases with the speed of the relative motion between sender and receiver.

The velocity of the relative motion can be measured by measuring such shifts in the frequency. This concept is used to measure blood flow in arteries.

5

HARDWARE: EQUIPMENT CONCEPTS

What Is a Typical UT Setup?

A UT setup comprises several pieces of hardware equipment, both electronic and mechanical. The core of a system is one or more transducers, which convert electrical signals into mechanical, and vice versa. The transducer (or transducers) is supplied with energy from a signal generator/pulser and a data acquisition system consisting of a signal detector, an analyzer, and a recorder, e.g., tape or disk drive. It is (they are) carried in some mechanical device, usually hand-held in medical and some mechanical applications, or a robotic system, typically a bridge or gantry which moves the transducer(s) across an object. The test object is often immersed in a water tank to facilitate transmission of waves into it. The parts of a system which contains these devices are illustrated in Fig. 74. Because the electrical signals are of high frequency, all connections between transducers and the electronic equipment must be coaxial cabling.

How Do Transducers Work?

A transducer is a device which converts one form of energy into another: here electrical current into mechanical motion, and vice versa.

There are several types:

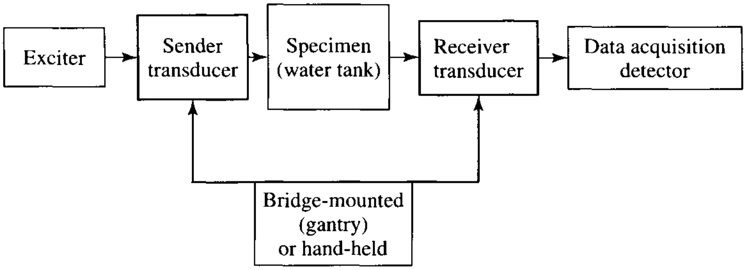


FIGURE 74 Illustration of a UT setup

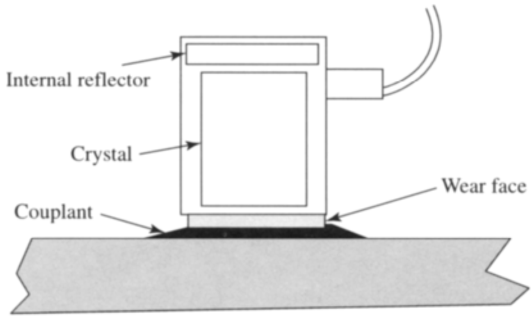


FIGURE 75 Typical construction of a crystal transducer

- Piezoelectric crystals
- Polarized plastic film
- Lasers
- Electromagnetic acoustic transducers (EMATs)

Piezoelectric Transducers

The piezoelectric (pz) effect creates a mechanical stress in a pz material, usually a crystal, when an electric field is applied as a voltage across it, or conversely, it creates a voltage when a mechanical stress is applied. An oscillating voltage produces an oscillating stress which excites propagating oscillatory waves as analyzed in Appendix 5.

The transducer contains a resonant crystal, a coupling or wear-face, and a case which includes a reflector, as illustrated in Fig. 75.

Piezoelectric crystals are commonly lead zirconate (PZT), barium titanate, lithium niobate, etc. The crystal resonance is determined by its size and shape in relation to the excitation. The reflector is a disk of suitable material which reflects waves back into the crystal, thereby reinforcing resonance.

Typical transducers of 0.5 to 2.0 MHz use crystals of about 0.5 to 1.0 in. in diameter. Low-frequency transducers are harder to make as they require large crystals, over 2 in. in diameter, and require large amounts of power to excite. High-frequency transducers are sometimes made of small deposits of pz material.

Shear waves can be excited by transducers with crystals which oscillate sideways. These require special coupling, such as adhesive bonding of the transducer to the test object. Alternatively, shear waves can be excited by using wave conversion in oblique interface transmission. Oblique waves are excited by shoe transducers (described later).

The crystals and their bonding to the case and faces are delicate and will be damaged if dropped. They use high voltage which can be hazardous.

Polarized Plastic Film Transducers

PVDF (polyvinylidene difluoride) is a polarizable polymer, which means that when it is rolled into thin sheets in a certain way, its molecules become electrically charged. Deformation, such as thinning or bending, changes the charge distribution, and these changes can be detected as capacitance changes by electrical circuits connected to metallized conductors on the faces of the sheet.

These sheets can be of any size, and they can be draped over curved surfaces. Many separate transducers can be built onto a sheet by appropriate design of the metallization, to make an array of transducers. One drawback is that they are sensitive to temperature changes, but compensation techniques can be developed using a second inactive (uncoupled) set of transducers placed alongside the active ones. They are cheap and can be mounted permanently.

Lasers

A laser when used as an exciter produces an intense but very brief pulse of light which heats a spot on the surface of an object under test. The sudden heat generates a thermal stress which excites a set of waves, as discussed in Appendix 5.

A laser used as a receiver produces a steady low-intensity beam which is split into two to serve as an interferometer: one beam (the primary) is reflected off a surface of the object under test, and the other (the reference) covers a fixed path among mirrors. The beams are brought together and the length of the reference path is adjusted so that the two coalescing beams oppose (interfere with) each other, leaving a spot of darkness. When the surface moves, the primary path changes length, so that the interference between the beams is destroyed and some light results. A photocell at the spot detects the intensity of the light, which is proportional to the motion of the surface. This concept is called an interferometer, of which there are many types; some measure displacement, and some measure velocity.

Electromagnetic–Acoustic Transducers, EMATs (for Electrically Conductive Test Pieces)

An EMAT uses the electromagnetic field created by a coil to excite an eddy current in an object of conducting material. The electric induction effect produces a shear stress which excites transverse waves. The coil can be placed a short distance from the surface, allowing its use on heated parts, or in locations where accessibility is difficult, or when a fixed mounting is undesirable.

What Is a Couplant ?

Because the surface of most objects is rough on a microscopic scale, the transducer makes contact on only a small area. Consequently the transmission of force between the surface and the transducer is severely reduced. (This problem in mechanical contact is unique to UT because the stresses and motions are so small that they do not flatten the irregularities as in engineering surface contacts.) To improve this, a couplant is always used. Common couplants are water, grease, or a soft polymer. An oil or a gel is used in medical applications.

For lasers, the roughness scatters the incident light and reduces the amount absorbed. To improve the energy transfer, a light-absorbant coating such as carbon black is required.

What Is an Immersion Tank and a Gantry?

For large objects of varying shape, coupling is often made through water by immersing the object in a tank of water. A transducer is (or two transducers are) mounted on a probe (or coupled probes) projecting into the water and carried on a gantry outside the tank as illustrated in Fig. 76. If only one transducer is used, a fixed reflector plate is mounted behind the test object. The gantry is operated robotically to move the transducer assembly over the object.

Immersion testing provides the advantages of good and uniform coupling over the test object, eliminating transducer contact (good for curved surfaces), and imposing a delay path (the water) between the transducer and the object. This separates the emitted and received signals and allows identification of pulses through the known front-face echo which passes only through water.

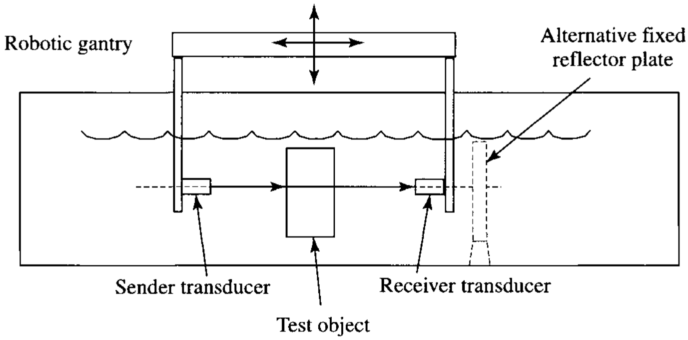


FIGURE 76 Typical immersion tank setup

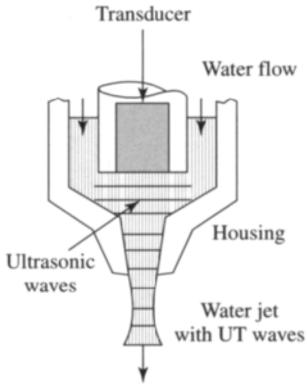


FIGURE 77 Schematic of water squirter

What Is a Water Squirter?

When a surface is not easily accessible or is not flat, a water squirter is sometimes used. This consists of a housing which holds a transducer in a chamber through which water flows. The water flows out of a nozzle under pressure, forming a jet which transports the UT waves onto the test object, as illustrated in Fig. 77. It is essential that a smooth steady flow be established.

What Is a Bubbler?

A bubbler is a device similar to a squirter but with a smaller and slower flow; it is used close to the surface.

What Is a Roller or Wheel Transducer?

A rolling wheel transducer consists of a wheel several inches in diameter, which has a liquid-filled rubber tire to provide coupling for waves from a transducer mounted internally on the axle, as illustrated in Fig. 78. These transducers are used in an automated system to cover large areas of an object.

How Are Transducers Set Up?

The major transducer configurations are as follows:

a. Pulse-echo (PE)
b. Through-transmission ultrasonics (TTU)
c. Pitch–catch
d. Angle shoe
e. Beam steering
f. Focusing horn

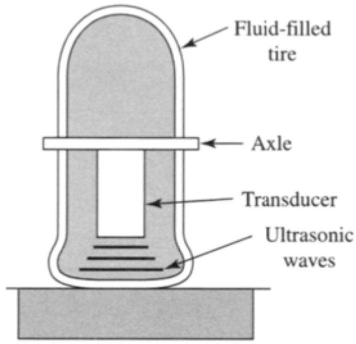


FIGURE 78 A roller, or wheel, transducer

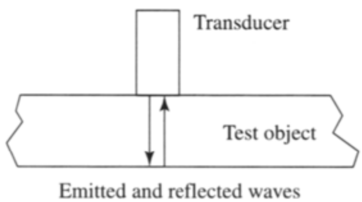


FIGURE 79 Pulse-echo concept

g. Fluid coupled lens
h. Bending waves

a. Pulse-Echo (PE)

In PE, a wave is emitted from a transducer on a surface of an object to be examined, and its echo (i.e. a reflection from another surface) is detected by the same transducer at a later time, after the wave has passed through the body and returned, as illustrated in Fig. 79.

This operation is viable only for use on relatively thick specimens such that the waves operate in the far field (described in Section 3) and the echo returns to the transducer after the excitation pulse is complete. Controlled damping in the circuitry and natural damping in the transducer improves the separation in pulses. When adequate separation is achieved, this setup has the advantage of a double transit, thus sampling the specimen twice. In addition, it has the benefit of requiring only single-sided access.

The transducer can be set back from the surface as in a water tank or a squirter setup. In a water tank a reflector plate can be placed behind the test object so that the wave transmitted through the object is reflected back for a second transit and a return to the transducer.

A form of PE is used for Doppler measurement of blood flow in a human body, as illustrated in Fig. 6. A transducer placed at an angle to the skin transmits signals into a vein or an artery. A shift in frequency of the backscattered signal measures the speed of flow.

b. Through-Transmission Ultrasonics (TTU)

In TTU, a wave is emitted from a transducer on one surface of an object to be examined and is detected by a second transducer on another surface, at a later time after the wave has passed one way through the body. This is shown in Fig. 80. The emitter and sender transducers can be set off the surface as in a water tank or a squirter.

This technique is useful when the object under test is thin or has a complex structure (such as a sandwich construction) which results in a complex echo. It requires double-sided access.

A form of TTU can be used as a Doppler flow meter for fluid in a pipe: two pairs of transducers are mounted on each side of the pipe, as illustrated in Fig. 81. One pair examines waves propagating toward the flow, and the other examines those propagating against the flow. Both waves experience Doppler shifts which can be measured spectrally.

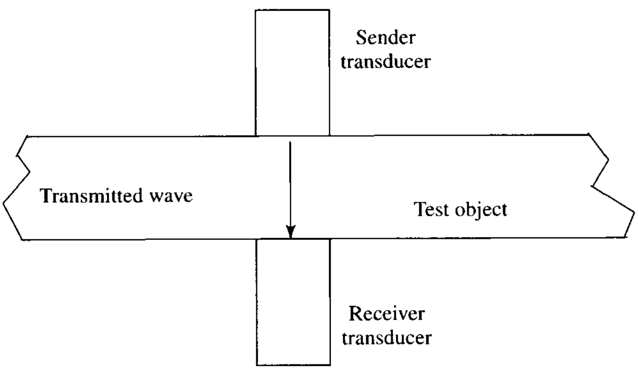


FIGURE 80 Through-transmission ultrasonics (TTU) concept

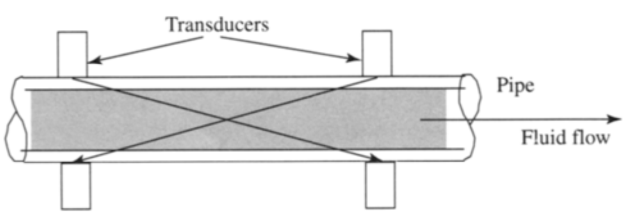


FIGURE 81 TTU Doppler flow meter

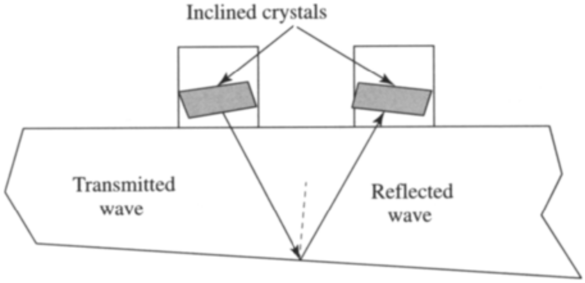


FIGURE 82 Pitch–catch arrangement

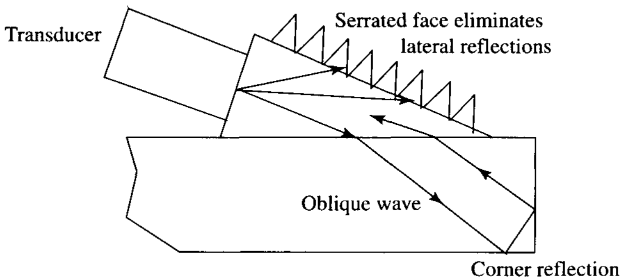


FIGURE 83 Angle shoe for exciting oblique or shear waves

c. Pitch–Catch

Two transducers are placed on the same side of an object, spaced a short distance apart. The sender excites a wave at a small angle to the normal, so that the reflection returns to the second transducer, as illustrated in Fig. 82.

This technique is a special case of a pulse-echo system. It can be useful when the back surface is not parallel to the front or in thin materials where it allows the excitation (main bang) to be isolated from the reflection signals.

d. Angle Shoe

Oblique waves are excited and detected by transducers placed at a large angle to the normal. To achieve this, a shoe is used, as illustrated in Fig. 83. A common transducer is attached to the shoe, whose material must be selected to address both Snell's Law (the wavespeed determines the transmitted wave angle, see Section 4) and the impedance, to ensure adequate transmission.

A special application of the angle shoe is to the inspection of a corner, such as around the periphery of a bolt hole. Double reflection assures that the wave reaches the transducer from any angle, as illustrated in the figure, unless a crack deflects the wave.

e. Beam Steering

An array of transducers on a surface is excited at differing times, so that the emitted waves coalesce to form a wave inclined to the surface, as illustrated in Fig. 84. The orientation of the wavefront is controlled (i.e., steered) by varying the excitation times across the array. The technique is usually operated in the pulse-echo mode. It is used extensively in medical imaging.

Alternate methods for steering a beam use an oscillating mount for a transducer, or a mechanical drive to rotate a sequence of transducers mounted on a wheel, as illustrated in Fig. 85. These methods are now out-of-date.

f. Focusing Horn

A horn of metal or plastic can be used in the fashion of a waveguide to focus waves from a large transducer onto a small area, increasing the intensity. The horn is conical with a suitably designed shape, commonly exponential or linear, as illustrated in Fig. 86.

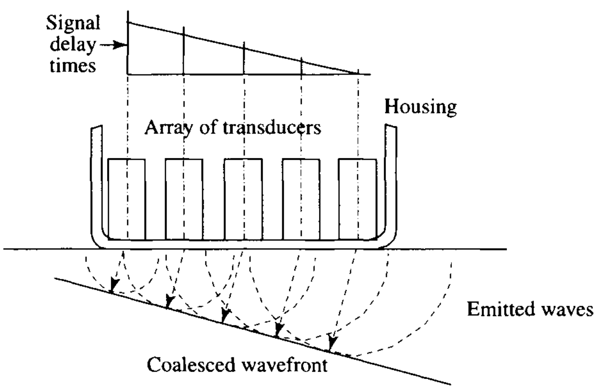


FIGURE 84 Steered wave

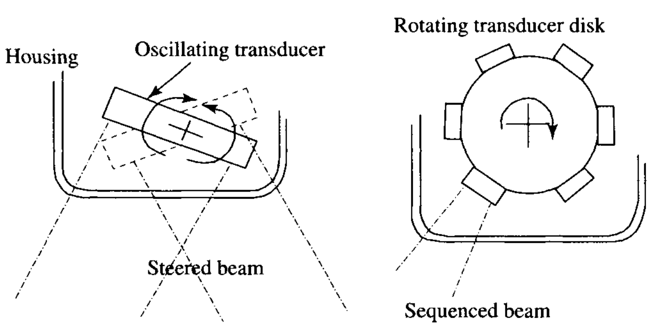


FIGURE 85 Mechanical beam steering

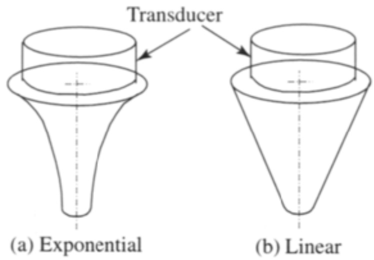


FIGURE 86 Typical focusing horns

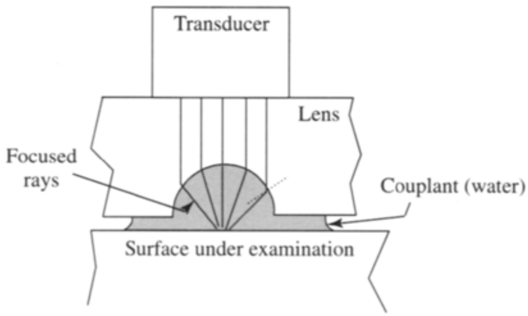


FIGURE 87 Fluid-coupled lens focusing concept

The design of a horn for maximum efficiency can be complex, but a simple machined shape can be effective. A transducer is bonded to the large end, and the smaller end is used to transmit and receive signals.

g. Fluid-Coupled Lens Focusing

The face of a transducer or a face plate is curved to form a lens. A spherical-concave shape is filled with a fluid, commonly water, so that waves emitted from different places on the face are bent (refracted) by differing amounts into the couplant where they meet at a point. The scheme is shown in Fig. 87. The transducer is operated in pulse-echo mode at a very high frequency with a very short wavelength, to view very small regions, and is mounted in a small and precise positioner.

This principle is used in the acoustic microscope for examining computer chips.

h. Bending Waves

Bending waves can be excited and detected by placing a sender and a receiver transducer at some distance apart on a thin plate or bar, as illustrated in Fig. 88.

In some cases the frequencies of interest are quite low so large transducers must be used. Alternatively, an impactor (e.g., a dropping weight) can be used for

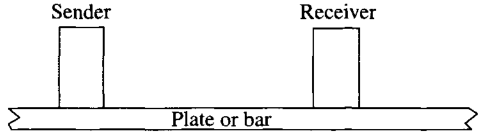


FIGURE 88 Transducer arrangement for bending waves

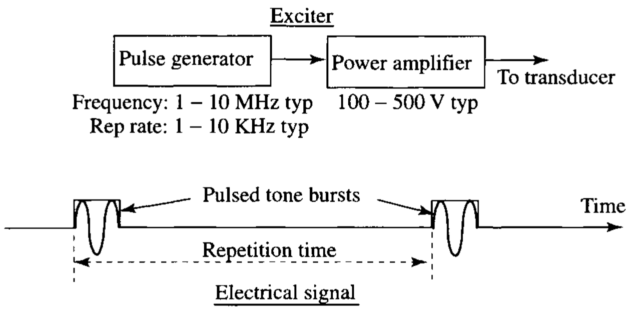


FIGURE 89 Electrical excitation and mechanical response of a transducer crystal

an exciter, and an accelerometer can be used for a detector. The transducer pair can be swept over the plate to map out regions of varying bending properties.

Since bending waves are strongly dispersive (see Section 3 and Appendices 10 and 12), it is desirable to take measurements with several spacings between the transducers to provide data on propagation speeds of various frequency components. If the transducers are of sufficiently broad-band response, it is also desirable that they be used at several frequencies. Alternatively, several alternate transducers could be used. The impact/accelerometer configuration is very broad-banded and excites/receives a wide range of frequencies.

How Is a Transducer Excited ?

A transducer is excited electrically by a high voltage, called the main bang, from a power amplifier driven by a pulse generator. The voltage can be dangerously high. The magnitude of this pulse is adjustable and is sometimes given as a gain in decibels (defined in Section 3). The excitation usually provides a repeated sequence of square pulses as illustrated in Fig. 89.

The mechanical response of the crystal to each pulse is a linear increase of oscillation (resonance ring-up) during the pulse, followed by an exponential decay, whose duration depends on the damping characteristics of the transducer and its circuitry, as illustrated in Fig. 90.

In a pulse-echo system, the excitation pulse (main bang) is recorded together with the echo signals. Generally it is so high that when the gain is set to

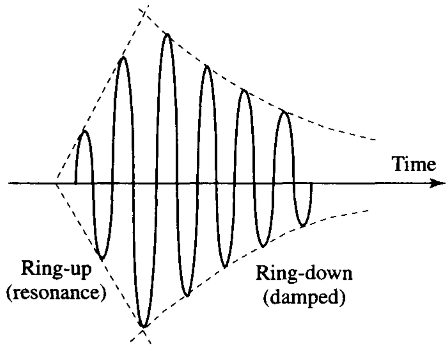


FIGURE 90 Typical transducer response

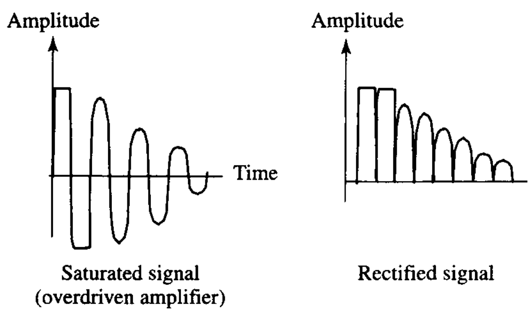


FIGURE 91 Cropping of main bang by overdriven amplifier

adequately display the echo, the main bang overdrives the receiver amplifier. The main bang is then cropped and possibly delayed, as illustrated in Fig. 91.

What Is the Repetition Rate and Averaging ?

The excitation of the sender is repeated many times to allow many wave transits to be averaged. This minimizes the effects of minor spurious fluctuations, such as electrical noise. The rate of repetition (rep rate) is selected to ensure that all relevant response in one echo has died down before the next arrives.

What Is a Trigger ?

To ensure that the receiver system (recording device) is switched on when a signal arrives, and not before or after, a short duration of the signal is continuously stored in a buffer. When a significant rise in signal is detected, a trigger is set off to start recording, including the current buffered segment. The trigger provides a precise timing of the recording, so that multiple pulses can be overlaid.

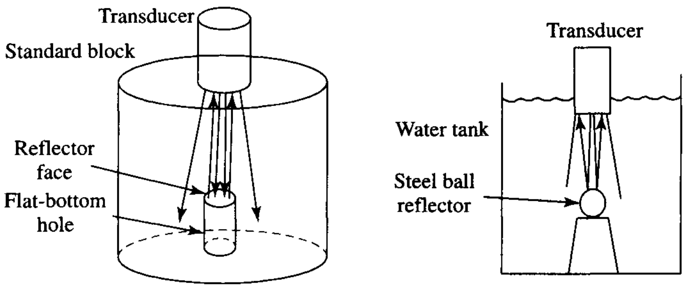


FIGURE 92 Typical UT standards

What Is a Digital or an Analog Receiver System ?

The signal generated by the receiver transducer from incident wave motions is a continuous (or analog) current. In an analog system the signal is displayed on some device such as an oscilloscope which shows a continuous trace. Processing of the signal is limited, so analog systems are limited to visual data interpretation.

The signal is usually converted into a discrete series of digital values by sampling at a high rate, as discussed in Section 6, to allow detailed analysis by computer.

What Is Standardization?

A UT system can be standardized to ensure repeatability in measuring amplitude or to validate images of a specific type. This operation is done with Standard or Reference blocks, designed for a specific task. They essentially function as comparators: results obtained with different systems can be compared when they are adjusted to give the same results on the standards.

This comparator concept is not totally accurate, because factors such as anisotropy modify the characteristics of a UT test in ways which may not be accommodated in the comparison. UT Standards are defined by the American Society for Testing Materials (ASTM).

To standardize a UT beam, a target, which essentially probes the beam by creating a selective backscatter, is placed in a propagating medium at some distance from the surface where a transducer is applied, as illustrated in Fig. 92. The medium can be a solid, typically Plexiglas or aluminum, or a liquid, such as water. In the solid, the target is a flat-bottomed hole drilled from the back face to serve as a small reflector. In the liquid, a steel ball is mounted on a pedestal.

To validate scans of a test object, another object, sometimes called a phantom, is constructed to simulate the object being tested, and specifically incorporating an out-of-specification configuration such as a shape or material insert or a defect. Scans of the phantom allow settings such as gate times and threshold levels (see Section 6) to be set.

What Is Calibration?

Calibration is a procedure used to establish quantitative measures of the response characteristics of the equipment. A UT system can produce quantitative measurements of time and amplitude in a waveform, but in most applications only changes in such quantities are measured. These are used to produce visual images, which are not precise measurements, and the response of the equipment is of no concern. Accuracy is required only when the results of measurements made with different systems are to be compared. The system must then be calibrated, as discussed below.

Time Measurements

The measurement of a time event depends on the accuracy of the oscillator used to determine the digitization sampling rate. This is usually controlled extremely well, so that time measurements are, of themselves, very precise. Digitization poses a discrete time interval for each measurement so that the resolution of a time measurement is one half of the interval between samples.

The issue of time accuracy is rather one of identifying the event, e.g., the onset or the peak, in a wave to be measured, particularly when the waveform changes during propagation as in a dispersive system (see Appendix 7). Methods for identifying an event are discussed under Software in Section 6. Factors of the test system, such as path length, identification of specific reflections, and material uniformity, must be considered.

Amplitude Measurements

The measurement of wave amplitude is usually made to evaluate attenuation, i.e., a change in amplitude during propagation, and not absolute amplitude. This is taken in terms of the electrical signal, e.g., volts, and not physical quantities such as stress or particle velocity and impedance.

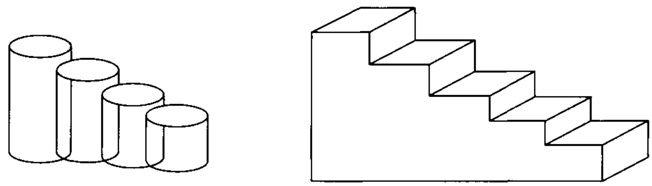
If attenuation as measured in two different systems is to be compared, then factors of the equipment which affect the UT waves must first be determined. These include beam formation effects, such as radial spread and axial decrease, and interface effects at the transducer, such as transducer properties and coupling. Measurements can be made on known objects to define these characteristics of the test system.

What Test Equipment Factors Must Be Measured?

The are two test-unique parameters which, once measured can be used to make corrections:

- The effective size of the transducer (significant mainly for squirters)
- The coupling impedance between the transducer and the test object

The diameter of a contact transducer is essentially that of the wear face, though usually slightly smaller because of nonplanarity, roughness, etc. In a water



(a) Set of calibration cylinders (b) Step wedge calibration block

FIGURE 93 Illustration of a cylinder set and step wedge for calibration

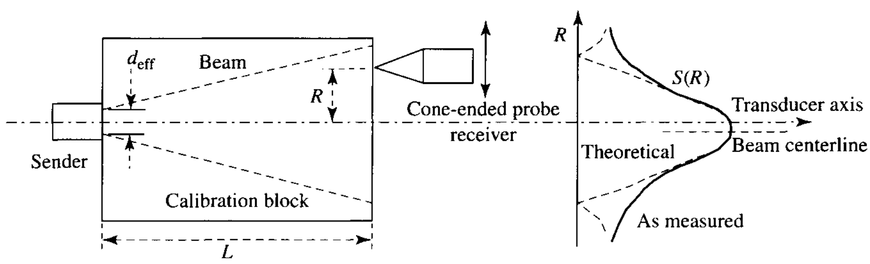


FIGURE 94 Illustration of method for calibrating effective transducer diameter

squirter or bubbler, the diameter is only roughly known because the water jet spreads and the UT waves become dispersed. In a water tank, the effective diameter of a transducer is known through the beam angle for water, not the test object.

How Is a UT System Calibrated?

To effect a calibration of these parameters, tests are made on calibration blocks, which are usually cylinders of a well-known material (e.g., Plexiglas) in several lengths, or a step wedge, which is a block with steps, as illustrated in Fig. 93. The use of the blocks and their design are discussed later.

1. Transducer Diameter

The size of a transducer affects mainly the beam-forming process (see Section 3 and Appendix 8) in which the beam spreads radially and diminishes in propagation axially. The radial beam spread can be measured to determine the effective transducer diameter, which can then be applied to analysis of the axial variation.

From Appendix 8, the radial distribution of a UT beam depends on the two ratios R/L, which is radial off-axis distance normalized to axial distance from the transducer, and d_{eff}/λ, which is the effective transducer diameter normalized to wavelength. These quantities are illustrated in Fig. 94.

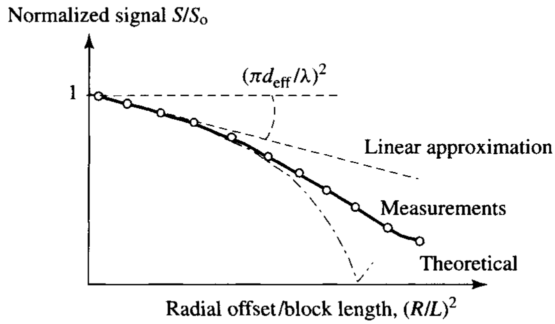


FIGURE 95 Linear fit to radial distribution for evaluating effective transducer diameter

The theoretical beam distribution is plotted as a function of R/L for several values of d_{eff}/λ in Fig. 40. An approximate formula from Appendix 8 for the signal strength at small radial offset is

$$S(R)/S_o = 1 - (1/8)(\pi d_{eff}/\lambda)^2(R/L)^2 + \ldots.$$

A conical end (typically Plexiglas machined to a small rounded point) is glued to a transducer to create a receiver probe. A calibration block is then excited at the center of the face at one end by another transducer, and the face at the opposite end is scanned with the probe, as illustrated in Fig. 94.

The center of the beam is identified carefully by locating its maximum, and the distribution about the center line is measured for several cylinders of differing lengths. The results are plotted as pairs of values of S/S_{max} vs $(R/L)^2$, leading to a straight line with slope $(1/8)(\pi d_{eff}/\lambda)^2$, as illustrated in Fig. 95. The wavelength must be known from the frequency and the wavespeed in the cylinder.

2. Transducer Impedance and Coupling

The interaction between a transducer and an object of isotropic material produces waves of normal transmission and reflection (except for anisotropic materials or the use of an angle shoe). The strengths of these waves are determined by the coefficients of reflection and transmission across the interface (see Section 4), which depend on the ratio of the impedances of the object and of the transducer. The impedance of the transducer is determined by the properties of the exciter crystal and of any intermediary material, such as a wear face or a water beam, which is generally not known.

The effective impedance can be determined by measuring amplitudes of multiple back-face echoes in blocks of several lengths.

The signals received by the transducer as generated by multiple reflections within the block are illustrated in Fig. 96.

The first signal is the front-face echo, S_{FF}, and succeeding signals, S_{Bn}, are back-face echoes after multiple passes through the block. The reflection and

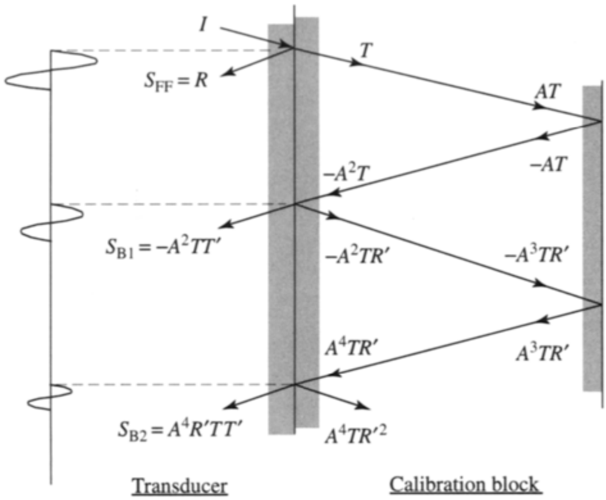


FIGURE 96 Transducer signals generated by multiple reflections

transmission coefficients for waves passing from the transducer to the block and vice versa, used in the figure, are (from Section 4)

$$R = (r - 1)/(r + 1),\ R' = -(r - 1)/(r + 1),$$
$$T = 2r/(r + 1),\ T' = 2/(r + 1)$$

where $r = z_{\text{block}}/z_{\text{xdcr}}$, z_{block} is the known impedance of the block, and z_{xdcr} is the unknown impedance of the transducer. The amplitude ratio from one end of a block to the other due to attenuation is

$$A_\alpha = e^{-\alpha L},$$

where α is the attenuation coefficient for the material of the block, and L is the length of the block. Through beam divergence, the signal strength at one end of the block after n passes is inversely proportional to the distance $nL : S_n \sim 1/nL$, and after one more pass it is $S_{n+1} \sim 1/(n + 1)L$. The amplitude ratio between these two passes is

$$A_L = n/(n + 1) = 1/[1 + (1/n)].$$

This ratio is approximately unity except for the first few passes. The total amplitude ratio is then $A = A_\alpha A_L \sim e^{-\alpha L}$.

Evidently, the nth echo can be written as

$$S_{Bn} = (-1)^n (TT'/R')(A^2R')^n$$

so that the attenuation in the nth signal is

$$\alpha_n = \log|S_{Bn}| = \log(TT'/R') + n\log(A^2R').$$

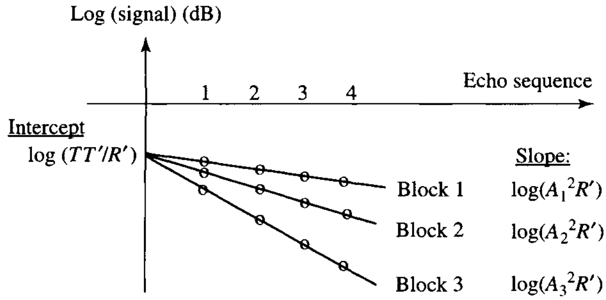


FIGURE 97 Measured back-face echoes and their interpretation

The attenuation is plotted vs the echo sequence number, as illustrated in Fig. 97. A correction can be made to the data points for the first few echoes to account for the factor $n/(n+1)$.

The slopes $d\alpha_n/dn = \log(A^2R')$ of the line for each block are different because the attenuation factor A depends on the length of each block. Now, the common intercept is

$$\alpha_o = \log(TT'/R')$$

so that

$$10^{\alpha_o} = TT'/R' = 4r^2/(1-r^2) \text{ and } r = [10^{\alpha_o}/(4+10^{\alpha_o})]^{1/2}.$$

Then

$$z_{\text{xdcr}} = z_{\text{block}}/r.$$

A check on the result can be made by calculating TT'/R', and then $A = e^{(d\alpha_n/dn)}/R'$, which should be essentially the same as the known attenuation of the cylinder given above, i.e., $A = e^{-\alpha L}$.

How Can Attenuation Measurements Be Corrected?

The amplitude of a signal can be corrected to account for the effective transducer diameter and impedance in three ways:

- Radial spread
- Axial reduction
- Interface coupling

1. Radial Spread

A transducer which receives a signal from a propagated beam will average the signal over its face, so that the radial distribution of the beam must be considered. If the beam is broad, then variation across the transducer will be small, but a large, low-frequency transducer will have appreciable variation. The effective diameter of the transducer depends on its operation as well as on its physical

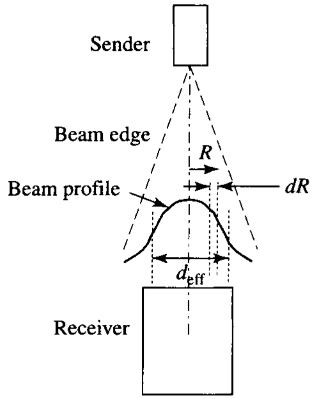


FIGURE 98 Signal averaging over transducer face

size. The average signal over a transducer face, as illustrated in Fig. 98, is

$$S_{av}/S_o \simeq 1 - \pi(\lambda/2L)^2(d_{eff}/2\lambda)^4.$$

The measured signal then departs from nominal by the increment $\pi(\lambda/2L)^2 (d_{eff}/2\lambda)^4$.

2. Axial Decrease

The signal decreases inversely with distance,

$$S \sim S_o 4N/x,$$

where $N = (d_{eff})^2/4\lambda$ is the near-field distance (see Appendix 8).

A change of effective transducer diameter introduces a change to the near-field distance and hence the signal strength at a given distance. Thus a correction can be made to relate the signal from one diameter to another according to the formula

$$S_{deff \cdot 1} = S_{deff \cdot 2}(d_{eff \cdot 1}/d_{eff \cdot 2})^2$$

3. Interface Coupling

The amplitude of an nth echo is dependent on the effective impedance z_{eff}, using the formulas quoted earlier:

$$S_{Bn}/I = (A^2 R'_{eff})^n (A^2 T_{eff} T'_{eff})$$

The front-face echo is simply

$$S_{FF}/I = R_{eff}.$$

Design of Calibration Blocks

The design of the blocks or wedges must account for reflections from the sides of the piece. They must be long enough to allow near-field effects to die down and

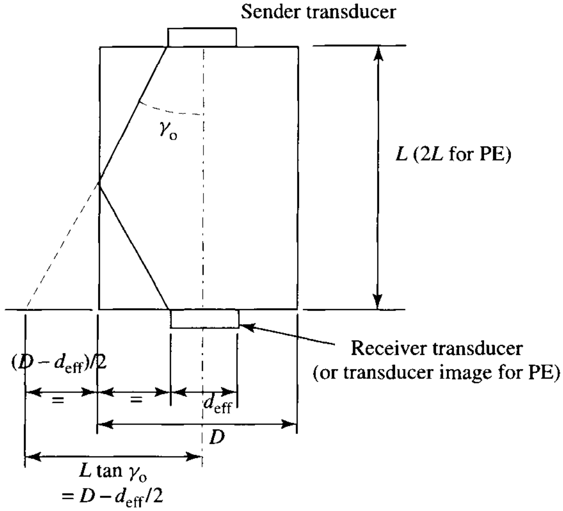


FIGURE 99 Design of calibration block to eliminate lateral reflection

for significant attenuation to develop, yet not so long that reflections from the sides interfere. The near-field effects die down for $L > 4N$, where $N = d_{eff}^2/4\lambda$ is the near-field distance, so that $L > d_{eff}^2/\lambda$.

Outside a direct propagation path to the receiver the beam reflects toward the receiver off the sides. For ray angles greater than atan$[(D - d_{eff}/2)/L]$, however, the side reflection does not reach the receiver, as illustrated in Fig. 99. Rays outside the beam, i.e., with an angle $\gamma_o >$ asin$[1.2\lambda/d_{eff}]$, are of no consequence, so that $[(D - d_{eff}/2)/L] > \tan \gamma_o$.

The diameter/length ratio for the block must be therefore be bounded by the following:

$$(D/d_{eff})/(d_{eff}/\lambda) > D/L > d_{eff}/2L + \tan \gamma_o.$$

6

SOFTWARE: DATA PROCESSING

WHAT DOES SOFTWARE DO?

Software manipulates waveform data to produce a desired display, or to provide a measurement of some aspect of the waveform. The manipulation is based on mathematical procedures (algorithms) and may be performed in the natural time domain or in the spectral frequency domain, as described below.

This section is subdivided into four parts:

A. Time Domain Analysis
B. Spectral Domain Analysis
C. Statistics
D. Imaging

What Is the Time Domain?

A waveform is a sequence of wave magnitudes in a time progression, defined in Section 2. The representation of such a set of measurements is referred to as the time domain. The wave motions can depend on time in any way.

What Is the Frequency Domain and the Spectrum?

The amplitudes and phases of a sequence of sine waves of various frequencies which represent a waveform (as introduced in Section 2) comprise the spectra of the waveform. The frequency space in which they are defined is the frequency domain.

There are two common forms of spectra: the discrete Fourier or harmonic series, and the continuous Fourier spectrum. These are discussed later.

What Are the Software Types?

The main types of software are constructed to produce the following results, and operate in the time or the frequency domain:

- Display of waveforms
- Measurement of a single parameter (e.g., time of flight, wavespeed, attenuation)
- Statistical representation of defects or properties
- Waveform conditioning and spectral analysis
- Imaging (mapping) of the distribution of defects or of variations in properties (cracks, porosity, wavespeed, thickness, attenuation, etc.)

A. WHAT ARE THE TYPES OF TIME-DOMAIN ANALYSES?

1. *Display.* Several ways of displaying the basic waveforms are the following:

    - A- and B-scans
    - RF and video waveforms
    - The analytic envelope

2. *Time-Domain Measurements.* Analyses in the time domain which provide an end result of a measurement are:

    - Overamplification
    - Gating and thresholding
    - Correlation and convolution
    - Dispersion analysis

3. *Waveform (Signal) Conditioning, Preparation for Subsequent Analysis.* These processes prepare waveform data for further analysis:

    - Waveform averaging
    - Detrending

- Smoothing
- Zero padding
- Truncation and windowing

Smoothing can also be accomplished in the frequency domain (as described later).

What Are the Time-Domain A-Scan and B-Scan?

Traditionally, there have been two types of waveform data display, the A- and B-scans. Many variations on these have been developed.

The traditional A-Scan is a waveform, or the entire history of the main bang and one or more echoes, useful for indicating what is happening in terms of noise, echoes, reflections, etc., as illustrated in Fig. 9.

The traditional B-scan (or line scan) is a sequence of A-scans taken at points along a line on the surface of an object, and displayed side-by-side to represent a cross-section of the boundaries of the object, as illustrated in Fig. 100. Because the reflected pulse is attenuated, the gain is sometimes increased by a preset fixed or time-varying amount for the echo.

A presentation similar to the B-scan is used in seismology, referred to as a seismic cross-section plot, to illuminate the arrival of echoes from various layers in the earth's crust. The positive half of each pulse is darkened as illustrated in Fig. 101, by drawing lines between the baseline and the positive value at each time. It is then visually easy to identify the locus of an interface, as illustrated in the figure.

How Are Waveforms Displayed?

Waveforms can be displayed in several ways, illustrated in Fig. 10 in Section 2, i.e., the RF signal, the video waveform, and the analytic envelope.

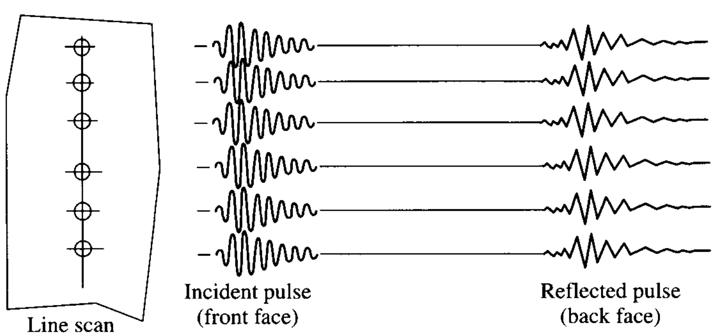


FIGURE 100 Illustration of a B-scan

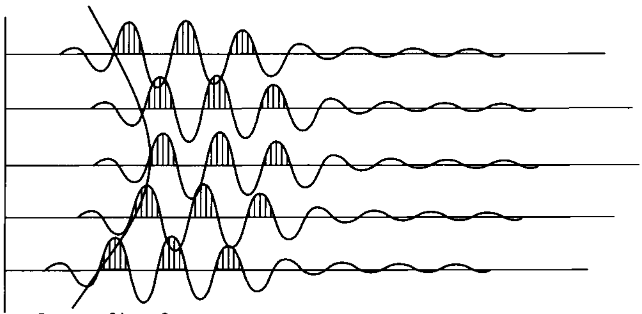


FIGURE 101 Illustration of seismology display

What Is the RF Waveform?

The signal received by the transducer is displayed directly. Since the UT frequency is typically in the kilohertz or megahertz range at which early-day radio waves propagated, the signal is called the radio-frequency (RF) waveform. Such signals must be handled with appropriate technology, such as the use of coaxial cabling, and high-frequency amplifiers.

What Is the Video Waveform?

The signal is rectified (negative values are changed to positive) and displayed possibly with smoothing or filtering to remove the high-frequency oscillations (discussed later). This waveform is referred to as video because, historically, it was displayed on an oscilloscope with low frequency response to produce a visual waveform.

What Is the Analytic Envelope?

A numerical procedure based on the Hilbert Transform (described next and in Appendix 13) is applied to the waveform in the time domain to create an envelope of the waveform. This is a representation of the amplitude modulation on the carrier wave frequency. The procedure also creates a time-domain phase function which has application to interpreting dispersion.

The Hilbert transform $H(t)$ of a function $f(t)$ can be thought of as an integration operator which results in a differentiation. Thus the transform of $\sin(t)$ is $\cos(t)$, and conversely, that of $\cos(t)$ is $-\sin(t)$. The points of a waveform are multiplied by a function (equivalent to $1/t$) which is antisymmetric about $t = 0$, and the result is integrated over a wide range. The envelope is then given by the square root of the sum of the squares of the waveform and of the transform, $E(t) = [f(t)^2 + H(t)^2]^{1/2}$, analogous to the relationship $(A^2 \sin^2 \theta + A^2 \cos^2 \theta)^{1/2} = A$.

What Measurements Are Taken from a Time-Domain Waveform?

Three types of numbers can be extracted from the time domain waveform:

- Event time—typically the signal arrival or onset
- Event amplitude—typically peak amplitude or a sequence of amplitudes
- Statistical data

Before any measurement can be made, an event must be identified: some aspect of the waveform must be selected and defined in relation to the waveform. A common example is the arrival or onset of a wave, such as the first motion in the signal. This can be confused with noise, particularly after a significant distance of travel.

Since a waveform can change by dispersion and attenuation, identifying the event often requires interpretation through an understanding of the intended measurement and of any ways in which it may change during propagation.

How Is a Time Event Measured?

The time of a selected event is determined by detecting the first time that the signal exceeds a threshold set to characterize the event. The onset of a pulse is detected with a threshold set to a value just above the noise level, which must be estimated from data taken with no signal present. For other events, the nature of the waveform and the intended feature must be considered to allow a suitable threshold to be set.

What Is Overamplification?

Overamplification provides contrast between the noise and the signal, to help identify the onset of a signal. The gain of the signal receiver amplifier is set to a large value, so that the signal is either small or saturated, as illustrated in Fig. 102. This is equivalent to enhancing the contrast in a picture so that the image is black or white.

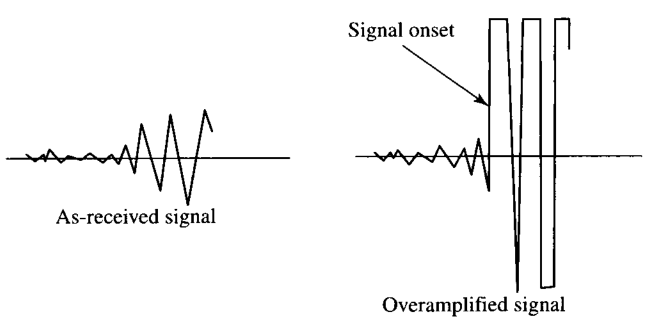


FIGURE 102 Illustration of overamplification to determine signal onset

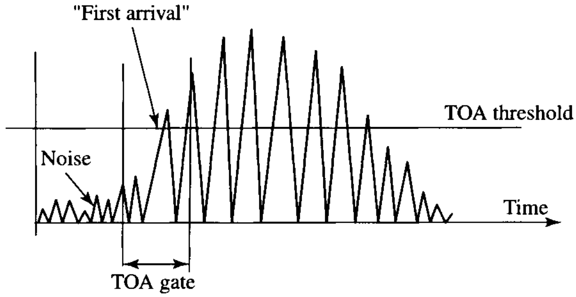


FIGURE 103 Gate-threshold technique for detecting wave arrival

What Is Gating and Thresholding?

A time interval, called a gate, or a window, is selected to ensure that only the event of interest in a pulse is examined, as illustrated in Fig. 103 for the arrival time. The signal is not examined outside the gate. Setting the gate requires a rough knowledge of the event time, for example, the propagation speed and distance, or the gate can be set by trial and error. Evidently a miscalculation here can lead to the wrong portion of the signal being selected.

A threshold is a signal level which is expected to be exceeded by the amplitude of the portion of signal of interest, for example a multiple of the noise level for the arrival time. Again, if the threshold is incorrectly set, the wrong event may be selected. Several thresholds can be used in one gate to narrow the selection of the event.

What Is the Time of Flight?

The time taken for a wave to travel a designated path is the time of flight or travel time. This is the difference between the onset time of the inserted wave and that of the received wave. In a TTU system, the path length is the dimension of an object, whereas in a PE system, the path is twice the distance to the echoing surface, of course. This may be an internal interface or crack, or the back face. The measurement of time of flight is illustrated in Fig. 104.

The figure illustrates the problem of determining the onset of the received signal: the echo is preceded by some noise, and the signal rise is distorted from that of the input signal. Alternatively the time difference between the peaks of the two wave forms can be used, but propagation affects these also.

How Is Time of Flight Used for Wavespeed or Path Length?

Wavespeed can be determined when the distance propagated (path length) is known, or path length can be found when the wavespeed is known. In a PE system, the path length for a back-face echo is twice the thickness; in a TTU

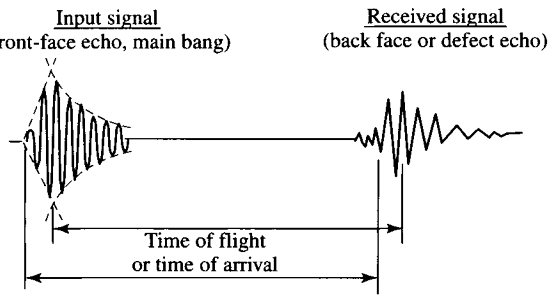


FIGURE 104 Illustration of time of flight, transit time

system, however, the path for the transmitted wave is just the thickness, but for subsequent reverberations it is twice that. In complex configurations where there may be sequences of reflections from multiple interfaces, or reflections may develop off side faces, the path length must be examined carefully.

If the wave path includes more than one material, or segments with mode conversion resulting in a different wavespeed, then the measured time includes the propagation time for each path segment:

$$t_{\text{meas}} = \sum x_i/c_i$$

where i is an index which identifies each segment, x_i is the length of each segment (thickness for a single pass, or twice for an echo), and c_i is the wavespeed. To determine wavespeed or thickness (path length) for any one segment, these must be known for all others.

The propagation time for a selected segment of the path is then

$$T_{\text{sel}} = t_{\text{meas}} - \sum x_i/c_i$$

where i now excludes the selected segment. The wavespeed in this segment is

$$c_{\text{sel}} = x_{\text{sel}}/t_{\text{sel}}.$$

This wavespeed must be considered with care to identify the type of wave. If the material is dispersive (e.g., particulate or directional, or there are surfaces along the path), then the wave is likely to give a group wavespeed measurement. The measurement should then be repeated at several different frequencies and in several directions to resolve the type. Of course, the appropriate speed must be used for the known segments.

How Is Amplitude Measured?

Commonly, the waveform event used for amplitude is the highest signal, or peak, lying within a preset gate, or time interval, as illustrated in Fig. 105. The discussion given earlier on the gate for time-of-arrival measurement is pertinent here.

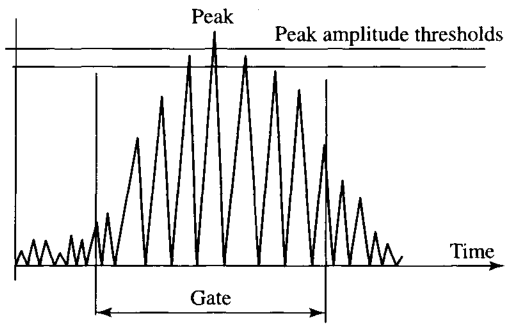


FIGURE 105 Gate threshold for peak amplitude

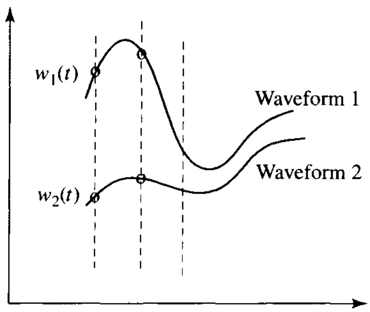


FIGURE 106 Illustration of correlation procedure

What Is Waveform Correlation and Convolution?

Any similarity between two waveforms is examined by determining the correlation between them (discussed later under Statistics).

In correlation, amplitudes at the points of one waveform $w_1(t)$ are multiplied by those of the second waveform $w_2(t)$ at each time point, as illustrated in Fig. 106, and the products are summed. The sum, $S = \sum w_1(t)w_2(t)$ measures the similarity or correlation between the two waveforms.

A series of correlations between the waves after shifting one in time relative to the other by increasing amounts is a convolution and is a function of the time shift:

$$C_{12}(t_s) = \sum w_1(t)w_2(t - t_s)$$

where t_s is the shift time and C_{12} is the convolution between waveforms 1 and 2. It is mathematically equivalent to an integration,

$$C_{12}(t) = \int w_1(\tau)w_2(t - \tau)d\tau,$$

which is usually evaluated by a spectral technique given later. Convolution between two waveforms allows the most accurate means of measuring the time difference between them, such as the time of flight, and also provides measures of periodicity by examining the sequence of peaks which result. This is useful in analysis of dispersed waves as discussed later. (Other applications of convolution are described later.)

How Is Dispersion Analyzed?

Dispersion, the frequency dependence of wavespeed, manifests itself in a waveform which stretches or shrinks in time, as discussed in Section 3 (see Fig. 46) and Appendix 7. It can be analyzed in the time domain through zero-crossing or convolution, or in the frequency domain through the phase spectrum described later.

What Is Zero-Crossing Analysis?

A dispersed time-domain waveform consists of oscillations whose period varies with time, as illustrated in Fig. 107.

A typical oscillation crosses the axis at times t_n and t_{n+1}, so that half the period of the nth cycle is $\Delta t = (t_{n+1} - t_n)$, and it is centered at the time $|t_n| = (t_{n+1} + t_n)/2$. The frequency is then

$$f_n = 1/2(t_{n+1} - t_n)$$

and the phase, which is 0 at the zero-crossing t_n, is

$$\phi = \kappa(x - ct_n) = 0$$

where k is the wave number, x is the distance propagated, and $c = c(f)$ is the wavespeed at frequency f, so that

$$c(f_n) = x/t_n.$$

A plot of $c(f_n)$ vs f_n is an approximation to the dispersion curve, as sketched in Fig. 108. This is only approximate because it assumes the wave shape to be constant over a half period.

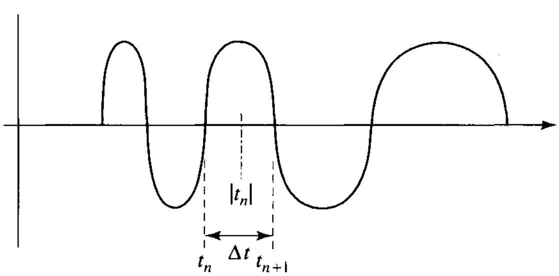


FIGURE 107 Illustration of zero-crossing analysis

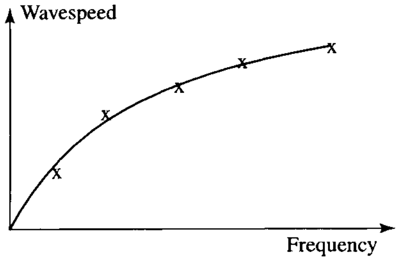


FIGURE 108 Dispersion curve as measured by zero-crossing analysis

What Is Waveform Conditioning?

Errors which are introduced by data acquisition systems can be mitigated by preconditioning the waveforms in the time domain prior to spectral analysis. The factors which arise and the methods for conditioning are described next.

What Factors Affect the Time-Domain Signal?

The time-domain signal is affected by the following factors:

- Noise
- Sampling
- Aliasing
- Drift
- Leakage
- Truncation

What Is Noise and the Signal-to-Noise Ratio, S/N?

There is always some small signal, even without any wave activity: random noise, caused by spurious electrical or mechanical disturbances. Noise is usually of high frequency in relation to the signal of interest. Wave motions generated by the excitation, but which are not part of the wave of interest, such as echoes, resonances, and mode conversions, may also considered to be noise as illustrated in Fig. 109, particularly if these motions overlap and interfere with the signal.

The ratio of the amplitude of the signal of interest to that of the noise is called the signal-to-noise (S/N) ratio. Although a signal is usually greater than the noise, i.e., $S/N > 1$, specialized procedures are available to extract a nonrandom signal from random noise through analysis of coherence, but this is not discussed here.

Noise is measured as a statistical variable so it is quantified by its root mean square (RMS) value, or better, through spectral intensities (discussed later). The noise spectrum is determined from data taken prior to the onset of the wave.

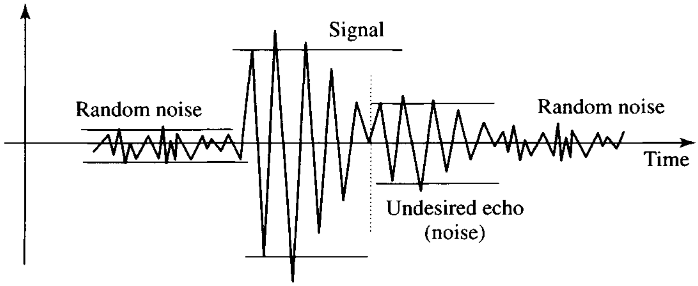


FIGURE 109 Illustration of signal and noise

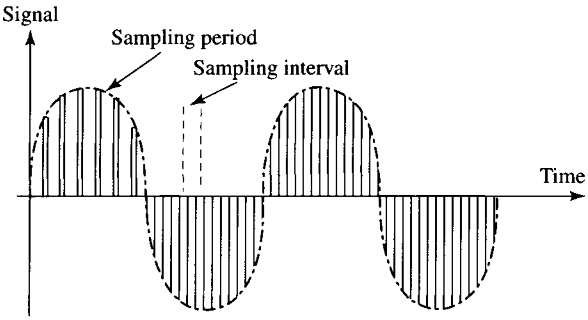


FIGURE 110 Digital signal sampling

The noise can be removed from the signal by subtracting its spectrum from that of the signal, or by a smoothing procedure (discussed later).

What Is Sampling and Digitizing?

In a digital system, the signal is sampled for very brief periods, repeated at a high frequency, as illustrated in Fig. 110. The amplitude of the signal in each sampling period is determined and is expressed as a digital number on a binary or equivalent scale. The waveform is then represented by a collection of numbers equally spaced in time, which can be manipulated by numerical processes in a computer.

The sampling (or digitization) rate or frequency, which is determined by the sampling interval, must be sufficiently high that the signal is adequately resolved. As an example, a minimum of 10 samples is required to resolve a half wave, so that for a 1 MHz signal, the digitization rate must be higher than 20 MHz. with a sampling interval of 0.05 μsec. The duration of the sample must be long enough for the circuitry to work, yet short in comparison to the sampling interval.

The selection of sampling rate is influenced by the factors described later, which include the total duration of the signal and the frequencies of interest in the signal.

What Is Aliasing?

When a pulse is sampled at time intervals Δt, the highest useful frequency, f_{max}, which can be attained in a spectral analysis is the Nyquist frequency:

$$f_{max} = f_N = 1/2\Delta t.$$

All frequencies higher than the Nyquist frequency are partially sampled at lower frequencies and the data are included in the transform for the lower frequencies. This effect is called aliasing and is illustrated in Fig. 111. Aliasing leads to errors in the intensity estimate for low frequencies. This can be avoided by making the Nyquist frequency high, i.e., by raising the sampling rate through small Δt.

The minimum frequency, f_{min}, which is also the frequency resolution increment obtainable from a spectral analysis (described in Appendix 14), since the frequencies obtained are usually multiples of f_{min}, is

$$f_{min} = 1/T_{tot} = 1/N\Delta t = (2/N)f_N$$

where T_{tot} is the total duration of a waveform, and N is the total number of digitized points. Evidently the frequency resolution and the Nyquist frequency are coupled by the total number of points.

What Are Offset and Drift?

Data offset occurs when the baseline, i.e., zero line, of the waveform is incorrectly set or has shifted through accident or varying conditions prior to data

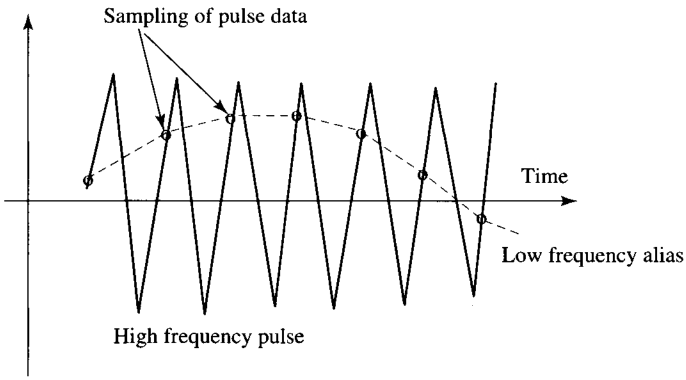


FIGURE 111 Alias sampling

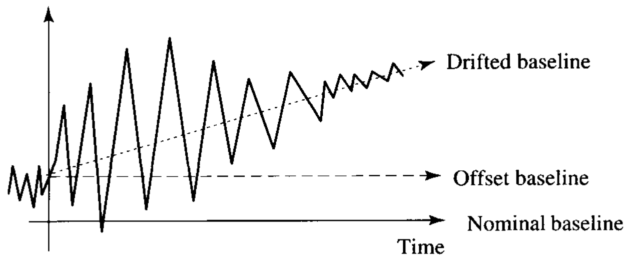


FIGURE 112 Baseline offset and drift

collection. Drift occurs when the baseline changes gradually in time from its initial line during data collection. These effects, illustrated in Fig. 112, produce unduly large low-frequency components in a spectral analysis and should be removed. The offset and drift can be subtracted from the waveform before spectral analysis is made (called preconditioning, described later).

Note that the frequency of most electric power sources (50 or 60 Hz), a common source of noise, is very low in relation to the frequencies in a UT signal so that it appears as a baseline shift which changes from test to test.

What Is Frequency Leakage?

The use of discrete frequencies in sampling causes the spectral content between the discrete time points to be spread onto those points. This is called leakage, and occurs for high frequency content, particularly if there are spurious echoes in the signal. Leakage is minimized by data padding, described later.

What Is Truncation and Sampling Bandwidth?

Another source of leakage is from truncating the data before the waveform is ended. This produces an artificial sharp step in the waveform, as illustrated in Fig. 113. The step introduces spurious high-frequency content into the spectrum through Gibb's Phenomenon, discussed later. This can be minimized by using a window which modifies the waveform values at the end of the sampling period. Windowing is discussed later.

Another effect of truncation arises because a signal is sampled over a finite time, t_s. The spectrum of a sine wave, which is theoretically a sharp line at the frequency of the sine, becomes broadened over a frequency band π / t_s with side lobes, as illustrated in Fig. 114. A short sampling time leads to a wide band, and therefore reduces the ability to isolate a frequency. This is the bandwidth associated with the signal. A narrow bandwidth is desirable in a spectral analysis, but in communications a broad bandwidth allows a wide range of amplitude modulation (superposition of low-frequency waves onto a high-frequency wave).

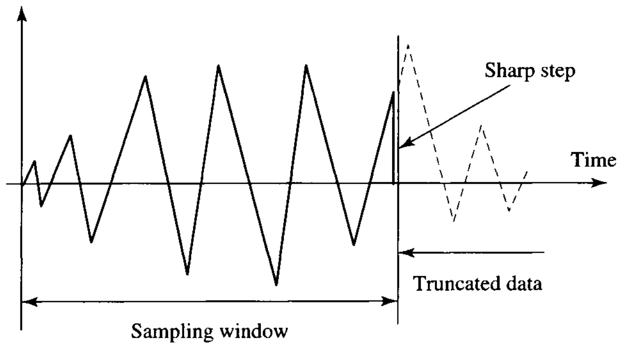


FIGURE 113 Data truncation

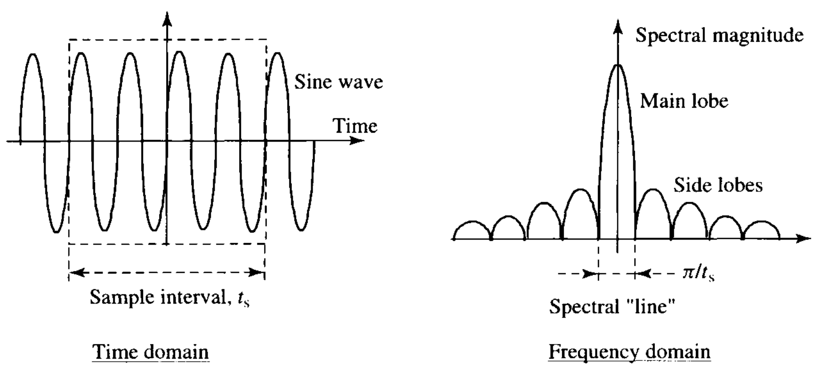


FIGURE 114 Bandwidth effect of finite sampling

How Are Time Domain Waveforms Conditioned?

Various procedures can be performed on a waveform to reduce the erroneous effects described above, as follows:

- Averaging
- Preconditioning (offset and drift correction)
- Smoothing (time domain)
- Windowing
- Zero padding

What Is Waveform Averaging?

Spurious random effects in a waveform, such as electrical or environmental variations, can be minimized by repeated acquisition of the signal from the test configuration. The assumption is made that these variations differ from waveform to waveform, whereas the signal does not. It is not desirable or necessary to store all the waveforms.

The excitation pulse is repeated by the exciter circuit at frequencies between 100 and 2000 Hz (the cause of any sound heard in a UT test), in essence producing multiple tests of the same configuration. The readings from each repetition at each time increment are summed continuously for a preset number of repeats, typically 100 to 1000. The sum is divided by this number to provide an average over the repetitions. In practice, an accumulating average procedure is used in which only the latest average is stored, as described in Appendix 13.

What Is Preconditioning?

To reduce the low-frequency effects introduced by offset and drift of the signal, the entire waveform is subject to a linear regression (described later under Statistics and in Appendix 13), i.e., fitting a straight line (refer to Fig. 112, baseline offset and drift) by the least squares method:

$$f_{\mathrm{o,d}}(t) = A + Bt$$

where $f_{\mathrm{o,d}}(t)$ is a time function which represents the offset and drift. Values calculated from this function are subtracted from the data points.

What Is Smoothing by Moving Averages?

A moving-point average is used to eliminate noise from the signal, assuming it to be of high frequency in relation to the signal itself. This process takes data from a time interval covering some number of data points and averages them to produce a new point assigned to the center of the interval, as illustrated in Fig. 115. The time interval is moved from one end of the signal to the other, in steps of one or more points to create a new waveform.

This procedure not only removes high-frequency oscillations superimposed on the waveform, but also reduces the peaks and troughs and flattens the slopes

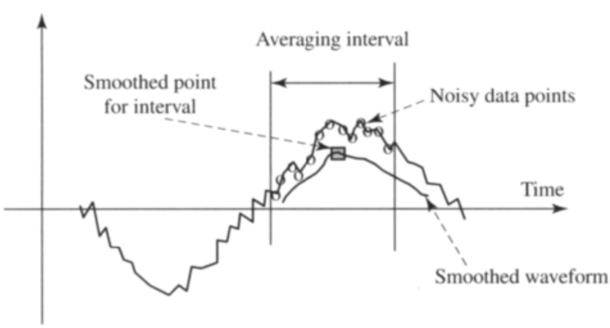


FIGURE 115 Moving average smoothing

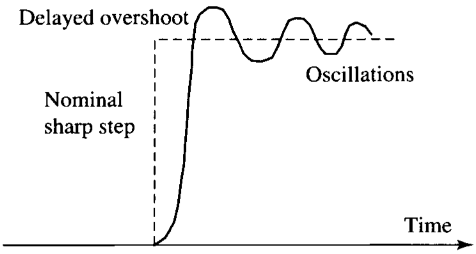


FIGURE 116 Illustration of Gibb's Phenomenon

of the waveform. Thus, while it can enhance the visual appearance, it can distort the waveform and hence its spectrum. As a check, the smoothed signal can be subtracted point-by-point from the original signal to display the difference, which should be unwanted noise showing no relationship to the smoothed signal.

These effects are serious if the averaging interval is so long that it spans the waveform oscillations of interest. To evaluate this, the duration of the moving interval should be changed and the process repeated to determine the best size for adequate smoothing without excessive distortion.

What Is Windowing?

Selecting a portion of a signal for analysis is called windowing. This process can be represented by a sharp window which truncates the signal before and after preset times, as described earlier and illustrated earlier in Fig. 113. Such a sharp window is called a square or box-car window. A sharp onset, i.e., a square cutoff, can only be achieved in a system with infinite frequency response. Any real system produces an effect called Gibbs' Phenomenon, illustrated in Fig. 116, in which a step is replaced by an overshoot followed by oscillations. These oscillations modify the frequency characteristics of the signal.

This effect is minimized by modifying the window to smooth out the transitions. Such windows are represented by time functions which can be applied mathematically to the signal.

Common windows which replace the square or box-car are illustrated in Fig. 117. These are the tapered step, and the Hamming and Hanning windows. They use various trigonometric or power-law functions to modify the jump.

What Is Data Padding?

For efficient spectral analysis and to improve the frequency resolution (discussed later), the number of data points should be a large power of 2 (e.g.,

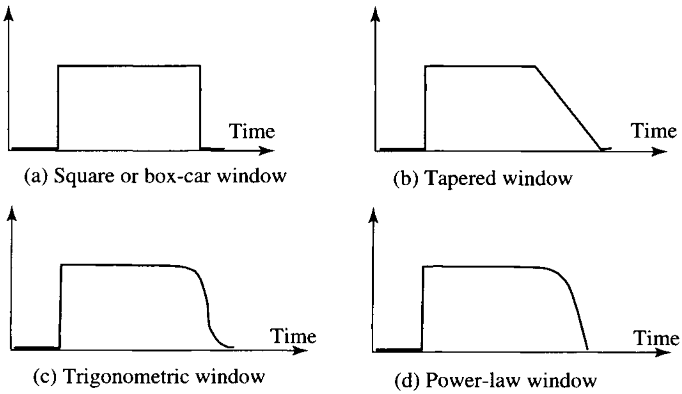


FIGURE 117 Window types for truncated signals

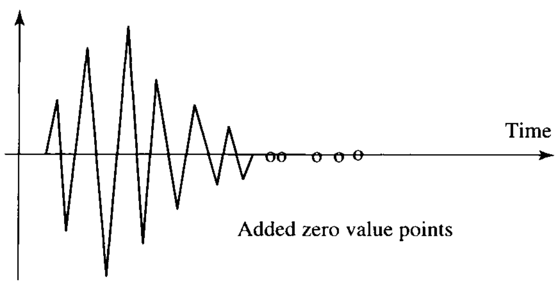


FIGURE 118 Pulse with zero padding

$2^9 = 512$ or $2^{10} = 1024$). It is then useful to extend the waveform by adding data points with the value 0 as illustrated in Fig. 118. This is called zero padding.

B. WHAT ARE THE TYPES OF SPECTRAL ANALYSIS?

Analysis of the spectral content of a waveform, i.e., the intensity of waves with all possible frequency components, provides a basis for physical interpretation of properties of an object under test. This is done through spectral analysis in the frequency domain using the Fourier transform, as discussed next and in Appendix 14. Major applications of the analysis include the following:

- Measurement of parameters—frequencies, spectral amplitude
- Waveform conditioning

- Feature extraction
- Dispersion analysis

Many of the analysis procedures used in the time domain discussed earlier can be performed efficiently in the frequency domain of spectral analysis.

What Is a Fourier Harmonic Series?

A periodic waveform is one which repeats itself over a finite duration, as sketched in Fig. 119. The frequency corresponding to this period is called the fundamental. Sine waves with frequencies which are multiples of the fundamental are called harmonics. The wave can be built up by combinations of the fundamental and harmonics with various amplitudes and phase shifts:

$$w(t) = \sum_{n=0}^{\infty} A_n \sin(2n\pi f t + \phi_n)$$

where A_n is the amplitude and ϕ_n is the phase shift of the nth component, as illustrated in Fig. 16.

The amplitudes are called the Fourier coefficients. Their units are the same as those of the measurements in the waveform (volts, mm/sec, etc.). The phase shifts may also be positive or negative and are in units of radians or degrees.

What Is a Fourier Spectrum?

An arbitrary (nonperiodic) waveform can be expressed as an integral combination of sines with all frequencies, in contrast to the restricted set of harmonics used in the Fourier series:

$$w(t) = \int_{-\infty}^{\infty} A(f) \sin[2\pi f t + \phi(f)] df$$

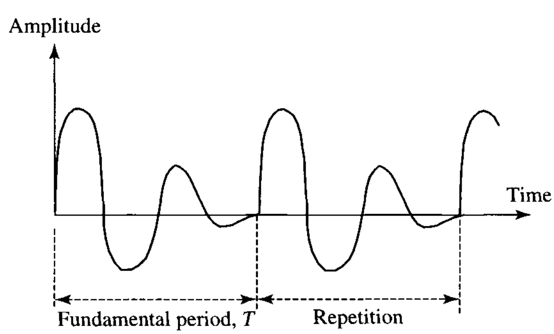


FIGURE 119 Periodic waveform

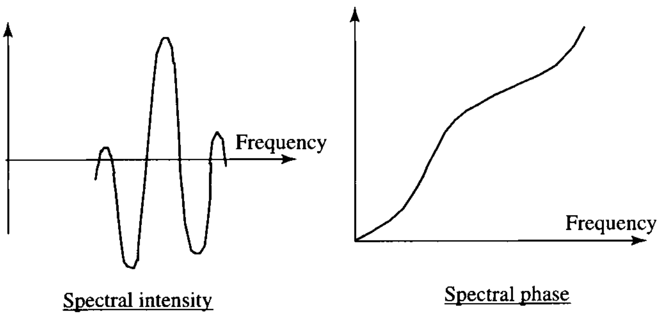


FIGURE 120 Spectra of a waveform

where $A(f)$ is the intensity spectrum and $\phi(f)$ is the phase spectrum, as sketched in Fig. 120. The frequency range extends to negative values denoting negative phases. For waves, these represent waves moving in opposite directions.

Because the integrand at a given frequency is the product of $A_n df$ with the sine function, the units of the spectral intensity A_n are those of amplitude divided by frequency (e.g., volts/Hz or mm/sec/Hz). Those of the spectral phase ϕ_n are radians or degrees.

What Is a Fourier Transform?

The intensity and phase spectra can be represented by the complex variable

$$F(\omega) = \mathrm{Re}[F(\omega)] + i\mathrm{Im}[F(\omega)],$$

where Re represents the real part, Im the imaginary part, and $i = (-1)^{1/2}$ is an imaginary unit. F is a function of the circular frequency $\omega = 2\pi f$ where f is the angular frequency. Again, the sine function can be replaced by the complex exponential $e^{-i\omega t} = \cos\omega t - i\sin\omega t$, so that the waveform time function $f(t)$ is given by the integral

$$f(t) = \int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega.$$

This integral is called the inverse Fourier transform, IFT, usually denoted by F^{-1}, which creates a time function from its Fourier Transform. The complex intensity function $F(\omega)$ can be derived through the integral

$$F(\omega) = (1/2\pi)\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

which is called the Fourier transform of the function $f(t)$. Note that some definitions of the transform pair, $F(t)$ and $F^{-1}(\omega)$, include a factor $1/(2\pi)^{1/2}$ in each, so that symmetric formulas are obtained for the transform pair, allowing for simplified algorithms. This factor should be considered when results derived by different software packages are being used.

The complex integrand can be expanded as

$$F(\omega)e^{-i\omega t} = (F_R \cos \omega t - F_I \sin \omega t) + i(F_R \sin \omega t + F_I \cos \omega t)$$

so that the integral is a real function of time, $f(t)$, if the imaginary part is zero. This occurs if the intensity F_R is symmetric in frequency about 0, and the intensity F_I is antisymmetric, so that, with the sine and cosine multipliers, they cancel and their integrals are zero. Then

$$\text{Re}[F(-\omega)] = \text{Re}[F(\omega)]$$

$$\text{Im}[F(-\omega)] = -\text{Im}[F(\omega)]$$

It is thus not necessary to quote the negative frequency parts of the spectra.

What Is a Fast Fourier Transform, FFT?

When the time-domain waveform is digital, the Fourier transform can be approximated by evaluating the time integral as a summation at discrete times $t_n = n\Delta t$ where n is an integer and Δt is the time interval, and at discrete frequencies $\omega_m = m\Delta\omega$ where m is another integer and $\Delta\omega$ is the frequency increment:

$$F_m = \sum_{n=0}^{N} f_n e^{-in\Delta t m\Delta\omega}.$$

Here n takes values from 0 to N, where N is the total number of time values, so that $t_{\max} = N\Delta t$. The set of N complex values (i.e., $2N$ numbers) is the discrete Fourier transform (DFT) of the time function f_n.

Because the N values of the time function produce N complex transform numbers (i.e., $2N$ numbers), all the information in a real time-domain pulse defined by N points is contained in two frequency-domain functions defined over half as many frequency points, $N/2$. Only half of the frequency range of the transform is useful, the other half being identical. Hence m takes values from 0 to $N/2$.

To simplify this summation, the lowest frequency, $f_{\min} = \Delta\omega/2\pi$, is taken to occupy the total time interval, $T = N\Delta t$, so that

$$f_{\min} = \Delta\omega/2\pi = 1/T = 1/N\Delta t, \text{ or } N\Delta t\Delta\omega = 2\pi$$

and the argument of the exponent is

$$mn\Delta t\Delta\omega = mn/N.$$

The range of frequencies, which is the maximum frequency, is then $f_{\max} = n\Delta\omega/2\pi = 1/\Delta t$.

Evidently fine frequency resolution demands a long time signal, $T = N\Delta t$. The sampling rate and number of points do not individually affect the frequency resolution, i.e., using more points, and thus a smaller time increment, in a fixed length of signal does not improve frequency resolution. There are other benefits to small time increments, however, as discussed later.

What Is the Nyquist Frequency?

Because N real values of a time function produce N complex transform numbers with $2N$ real values, and because the intensity spectrum of a real time function is symmetric in frequency (negative frequency gives the same values as positive), only half the frequency range of the transform is useful. All the information in a real time-domain pulse defined by N points is contained in two frequency-domain transforms defined over half as many points, $N/2$.

Frequencies above the center frequency of the transform, $f_N = 1/2\Delta t$, called the Nyquist frequency, reflect the same spectrum as those below.

What Is the Cooley–Tukey FFT Algorithm?

An efficient numerical procedure called the Cooley–Tukey algorithm is commonly used, based on an ingenious binary representation and manipulation of the arguments of the exponential, which are defined by the integers m and n:

$$mn\Delta\omega\Delta t = 2\pi mn/N.$$

N is taken as a power of 2, i.e., $N = 2^p$, where p is an integer typically between 6 and 12.

The transform produced this way is one estimate of the spectrum and contains errors described later. Another commonly used spectral estimate is the Maximum Entropy algorithm.

What are the Power Spectral Density (PSD) and Root Mean Square (RMS) Density?

When the phase is not important, the square of the intensity, $\text{PSD} = \text{Re}(F)^2 + \text{Im}(F)^2$, called the power spectral density, is sometimes used as an indication of the energy spectrum. The units of the PSD are amplitude squared divided by frequency (e.g., volts^2/Hz). The RMS is the square root of the PSD and so has the units of the amplitude divided by square root of frequency (e.g., volts/Hz$^{1/2}$).

What Is a White Spectrum (White Noise)?

An intensity spectrum which is uniform over all frequencies, i.e., $F(\omega) = \text{Const}$, is called white (from the analysis of white light, which contains light of all colors, i.e., frequencies, in equal amplitude and with random phases). A spectrum which is almost white is called pink.

What Is the Phase Spectrum?

The phase spectrum can be determined from the real and imaginary spectral functions:

$$\phi(\omega) = \text{atan}[\text{Im}(F)/\text{Re}(F)].$$

This definition generally results in phases which lie in the range $-\pi$ to π. Realistically, the phase can cover any value; for example, the phase can represent the propagated distance such that $\phi = \kappa x = 2\pi x/\lambda$, where κ is the wave number, $\lambda = c/f$ is the wavelength, c is the wavespeed, f is the frequency, and x is the distance. This relationship of phase to distance is its most useful aspect. It can be used to determine the wave number κ, and hence the wavespeed, when distance is known or it allows the propagated distance x to be determined when the wave number (or frequency and wavespeed) is known.

When the phase passes beyond the range $-\pi$ to π, it is truncated by numerical procedures and "wraps around" the range, as indicated in Fig. 121. It is then necessary to unwrap the values.

A common procedure for this unwrapping is to examine the slope of the wrapped phase. When the slope changes sign and the phase, extrapolated backward and forward to a common frequency point, changes by $\pm 2\pi$ over two or three frequency steps, it can be assumed to have jumped by $- \pm 2\pi$, as illustrated in Fig. 122. For all frequencies higher than each jump, the phase points can be adjusted by $- \pm 2N\pi$ where N is the number of prior jumps.

A common use of the spectral phase is for dispersion analysis (determining the frequency dependence of wavespeed) as described below.

What Are the Common Spectral Analyses?

An FFT of a waveform can be analyzed in several ways:

- Measurement of frequencies, spectral amplitude
- Smoothing, filtering
- Doppler shift

The FFTs of two or more waveforms can be analyzed in several ways:

- Cross-correlation (convolution)
- Transfer function
- Coherence

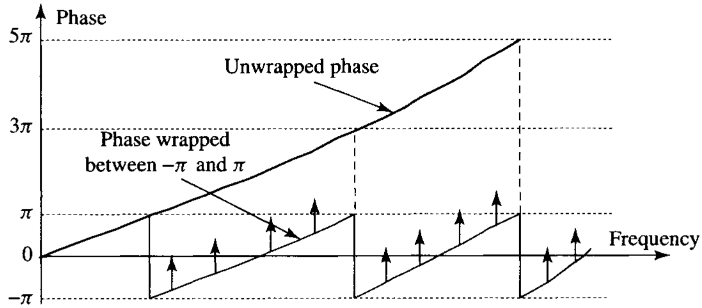


FIGURE 121 "Wrapped" and "unwrapped" phase spectra

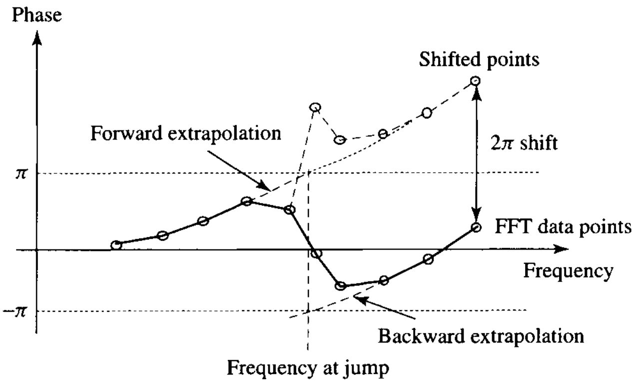


FIGURE 122 Unwrapping of phase

What Are the Common Spectral Measurements?

The major use of a spectrum is to identify those frequencies which contribute most to a waveform. Such frequencies can represent vibration modes of an object, which are governed by size, shape, and properties, or they can indicate the transit time of multiple passes of a reverberating wave. Any shift in these frequencies between measurements taken on a moving object and on the object when stationary define its velocity through the Doppler shift formula

$$v = c\Delta f/f$$

where c is the wavespeed.

The frequencies and spectral intensities can be identified by the gating/thresholding method described earlier for time-domain analysis.

Attenuation is determined as a function of frequency by obtaining spectra for a waveform at several distances of propagation. This can be done either by using objects of different size or by analyzing individually a sequence of echoes in a single object.

In the linearly elastic materials typical of UT test objects, there can be no transfer of energy between spectral components. Energy at one frequency can only be dissipated as heat, i.e., attenuation. Thus a change of intensity at a fixed frequency is due entirely to attenuation.

How Is a Waveform Smoothed , Filtered Spectrally?

A method of smoothing a waveform, alternative to the moving-average method in the time domain described earlier, is to filter out undesirable spectral components: the FFT of the waveform is windowed (see the earlier discussion of windowing) to remove high-frequency components, and an inverse FFT is performed.

What Is a Frequency-Domain Dispersion Analysis?

Application of the Fourier transform to a waveform described by the space-time function for a propagating wave $w(x,t) = Ae^{i(\kappa x - \omega t)} = Ae^{i\kappa x}e^{-i\omega t}$ results in a transform of the type $W(x,\omega) = e^{i\kappa x}F(\omega)$. The phase spectrum of this transform is then

$$\phi(f) = \kappa x + \phi_o = 2\pi f x/c + \phi_o$$

where $\tan \phi_o = F_I/F_R$ is the phase of the time transform. By determining the phase at several distances, x, the phase spectra at each frequency can be fitted to a straight line in distance, whose slope is $c(f) = 2\pi f/c(f)$.

C. WHAT STATISTICAL MEASUREMENTS ARE TAKEN?

Statistical measurements which are taken in either the time or the frequency domain are the following:

- Elementary statistics: means, standard deviations, histograms of variance
- Distributions, cumulative distributions
- Linear regression (curve fitting)
- Correlation (convolution)
- Feature analysis

What Are the Elementary Statistics?

Elementary statistical evaluations are commonly made when an object is scanned and variations over the object are to be quantified. For example, the wavespeed is measured over a grid of points on an object, and the results are to be compared to those from another similar object (see later discussions of C-scans and imaging).

A set of multiple measurements of a quantity with random variations (errors) is characterized by two statistics, defined in Appendix 13: the average, or mean,

$$\mu = (\Sigma m)/N,$$

of a set of nominally identical measurements m, and the standard deviation,

$$\sigma = \{[(\Sigma m^2) - N\mu^2]/(N-1)\}^{1/2}.$$

Note that σ is defined with a divisor of $(N-1)$ because the inclusion of the statistic μ reduces the effective number of measurements, or degrees of freedom of the measurements, by 1.

What Is a Histogram?

The measurements are assigned to intervals (bins) which span a small range of the variable, Δm_{bin}, and are centered at measurement values m_{bin}. The number of times the measurement falls into each of these bins is plotted against the center value for the bin, as illustrated in Fig. 123.

If the errors are random, the distribution of the measurements is modeled by the normal probability function $\exp[-(m_{\text{bin}} - \mu)^2/2\sigma^2]$ (called the bell curve). When the measurements conform to the bell curve, they represent two parameters: the mean, μ, and the standard deviation, σ, of random variations about the mean. The width of the bell where the number of occurrences is 0.6065 ($= e^{-1/2}$) times the peak is twice the standard deviation.

A convenient way to determine μ and σ is based on the cumulative distribution described next.

The extent to which the data depart from the bell curve reveals any nonrandomness in the deviation. For instance, if the data display a second hump to one side of the main hump, the measurements may represent two distinct parameters. This is called a binary distribution, as illustrated in Fig. 124. A common example is the measurement of wavespeed through an object containing a crack, where

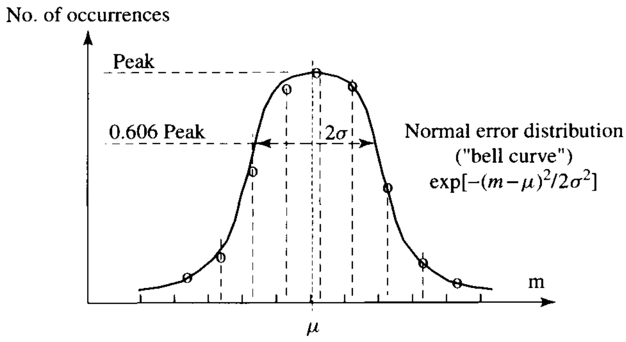


FIGURE 123 Statistics of typical measurements

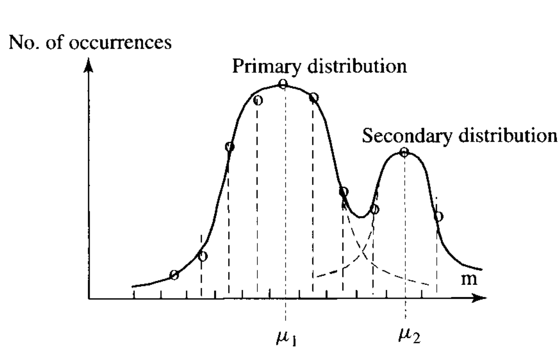


FIGURE 124 Illustration of two-parameter distribution

early reflections occur from the crack. These can be incorrectly interpreted as indicating higher wavespeed. An alternative conclusion is that there is a shorter path, i.e., the presence of a crack, giving a shorter travel time. The depth h_{cr} of the crack can be estimated to be $(\mu_2/\mu_1)h$, where h is the total thickness and μ_1 and μ_2 are the means of the "apparent wavespeeds" calculated from the formula $c_{app} = h/t$.

Another common statistic which is often used to describe a random variable is its root mean square (RMS) value,

$$R_{\mu=0} = [(\Sigma m^2)/N]^{1/2} = \{[(N-1)/N\sigma^2 + \mu^2)^{1/2}.$$

When the mean is zero, the RMS is almost the same as the standard deviation:

$$R_{\mu=0} = [(N-1)/N]\sigma.$$

What Is a Cumulative Distribution?

The histogram data on the number of occurrences of bin measurements can be summed from one bin to the next, starting with 0 at the lowest bin, signifying the total number which fall below the current bin. These accumulated numbers are divided by the total number of measurements and expressed as percentages, so that the highest bin has 100% of the measurements. Such a distribution is sketched in Fig. 125.

The integral of the normal distribution is a model for the cumulative distribution with random errors. The probable numbers for the mean (50%), and for one and two standard deviations away from the mean, are shown in the figure.

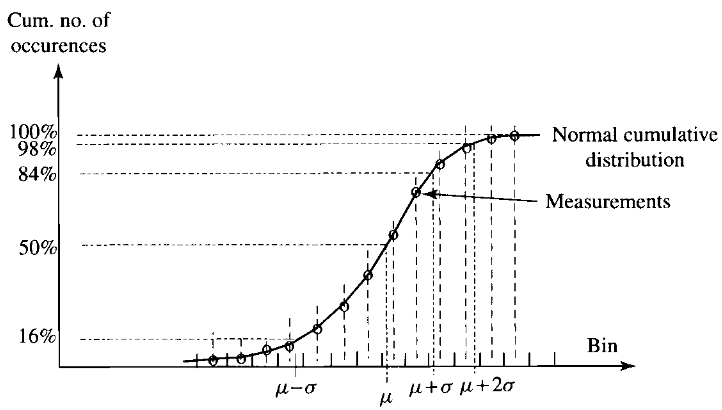


FIGURE 125 A cumulative distribution

What Is Linear Regression (Curve Fitting)?

The mean and standard deviation take no account of the sequence in which the measurements are taken, or any other factor which distinguishes each measurement. If, however, there is a relationship between each measurement, called the dependent variable, and a factor which defines the individual measurements, called the independent variable, these two statistics are not sufficient to represent the data. In that case, a theoretical model of the dependence on the independent variable must be considered.

The simplest model is a linear relationship between the measurements m and the independent variable, typically time t or a dimension such as thickness:

$$m_{\text{model}} = a + bt$$

where m_{model} is a hypothetical model of the measurement, and a and b are two constants to be determined statistically. This is done by considering the deviation Δ between each measurement and its model, as illustrated in Fig. 126.

$$(\Delta = m - m_{\text{model}}).$$

The following equations are obtained by minimizing the sum of the squares of the deviation, to give equal weight to positive and negative deviations, in variations of a and b (see Appendix 13):

$$a = \mu_m - b\mu_t$$

$$b = [(\Sigma mt - N\mu_m\mu_t)/(\Sigma t^2 - N\mu_t^2)]$$

where $\mu_{\text{m}} = (\Sigma m)/N$ is the mean of all N dependent variable measurements and $\mu_{\text{t}} = (\Sigma t)/N$ is the mean of all the independent variable values (e.g., times).

The regression procedure can be extended to model a dependence on multiple independent variables, for example, time, t, and thickness, h:

$$m = a + bt + ch.$$

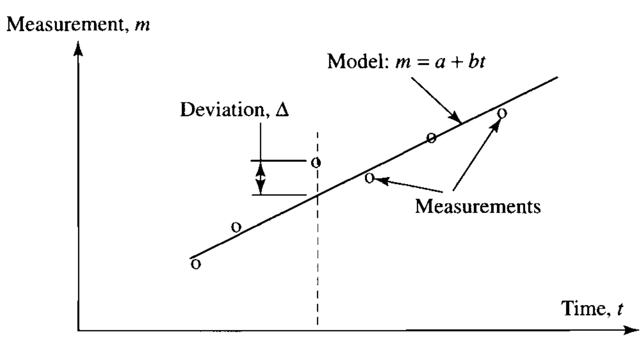


FIGURE 126 Illustration of deviations between data and statistical model

The procedure can also be used to represent nonlinear dependence by substituting functions of the independent variables, for example,

$$m = a + b\cos t + ch^2$$

where cos t and h^2 are treated as independent variables.

What Is Correlation?

Correlation is a way to provide a statistical definition of any similarity between two sets of measurements taken under the same conditions on two different parameters which contain errors.

Consider the measurement of two quantities X and Y. Several measurements of each are made giving the values $X_1, X_2, \ldots X_n$, and $Y_1, Y_2, \ldots Y_n$, with means μ_x and μ_y, as illustrated in Fig. 127.

The mean $|X|$ (and similarly $|Y|$) is

$$|X| = (1/n)\sum_i X_i.$$

The variance of the measurements of X is

$$\sigma_X^2 = [1/(n-1)]\sum(X_i - |X|)^2$$

where σ_X is the standard deviation for X.

The correlation between X and Y as determined from the measurements is

$$r = \left[\sum_i (X_i - |X|)(Y_i - |Y|)\right] \Big/ (n-1)\sigma_X\sigma_Y$$

$$= \left\{\left[\sum_i (X_i Y_i)\right] - n|X||Y|\right\} \Big/ (n-1)\sigma_X\sigma_Y.$$

Considering a waveform as a set of amplitude measurements, any similarity between two such waveforms w_A and w_B can be identified through correlation. This is done by summing the products of the values of the two waveforms for each time:

$$S = w_{A1}w_{B1} + w_{A2}w_{B2}\ldots$$

as illustrated in Fig. 106.

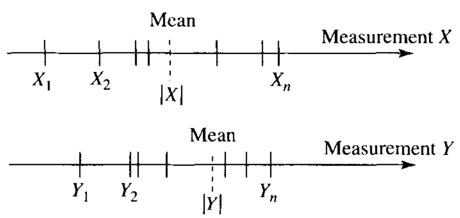


FIGURE 127 Illustration of multiple measurements

If the peaks of the two waveforms A and B do not correspond in time, but are otherwise similar, as illustrated in Fig. 128, then the correlation would be small. However, if waveform B were to be shifted until the peaks came into alignment, the correlation would be increased. This allows the construction of the correlation function, in which two waveforms are aligned successively at each data time point, and a correlation determined for that alignment. The resulting function, which has as many points as the waveforms, is the correlation function, as illustrated in Fig. 129.

Evidently a waveform can be correlated this way against itself, producing the autocorrelation function. This process identifies periodicity in the waveform, since correlation would be increased when a pulse aligns with a similar later pulse, as illustrated in Fig. 130.

At position A in the figure, the negative amplitudes in the shifted waveform align with positive amplitudes in the original waveform, giving a negative correlation. At position B, the positives coincide, giving a positive correlation.

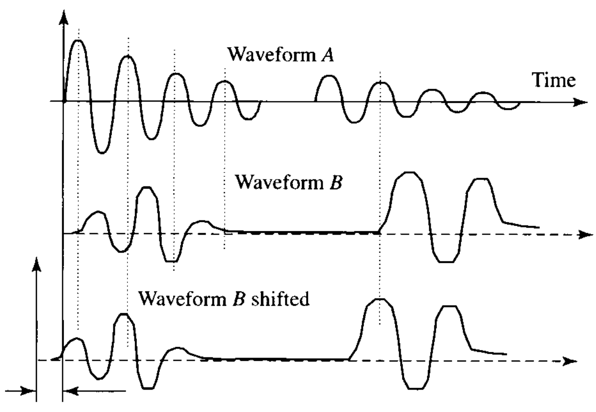


FIGURE 128 Correlation between two waveforms

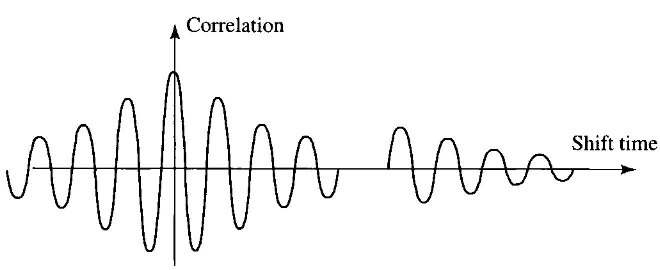


FIGURE 129 Correlation function between two waveforms

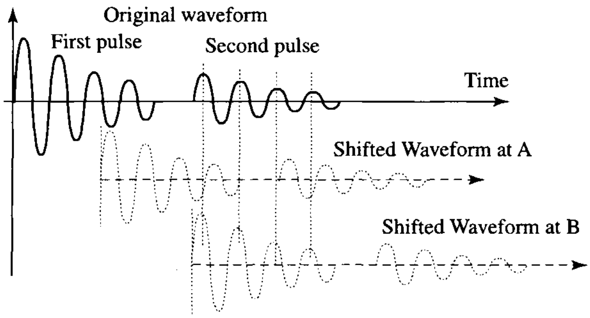


FIGURE 130 Illustration of waveform autocorrelation

The peak of an autocorrelation occurs at zero time shift, and the autocorrelation function is similar to the original waveform.

What Is Convolution?

A convolution is an integral over time of the shifting product of two time functions $f(t)$ and $g(t)$:

$$c_{f,g}(t) = \int_0^t f(\tau)g(t-\tau)d\tau$$

where t is the time shift, and τ is a dummy time variable denoting the integration over the product of $f(t)$ and the time-shifted function $g(t-\tau)$. This algorithm can be most conveniently evaluated in the frequency domain, as discussed later.

What Is Spectral Correlation?

The transform of the convolution between two time functions is

$$F\left\{\int f(t)g(t-\tau)d\tau\right\} = F^*(\omega) * G(\omega)$$

where the complex conjugate of a spectrum $F(\omega) = F_R(\omega) + iF_I(\omega)$ is defined by reversing the sign of its imaginary part:

$$F^*(\omega) = F_R(\omega) - iF_I(\omega).$$

Note that either spectrum can be taken in conjugate form since $F(\omega) * G^*(\omega) = F^*(\omega) * G(\omega)$. This is a point-by-point multiplication of the values of one spectrum with those of another which can readily be obtained.

Convolution is also used to define the response $R(t)$ of a system to an excitation or forcing function $F(t)$ through the product of the fundamental system response or transfer function, $f(t)$:

$$R(t) = F(0)f(t) + \int_0^t [(dF/dt)_\tau]f(t-\tau)d\tau$$

What Is a Transfer Function?

The transfer function (or fundamental response) $f(t)$ of a system is its response to a step function in time of unit magnitude. The response, $r(t)$, of the system to an excitation $F(t)$ is described by the convolution

$$r(t) = \int_0^t F(\tau) f(t-\tau) d\tau.$$

The formation of this integral is illustrated in Fig. 131.

These calculations can best be done in the frequency domain. The output waveform $w_{\text{out}}(t)$ from transmission of an input waveform $w_{\text{in}}(t)$ through a transfer function described by $f(t)$ is the inverse transform of the frequency domain convolution:

$$W_{\text{out}}(t) = F^{-1}\{W_{\text{in}}(\omega)^* F^*(\omega)\}$$

where F^{-1} denotes the inverse transform of the terms in brackets, $F(\omega)$ is the transform of the transfer function $f(t)$, and $W(\omega)$ is the transform of the input waveform $w_{\text{in}}(t)$.

This theorem has a most useful corollary: it can be inverted. Thus if the output waveform and the transfer function are known, the input waveform can be found:

$$w_{\text{in}}(t) = F^{-1}\{F(\omega)/W^*_{\text{out}}(\omega)\}$$

Similarly, the transfer function can be determined from measurements of the input and output waveforms:

$$F(t) = F^{-1}\{W^*_{\text{out}}(\omega)/W_{\text{in}}(\omega)\}$$

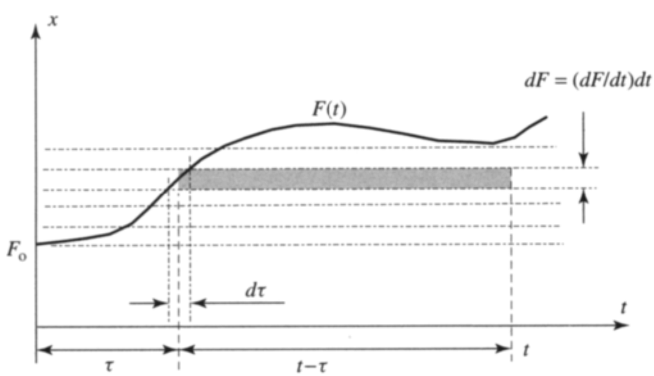


FIGURE 131 Illustration of convolution for response to an excitation

What Is a Waveform Feature?

A feature is a derived quantity whose values are related physically or correlated statistically to some quantitative aspect of an object under test and can be used to distinguish between two or more test configurations.

Simple features include the arrival or onset of a pulse (related to wavespeed), the largest peak (related to attenuation), and the dominant frequency (related to the size and shape of an object). A group of statistics can form a combined or a multivariable feature. Some typical features are the ratio of two peaks, the number of peaks, the number of zero crossings in a pulse, the width of the largest peak and the slope of its rise, the area under a peak, and its moment about the time axis. Such features may be created from physical reasoning, or they may have no obvious attribute. Each would be derived statistically as described next. Features are usually influenced by the measuring system.

The detection of a feature and its display is the most important facet of modern UT techniques and requires suitable software to analyze the waveform.

What Is Feature Extraction Analysis and Discrimination?

For each test configuration to be compared, typical waveforms (single or several averaged and processed) are acquired and analyzed to determine many parameters. Such test configurations could be different test objects or different test setups.

To illustrate the feature extraction process, consider two parameters, say time of flight t_f and peak amplitude A_p, for multiple tests on each of two test configurations 1 and 2. The data can be plotted as shown in Fig. 132.

For each test configuration, the means and standard deviations of the two parameters A_p and t_f are found and the bivariate normal probability distributions

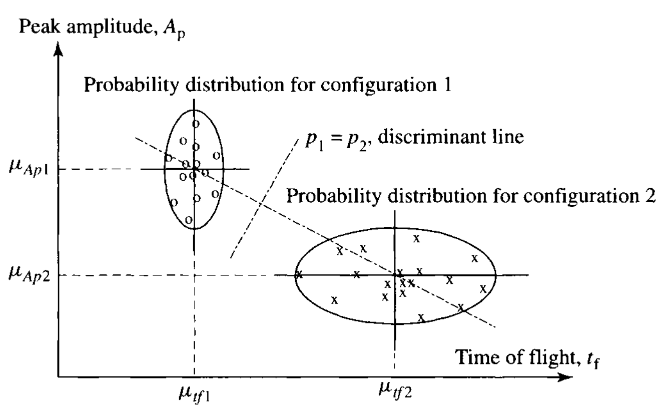


FIGURE 132 Illustration of multivariate statistical plot

for the two sets of data are set up:

$$p_1 = \exp[-(\delta_A^2/\sigma_A^2) - (\delta_t^2/\sigma_t^2)]$$

where $\delta_A = (A_p - \mu_A)$ and $\delta_t = (t_f - \mu_t)$, and similarly for p_2. These are the probabilities for a specified pair of measurements (A_p, t_f) belonging to one set or the other. Once these distributions are determined on sequences of tests on known configurations, similar measurements on an unknown configuration can be identified as belonging to one or the other set by comparing these probabilities for that pair of measurements.

The pair (A_p, t_f) belongs to the set 1 or 2 according to whether p_1 is greater or less than p_2. Consequently, a discrimination feature can be defined as

$$F_{12} = p_1/p_2, \text{ so that } F_{12} > 1 \text{ implies } p \text{ belongs to 1, and } < 1 \text{ to 2.}$$

Alternatively, the distance of a point (A_p, t_f) obtained in a new test from the two points representing the means for configurations 1 and 2 can be used as a feature. A line perpendicular to the line joining the means for configurations 1 and 2 can be found on which $p_1 = p_2$. This line is the boundary in the A_p–t_f space which separates the configurations. Such a line is called a discriminant.

When such an analysis is made on a group of measurements and a multi-dimensional analysis is made, a complex discriminant feature can be derived.

D. WHAT IS IMAGING?

Images are two-dimensional representations of the internal configuration of an object, i.e., a map of variations in some measurement. Direct display of wave-forms as described earlier (A- and B-scans) are not considered here to be images.

What Are the Types of Image?

Common images are:

- C-scan (or area scan)
- Tomography
- Steered image

What Is a C-Scan?

A sequence of waveforms is taken, by PE or by TTU, at points on a grid overlaid on the surface of an object, which is usually of large size in comparison to its thickness, such as a sheet or panel, as sketched in Fig. 133.

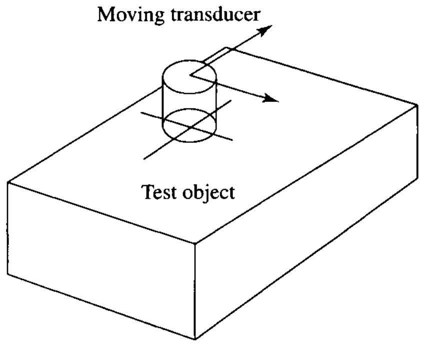


FIGURE 133 Illustration of test arrangement for C-scan

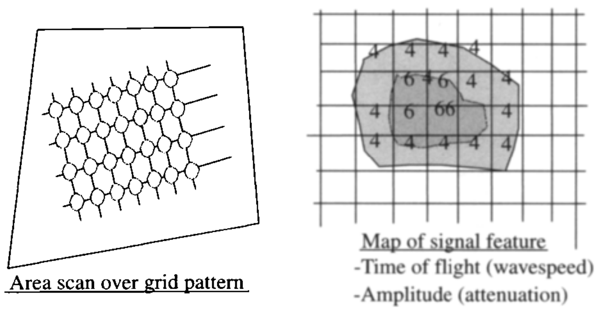


FIGURE 134 Illustration of a C-scan

What Is a Pixel?

Single feature data, usually travel time or peak signal, are derived from the waveforms and are assigned to the cells of an image of the grid (each grid box is referred to as a picture cell, or pixel). Contours of constant values of the data are connected to form a contour map which represents variation in the properties through the thickness of the entire object, as illustrated in Fig. 134. Any property obtainable from a waveform can be used in a C-scan, but most commonly the transit time or attenuation are used.

A C-scan can also be called a data map and can be enhanced by drawing contours and by using color. It can be updated continuously to show motion.

What Is Tomography?

A cross-sectional view of the interior properties of an object can be constructed from scans taken from many positions around its periphery, as illustrated in Fig. 135.

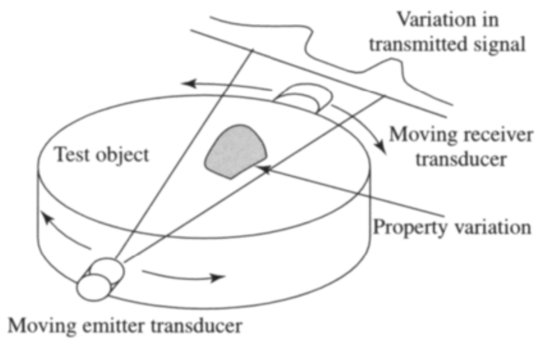


FIGURE 135 Illustration of test arrangement for tomographic scan

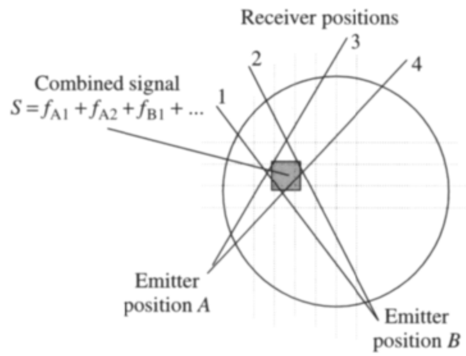


FIGURE 136 Illustration of image reconstruction

A single quantity is extracted from the waveform for each emitter–receiver path. These are added in each pixel with a weighting factor determined from the pixel-path geometry allowing for differing extent of participation. The result is assigned to the pixel for plotting as a tomographic scan, as illustrated in Fig. 136. This process is called back-projection reconstruction.

Tomography has been attempted in various UT applications, but its complications have precluded all but a few research efforts. It is used extensively in medicine as computed tomography (CT) (originally computer-aided tomography, CAT) with X-rays, or with microwaves in magnetic resonance imaging (MRI). Recordings of surface motions of the sun have been interpreted as sound waves and used to construct tomographic images of the solar interior.

What Is a Swept Image?

In medical UT, a linear array of transducers provides a steered sweeping beam which scans an object. It operates in the PE mode so that there is no separate moving receiver. The beam is steered by introducing successively increasing time delays across the transducer array (described earlier in Section 5, Hardware). The waveforms from each transducer are analyzed to produce one measure such as attenuation, and this is subjected to a form of tomographic reconstruction.

APPENDIX I

STRESS, STRAIN, AND ELASTICITY (ALSO VECTORS AND TENSORS)

A. ELASTICITY

The following discussion is intended to provide an intuitive description, yet retaining all facts pertinent to UT. Many texts (e.g., Love, 1944, or Sokolnikoff and Specht, 1946) deal with this subject in varying degrees of detail and completeness.

A material will deform when subjected to forces. If the deformation is dependent directly on the force, so that it follows any change in the force and disappears when the force is removed (reversibility), then the material is elastic. This elastic deformation may be linearly proportional to the force, typical of low to moderate forces, or it may be nonlinear but still reversible under high forces (in materials which do not yield plastically) as illustrated in Fig. 19 of Section 2.

Nonelastic deformations can appear under high forces, where some or all the deformation remains after removal of the force (plasticity, irreversibility), or when the deformation changes with time under constant, even zero, force (creep, viscosity, shrinkage). Deformation can appear to be nonlinear and irreversible under low or high forces when the force is not applied correctly (slippage, mechanical backlash, etc.) or when the measurement of force or deformation is inadequate (incorrectly initialized, etc.)

In linear elasticity, deformation δ (i.e., change of length, $l - l_0$) is proportional to force F:

$$F = K\delta$$

where K is the stiffness. This stiffness depends on extrinsic factors—the global configuration of the material, the distribution of the forces, and the measurement of deformation—in addition to the intrinsic nature of the material.

This relationship, called Hooke's Law, was discovered by Robert Hooke while studying clock springs. He published it in 1660 as the anagram "*ceiiiosssttuv*," which no one understood. In 1676 he published the corresponding Latin phrase "*Ut tensio sic vis*," which translates roughly into "As the stretch so the force" (Love, 1944). Generalizations of this law were developed over the following 200 years by such scientists as Navier and Cauchy. They are discussed in Appendix 2. Note that Sokolnikoff and Specht (1946) give the date of Hooke's first publication as 1676, and the second as two years later!

Commonly in engineering, tensile forces are taken to be positive while compressive forces are negative. Extension (stretch) is positive, and shortening (compression or contraction) is negative. Pressure as applied to a fluid, however, is usually taken as positive, contradicting the engineering definition.

The dependence of stiffness on extrinsic factors can be accounted for, leaving only intrinsic material properties, by defining local force and deformation measures: stress σ (force divided by area, F/A) and strain ε (essentially deformation divided by initial length, δ/l), as discussed later. In an elastic material. Hooke's Law relates these linearly:

$$\sigma = E\varepsilon$$

so that $K = EA/l$, where E is the elastic modulus, a material property, A is the area carrying the force, and l is the length. This relationship is generalized to account for three-dimensional aspects in Appendix 2.

B. STRESS

There are two types of force: surface and body forces. Surface forces are applied over all or part of a surface. Stress is defined as the surface force per unit area:

$$\sigma = F/A.$$

This elementary one-dimensional definition of stress is generalized into three dimensions later.

Body forces are applied to an interior-volume region of the material. Body forces are represented by force per unit volume and are not stresses. Inertia force

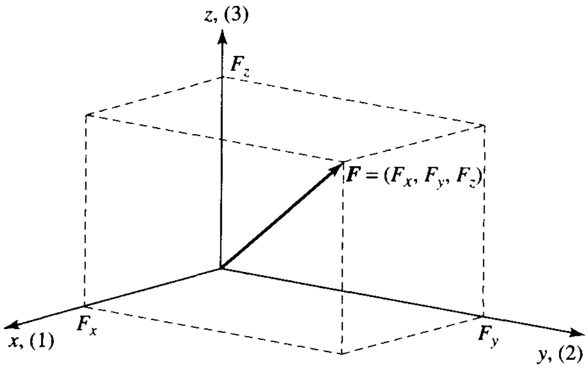


FIGURE 137 Directionality of force

(discussed in Appendix 4), which is proportional to density and acceleration, is an example of a body force:

$$f = \rho a$$

where f is a point body force, ρ is density, and a is acceleration. Other examples are gravitational and electromagnetic forces.

Force, Area, and Stress in Three Dimensions: Vectors and Tensors

To this point, directionality of force has not been considered. In fact, force, as well as area, are associated with directions: they are vectors (written here as bold symbols) and must be described by three components in relation to a designated coordinate system. For example, a force vector, **F**, has components F_x, F_y, F_z in a Cartesian coordinate system of x-, y-, z-axes, so that the indices x, y, z on force represent the components of force in the directions of the axes, as illustrated in Fig. 137.

A convenient shorthand way to describe the components of a vector is to use numbers 1, 2, or 3 to represent each of the three coordinate components. A letter, called an index, is then used to represent these numbers as variables when arbitrary directions are considered, essentially an algebraic notation. For example, the components of a force vector **F** can be represented by the single symbol F_i, with i taking values 1, 2, or 3 according to context, such that $\mathbf{F} = (F_1, F_2, F_3)$. Similarly, area is $\mathbf{A} = A_i$, where the directionality is that of the normal to the surface, as illustrated in Fig. 138.

In general, the force may not be normal to the area, so that the components of force and area do not lie parallel and must be described separately. The components are given different letters, e.g., i and j, since these can each take a value of 1, 2, or 3 independently in considering different components of force and area. The components of stress are then

$$\sigma_{ij} = F_i / A_j.$$

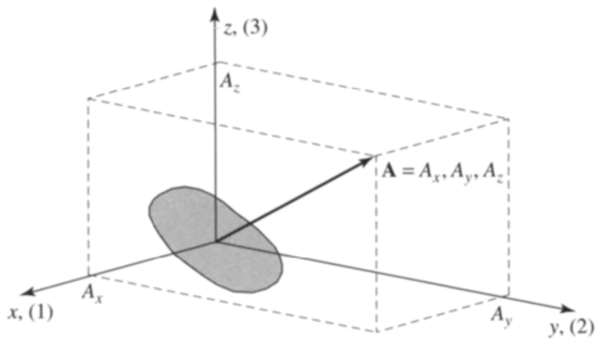


FIGURE 138 Directionality of area

(Note that, to purist mathematicians, it is not correct to divide anything by a vector. For the present purpose, however, the result is sufficient and is descriptive.)

The components of stress form a 3 × 3 matrix:

$$\sigma_{ij} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$$

Stress is a quantity which involves two vector directions and is called a tensor.

The description of a physical quantity by a vector or a tensor is intimately tied to the use of a coordinate system. A different coordinate system results in different components of the same physical quantity.

Change of Coordinate System

In an alternate coordinate system, where the axes are each at some angle (possibly all different) to the original axes, each of the new axes can be described by a unit vector u_i (i.e., one having unit length). In a Cartesian coordinate system, the length of a vector v_i is given by the Euclidean formula

$$L = (v_x^2 + v_y^2 + v_z^2)^{1/2} = \sum_{i=1}^{3} (v_i v_i)^{1/2} = (v_i v_i)^{1/2}$$

The third equation introduces Einstein's shorthand summation convention in which the repeated index i indicates a summation over all three values, 1 to 3, the summation symbol being dropped. In that case, the letter used for an index is not unique, and any other letter can be substituted. Unless specifically stated, a letter can never be used more than twice in any expression. Then for a unit vector,

$$u_i u_i = 1.$$

This has components in the original system, as illustrated in Fig. 139.

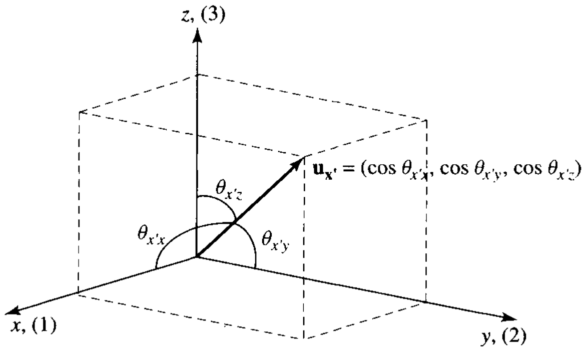


FIGURE 139 Definition of new coordinate axes

A unit vector $\mathbf{u}_{\mathbf{x}'}$ which defines a new x-axis has the components

$$\mathbf{u}_{\mathbf{x}'} = (u_{x'x}, u_{x'y}, u_{x'z}) = (\cos\theta_{x'x}, \cos\theta_{x'y}, \cos\theta_{x'z})$$

where $\theta_{x'x}$, etc., are the angles between the new x' axis and the three original axes, x, y, and z. The new coordinate system is described by three such unit vectors each having three components, forming a 3×3 matrix. The matrix of direction cosines is written as the matrix $l_{i'j}$,

$$l_{i'j} = \begin{pmatrix} \cos\theta_{x'x} & \cos\theta_{x'y} & \cos\theta_{x'z} \\ \cos\theta_{y'x} & \cos\theta_{y'y} & \cos\theta_{y'z} \\ \cos\theta_{z'x} & \cos\theta_{z'y} & \cos\theta_{z'z} \end{pmatrix}.$$

The angles $\theta_{y'x}$, etc., are the angles between the new y' axis and the three original axes. In general, the new axes may not be orthogonal (at right angles to one another) so that the angles are not necessarily symmetric:

$$\theta_{x'y} \neq \theta_{y'x}.$$

For rotation of the axes through an angle θ about an axis defined by the unit vector $a_i (a_i a_i = 1)$, the directional cosines are

$$l_{ij} = \delta_{ij} + a_i a_j (1 - \cos\theta) + \varepsilon_{ijk} a_k \sin\theta$$

where ε_{ijk}, called the alternating symbol, has the following values:

+1 when i, j, k are in cyclic sequence of 1, 2, 3 (1, 2, 3 or 2, 3, 1 or 3, 1, 2)

−1 when i, j, k are in countercyclic sequence of 1, 2, 3 (3, 2, 1 or 2, 1, 3 or 1, 3, 2)

0 when any two or more of i, j, k are the same.

For small angles, say $\theta < 10°$, so that $\cos\theta \sim 1$ and $\sin\theta \sim \theta$ this expression is simplified to $l_{ij} \approx \delta_{ij} + \varepsilon_{ijk}\theta_k$, where $\theta_k = a_k\theta$ called the rotation vector, which is a rotation through an angle θ about an axis a_i.

The components of a vector, such as force, in the new coordinate system are then given by the contributions of each of its original components along the new axes:

$$F_{x'} = F_x \cos\theta_{x'x} + F_y \cos\theta_{x'y} + F_z \cos\theta_{x'z}, \text{ etc.}$$

The transformation of the components of force is then

$$F_{i'} = \sum_{j=1} l_{i'j} F_j = l_{i'j} F_j$$

This transformation formula is used as a test to ascertain whether a group of three quantities is a vector. For example, whereas the three components of a force form a vector, a group of three measurements of distance in one direction do not.

Properties of Stress

It can be shown through consideration of angular momentum that the stress is symmetric:

$$\sigma_{ij} = \sigma_{ji}$$

so that only six of the nine possible components are independent. They are the normal stresses σ_{11}, σ_{22}, and σ_{33}, and the shear stresses $\sigma_{12} = \sigma_{21}$, $\sigma_{23} = \sigma_{32}$, $\sigma_{31} = \sigma_{13}$. In normal stress, the force component is parallel to the component of the surface normal. In shear stress, the force is perpendicular to the normal (in the direction of one of the other coordinates) and thus they are tangential to the surface. Shear stress is often designated by the symbol τ_{ij}.

The average of the normal stresses, called the hydrostatic stress, is the negative of the pressure:

$$(1/3)(\sigma_{11} + \sigma_{22} + \sigma_{33}) = \sigma_{ii}/3 = -p(\text{i.e., } p = -3\sigma_{ii}).$$

A purely hydrostatic stress has three equal normal components and no shear components and can be written as

$$\sigma_{ij} = -p\delta_{ij}$$

where δ_{ij} is called the Kronecker delta or the unit tensor, having components equal to 1 when the two indices are the same, and zero when not:

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The algebraic difference between the stress and its hydrostatic part is called the stress deviator (note that pressure is negative in this sign convention):

$$s_{ij} = \sigma_{ij} + p\,\delta_{ij}.$$

The trace of a matrix is the sum of its diagonal elements:

$$\delta_{ii} = 3 \text{ (for three dimensions)},$$

$$\sigma_{ii} = -3p.$$

The trace of the deviator is zero by the definition of pressure. The deviatoric stress is essentially a shear stress. The stress can then be separated into two parts, the hydrostatic and the deviatoric:

$$\sigma_{ij} = -p\,\delta_{ij} + s_{ij}.$$

Stress in Rotated Axes

It is often necessary to convert the stress components to coordinates different from those under which they have been determined, such as in a rotated system.

Both the force and area features of the components of stress must be transformed in a similar manner, so that the components of stress are

$$\sigma_{i'j'} = l_{i'k} l_{j'l} \sigma_{kl}.$$

This transformation is the definition of a second-rank tensor. A vector is a first-rank tensor.

Two-Dimensional Case: Mohr's Circle and Principal Stress

When only two dimensions are considered, there are three of the foregoing transformation equations, which can be written out as

$$\begin{aligned}
\sigma_{x'x'} &= \sigma_{xx}\cos^2\theta + \sigma_{yy}\sin^2\theta + 2\sigma_{xy}\sin\theta\cos\theta \\
&= (1/2)(\sigma_{xx} + \sigma_{yy}) + (1/2)(\sigma_{xx} - \sigma_{yy})\cos 2\theta + \sigma_{xy}\sin 2\theta \\
\sigma_{y'y'} &= \sigma_{yy}\cos^2\theta + \sigma_{xx}\sin^2\theta + 2\sigma_{xy}\sin\theta\cos\theta \\
&= (1/2)(\sigma_{xx} + \sigma_{yy}) - (1/2)(\sigma_{xx} - \sigma_{yy})\cos 2\theta + \sigma_{xy}\sin 2\theta \\
\sigma_{x'y'} &= \sigma_{xy}(\cos^2\theta - \sin^2\theta) = \sigma_{xy}\cos 2\theta
\end{aligned}$$

where $\theta x'x = \theta y'y = \theta$ is the angle between the x and x' or y and y' axes, and $\theta x'y = \theta y'x = \pi/2 - \theta$.

These relationships are conveniently represented by Mohr's circle illustrated in Fig. 140. A circle is drawn through the points $(\sigma_{xx}, \sigma_{xy})$ and $(\sigma_{yy}, -\sigma_{xy})$, on a center at $(1/2)(\sigma_{xx} + \sigma_{yy})$.

The intercept of a diameter drawn at an angle 2θ with the circle defines the normal stresses $\sigma_{x'x'}$ and $\sigma_{y'y'}$ and the shear stress $\sigma_{x'y'}$ as shown. The circle shows that the normal stresses have a maximum and a minimum value at an angle where there is no shear stress. These are called the principal stresses, and the angle is the principal angle. The shear stress has a maximum when the two normal stresses are equal and occurs at a 45° angle.

Three such circles can be combined to represent a three-dimensional state.

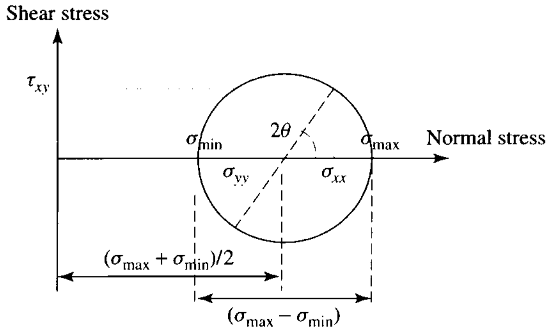


FIGURE 140 Mohr's circle of stress

C. STRAIN

Strain is defined as the ratio of deformation to original length:

$$\varepsilon = \delta / l_o.$$

When the deformation is a displacement field $u(x)$ which varies over a region of the material, a local strain is defined from the difference of displacement at two ever-closer points on an element of position dx:

$$\varepsilon = dx \overset{\lim}{\longrightarrow} 0 \left\{ \frac{u(x+dx) - u(x)}{dx} \right\} = \frac{\partial u}{\partial x}$$

The definition of strain can be extended to displacements with three components, and with variation in three directions. Because of energy considerations, this definition of strain, like stress, must be symmetric:

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$$

Properties of Strain

There are only six independent components of strain. The components where the variation is in the same direction as the displacement, i.e., $\varepsilon_{11} = \partial u_1/\partial x_1$, $\varepsilon_{22} = \partial u_2/\partial x_2$, and $\varepsilon_{33} = \partial u_3/\partial x_3$, are called direct or axial strains. Those where variation is across the direction of displacement, i.e., $\varepsilon_{23} = \varepsilon_{32} = (1/2)(\partial u_2/\partial x_3 + \partial u_3/\partial x_2)$, $\varepsilon_{31} = \varepsilon_{13} = (1/2)(\partial u_3/\partial x_1 + \partial u_1/\partial x_3)$, and $\varepsilon_{12} = \varepsilon_{21} = (1/2)(\partial u_1/\partial x_2 + \partial u_2/\partial x_1)$, are called shear strains. Shear strain is sometimes represented by the symbol γ_{ij} instead of ε_{ij}. Axial strain represents changes in length whereas shear strain represents changes in angle. Note that rotation changes the orientation of lines, but not the angle between them.

The sum (not average) of the three axial strains is the volumetric strain:

$$\varepsilon_{ii} = \Delta V / V_0.$$

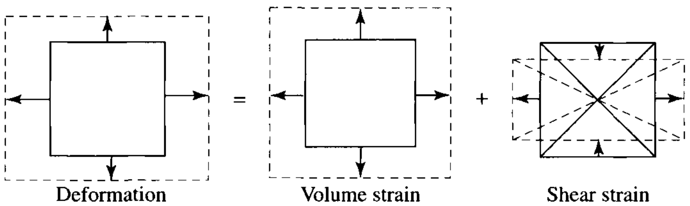


FIGURE 141 Decomposition of deformation

Volume strain is sometimes represented by the condensation, $\theta = -\Delta V/V_0$, which is taken as positive for compression, i.e., reducing volume.

Any deformation can be separated into the sum of a volume change and a shear deviator,

$$\varepsilon_{ij} = (\Delta V/V_0)\delta_{ij} + d_{ij}$$

as illustrated in Fig. 141.

The preceding discussion relates to small (infinitesimal) strains typical of ultrasonics. It can be redefined to include large (finite) strains using the methods of nonlinear differential geometry.

The time derivative of the strain shows that a velocity gradient produces a strain rate:

$$\partial\varepsilon_{ij}/\partial t = (\partial v_i/\partial x_j + \partial v_j/\partial x_i)/2$$

where $v_i = \partial u_i/\partial t$ is the velocity.

Rotation

Strain considers only the symmetric part of the derivatives of deformation. The antisymmetric part is the rotation:

$$\omega_{ij} = (1/2)(\partial u_i/\partial x_j + \partial u_j/\partial x_i).$$

The deformation gradient is then the sum of the strain and rotation:

$$\partial u_i/\partial x_j = \varepsilon_{ij} + \omega_{ij}.$$

In general, a deformation field produces both a strain and a rotation, and a pure strain or a pure rotation are unusual.

Rotation is a second-rank tensor. A pseudo-vector, or polar vector (also called the dual), can be defined using the alternating symbol ε_{ijk}:

$$\omega_i = \varepsilon_{ijk}\omega_{jk} = \varepsilon_{ijk}\partial u_j/\partial x_k.$$

Strain in Rotated Axes

When the strain is measured in different directions than used for an analysis, it is necessary to convert one to the other. The rotation of axes then leads to strain components which obey the same formulas used for stress.

D. ALTERNATIVE INDICIAL NOTATION

An alternate notation for stress and strain is defined to take advantage of the symmetry of the stress and strain components. It describes the six independent components of stress or strain, which use a pair of indices ranging from 1 to 3, by a single index which ranges from 1 to 6. Thus the matrix of three-dimensional components is represented by a six-dimensional group of components. The correspondence between these is as follows:

σ_{ij}:	ij	=	11	22	33	23	31	12
σ_a:	a	=	1	2	3	4	5	6

In this notation, the components of stress, for example, are

$$(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6) = \begin{pmatrix} \sigma_1 & \sigma_6 & \sigma_5 \\ \sigma_6 & \sigma_2 & \sigma_4 \\ \sigma_5 & \sigma_4 & \sigma_3 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix}.$$

This notation does not support coordinate transformation relationships. Thus the six-dimensional component array is not a six-dimensional vector with regard to transformation of coordinates. A typical transformation of a component of stress is as follows:

$$\sigma_1' = l^2\sigma_1 + m^2\sigma_2 + n^2\sigma_3 + 2lm\sigma_6 + 2mn\sigma_4 + 2nl\sigma_5$$

where (l, m, n) are the direction cosines of the vector transforming the 1-axis, $(l_{11}, 1_{12}, l_{13})$.

The correspondence between the two index systems can be written, for purposes of numerical calculations, as follows:

$i, j \rightarrow$ a: if $i = j$, then $a = i = j$
else $a = 9 - i - j$

$a \rightarrow i, j$: If $a < 4$ then $i = j = a$
else $i = a - 2$, unless $a = 6$, then $i = a - 5$
and $j = a - 4$, unless $a = 4$, then $i = a - 1$

APPENDIX 2

THE GENERALIZED HOOKE'S LAW

A. HOOKE'S LAW

Hooke's Law is generalized to relate linearly all components of stress to all components of strain and is written as

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl}, \text{ or } \sigma_a = C'_{ab}\varepsilon_b$$

where C_{ijkl} in the three-dimensional tensor form, or C'_{ab} in the alternate non-tensor six-dimensional form in the notation of Appendix 1, is called the elastic stiffness.

Examples of terms in the two forms are as follows:

$$\begin{array}{llll} C' & = C_{1111}\ C_{1122}\ C_{1133}\ C_{1123}\ C_{1131}\ C_{1112} & = & C_{11}\ C_{12}\ C_{13}\ C_{14}\ C_{15}\ C_{16} \\ & \ C_{2211}\ C_{2222}\ C_{2233}\ C_{2223}\ C_{2231}\ C_{2212} & & C_{12}\ C_{22}\ C_{23}\ C_{24}\ C_{25}\ C_{26} \\ & \ \text{etc.} & & \text{etc.} \end{array}$$

These equations can be inverted:

$$\varepsilon_{ij} = S_{ijkl}\sigma_{kl}, \quad \text{or } \varepsilon_a = S'_{ab}\sigma_b$$

where S_{ijkl} or S'_{ab} is called the compliance. It is a historic absurdity that stiffness is denoted by C and compliance by S!

The expressions for these inverses, in cases up to orthotropic, are as follows:

$$\begin{array}{lll}
C_{11} = (S_{22}S_{33} - S_{23}^2)/S & C_{12} = (S_{13}S_{23} - S_{12}S_{33})/S & C_{44} = 1/S_{44} \\
C_{22} = (S_{33}S_{11} - S_{13}^2)/S & C_{13} = (S_{12}S_{32} - S_{13}S_{22})/S & C_{55} = 1/S_{55} \\
C_{33} = (S_{11}S_{22} - S_{12}^2)/S & C_{23} = (S_{21}S_{31} - S_{23}S_{11})/S & C_{66} = 1/S_{66}
\end{array}$$

where $S = S_{11}S_{22}S_{33} - S_{11}S_{23}^2 - S_{22}S_{31}^2 - S_{33}S_{12}^2 + 2S_{12}S_{23}S_{31}$. (Compare also the relationships given below.)

The stiffness C_{ijkl} appears to have $3^4 = 81$ components, but because of the symmetries of stress and strain expressed by the six-component notation, C'_{ab} can have no more than $6^2 = 36$ components. Furthermore, energy considerations show that C'_{ab} must itself be symmetric so that there are at most 21 independent components (the 6 components which have a pair of the same index, and half of the remaining 30, i.e., 15, which have pairs of differing indices).

The most general form of material has 21 independent stiffness components. These are some crystals and some man-made materials, in which the material structure exhibits strong directional dependence with three nonorthogonal axes (axes that are not perpendicular). They are called anisotropic, aeolotropic, or rhombohedral materials. Symmetries of the material structure usually reduce the number of independent components considerably. Most common materials require many fewer components, and these may be described by a small number of independent parameters.

The simplest solid is isotropic, in which there is no identifiable direction, requiring only two properties to define nine nonzero components. A fluid (liquid) can be described by only one property and three components. These and other classes of materials are described later.

B. STIFFNESS IN ROTATED AXES

When the elastic coefficients must be used in a different set of axes than the natural axes relating to the material, a rotational transformation must be used.

The general expression for rotation about an arbitrary axis at angles θ_x, θ_y, θ_z to the three axes can be written with the use of the matrix of direction cosines l_{ij} (defined in Appendix 1) as follows:

$$C_{i'j'k'l'} = C_{pqrs}l_{i'p}l_{j'q}l_{k'r}l_{l's}.$$

In the alternative notation, the stiffness is given by the following equation:

$$C'_{a'b'} = L_{a'c}L_{b'd}C'_{cd}$$

where the 6 × 6 matrix $L_{a'b}$ has the following components in terms of three direction cosines l_i, m_i, n_i:

$$L = \begin{pmatrix} l_1^2 & m_1^2 & n_1^2 & 2m_1n_1 & 2n_1l_1 & 2l_1m_1 \\ l_2^2 & m_2^2 & n_2^2 & 2m_2n_2 & 2n_2l_2 & 2l_2m_2 \\ l_3^2 & m_3^2 & n_3^2 & 2m_3n_3 & 2n_3l_3 & 2l_3m_3 \\ l_2l_3 & m_2m_3 & n_2n_3 & m_2n_3+m_3n_2 & n_2l_3+n_3l_2 & l_2m_3+l_2m_3 \\ l_3l_1 & m_3m_1 & n_3n_1 & m_3n_1+m_1n_3 & n_3l_1+n_1l_3 & l_3m_1+l_1m_3 \\ l_1l_2 & m_1m_2 & n_1n_2 & m_1n_2+m_2n_1 & n_1l_2+n_2l_1 & l_1m_2+l_2m_1 \end{pmatrix}.$$

As an example, the coefficients C'_{11}, C'_{12}, and C'_{13} are transformed by a rotation θ about the z- (or 3-) axis into $C'_{1'1'}$, $C'_{1'2'}$, and $C'_{1'3'}$ as follows:

$$\begin{aligned} C'_{1'1'} &= C'_{11}\cos^4\theta + 2C'_{12}\sin^2\theta\cos^2\theta + C'_{22}\sin^4\theta \\ &\quad -4C'_{16}\cos^3\theta\sin\theta - 4C'_{26}\sin^3\theta\cos\theta + 4C'_{66}\sin^2\theta\cos^2\theta \\ C'_{1'2'} &= (C'_{12} + C'_{22})\sin^2\theta\cos^2\theta + C'_{12}(\sin^4\theta + \cos^4\theta) \\ &\quad +2(C'_{16} - C'_{26})\sin\theta\cos\theta(\cos^2\theta - \sin^2\theta) - 4C'_{66}\sin^2\theta\cos^2\theta \\ C'_{1'3'} &= C'_{13}\cos^2\theta + C'_{23}\sin^2\theta - 2C'_{36}\sin\theta\cos\theta. \end{aligned}$$

Note that some of these formulas contain fourth powers of sines and cosines in several combinations. This makes for difficult analysis of stiffness data taken from various angles to obtain the stiffnesses in material axes. Convenient relationships between stiffnesses and single terms of sines and cosines, which can be used in direct linear data fitting, are as follows:

$$\begin{aligned} C'_{11} + C'_{13} &= [(C_{13} + C_{33}) + (C_{11} - C_{33})\cos^2\alpha] \\ C'_{12} &= [C_{32} + (C_{12} - C_{32})\cos^2\alpha] \\ C'_{13} &= [C_{13} + (C_{11} + C_{33} - C_{55} - 2C_{12})\sin^2\alpha\cos^2\alpha]. \end{aligned}$$

Identical relationships hold for compliances.

C. SPECIAL CLASSES OF ELASTIC MATERIALS

The following classes of material, described in the six-index form, are common.

1. Fluid

A fluid does not sustain shear stress: it can only support a stress with normal components equal in all directions, i.e., a pressure. Shear deformation is unrestricted except by geometric constraints. There is no shear elasticity, and only one independent elastic constant—the compressibility, K.

The stiffness (and the compliance) has three components:

$$C'_{11} = C'_{22} = C'_{33} = K.$$

All other components are zero. This can be regarded as a special case of an isotropic solid (see later discussion) with Poisson's ratio $\nu = -1$, so that the shear modulus is $G = 0$.

Only triaxial hydrostatic stress is possible:

$$p = -K\Delta V/V = -K\partial u_i/\partial x_i.$$

The deformation is not restricted. In uniaxial deformation (for example, flow along the x-axis in a pipe or a plane wave propagating along the x-axis), the stress–strain relationship is

$$p = -K\partial u_x/\partial x.$$

2. Isotropic Solid

A solid can sustain shear stress, responding with a definite shear strain. Also, Poisson showed that the strain transverse to an applied stress is not the same as that in the direction of the stress (see Section 2). A solid thus has one or more elastic constants related to shear response.

In an isotropic solid there is no directionality and only two independent elastic constants—commonly taken as Young's modulus E and Poisson's ratio ν.

The compliance has nine components:

$$S_{11} = S_{22} = S_{33} = 1/E$$

$$S_{12} = S_{23} = S_{31} = -\nu/E$$

$$S_{44} = S_{55} = S_{66} = 2(1+\nu)/E$$

so that

$$\varepsilon_1 = (1/E)[\sigma_1 - \nu(\sigma_2 + \sigma_3)]$$

(and similarly for the 2 and 3 components),

$$\varepsilon_4 = [2(1+\nu)/E]\sigma_4$$

(and similarly for the 5 and 6 components).

The stiffness is

$$C'_{11} = C'_{22} = C'_{33} = E(1-\nu)/(1+\nu)(1-2\nu)$$

$$C'_{12} = C'_{23} = C'_{31} = E\nu/(1+\nu)(1-2\nu)$$

$$C'_{44} = C'_{55} = C'_{66} = E/2(1+\nu)$$

so that

$$\sigma_1 = [E(1-\nu)/(1+\nu)(1-2\nu)]\{\varepsilon_1 + [\nu/(1-\nu)](\varepsilon_2 + \varepsilon_3)\}$$

(and similarly for the 2 and 3 components),

$$\sigma_4 = [E/2(1+\nu)]\varepsilon_4$$

(and similarly for the 5 and 6 components).

The stiffness $E(1-\nu)/(1+\nu)(1-2\nu)$ is referred to later as the longitudinal wave stiffness C_l because of its role in longitudinal waves, as discussed in Appendices 3, 6, and 10.

Any two of the following alternate moduli can be used:

$$\text{Shear modulus } G = E/2(1+\nu)$$

$$\text{Bulk modulus } K = E/3(1-2\nu)$$

$$\text{Lamés constants } \lambda = E\nu/(1+\nu)(1-2\nu) \text{ and } \mu = G.$$

Some useful relationships among these moduli are

$$E = 2G(1+\nu) = 3K(1-2\nu) = 9KG/(3K+G)[= 3/\{(1/3K)+(1/G)\}]$$
$$= \mu(3\lambda+2\mu)/(\lambda+\mu)$$
$$G = \mu = E/2(1+\nu) = 3K(1-2\nu)/2(1+\nu) = \lambda(1-2\nu)/2\nu$$
$$\nu = \lambda/2(\lambda+\mu) = (E/2G) - 1 = (3K-2G)/2(3K+G)$$
$$\lambda = G(E-2G)/(3G-E) = K - 2G/3 = E\nu/(1+\nu)(1-2\nu)$$
$$K = E/3(1-2\nu) = \lambda + 2\mu/3$$
$$C_l = C'_{11} = E(1-\nu)/(1+\nu)(1-2\nu) = \lambda + 2\mu = K + (4/3)G$$
$$\lambda + \mu = E/2(1+\nu)(1-2\nu)$$
$$C_l/G = 2(1-\nu)/(1-2\nu)$$

The Lamé coefficients allow the isotropic stress–strain relationship to be put into a simple and convenient tensor form (Cauchy's law):

$$\sigma_{ij} = \lambda\varepsilon_{kk}\delta_{ij} + 2\mu\varepsilon_{ij}$$

so that

$$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}).$$

Limitations

Energy considerations which require that the shear modulus G and the compressibility K be positive show that Poisson's ratio is restricted to the range

$$-1 < \nu < 1/2.$$

Negative values are thus admissible, but not common.

3. Cubic Symmetry

This is similar to the isotropic material, but the shear modulus is not related to Poisson's ratio. There are three independent elastic constants—Young's modulus E, Poisson's ratio ν, and shear modulus G.

The compliance has nine components:

$$S_{11} = S_{22} = S_{33} = 1/E$$
$$S_{12} = S_{23} = S_{31} = -\nu/E$$
$$S_{44} = S_{55} = S_{66} = 1/G$$

so that

$$\varepsilon_1 = (1/E)[\sigma_1 - \nu(\sigma_2 + \sigma_3)]$$

(and similarly for the 2 and 3 components),

$$\varepsilon_4 = \sigma_4/G$$

(and similarly for the 5 and 6 components).

The stiffness is

$$C'_{11} = C'_{22} = C'_{33} = E(1-\nu)/(1+\nu)(1-2\nu)$$
$$C'_{12} = C'_{23} = C'_{31} = E\nu/(1+\nu)(1-2\nu)$$
$$C'_{44} = C'_{55} = C'_{66} = G$$

so that

$$\sigma_1 = [E(1-\nu)/(1+\nu)(1-2\nu)]\{\varepsilon_1 + [\nu/(1-\nu)](\varepsilon_2 + \varepsilon_3)\}$$

(and similarly for the 2 and 3 components),

$$\sigma_4 = G\varepsilon_4$$

(and similarly for the 5 and 6 components).

4. Transversely Isotropic Solid (Hexagonal Symmetry)

A transversely isotropic solid has one axis of symmetry (taken here as the 3-axis). Examples are hexagonal crystals, unidirectionally reinforced composites, and materials with a layered structure, such as plywood and rolled metal sheets, the properties in these cases are isotropic in the plane, but differ through the thickness.

There are five independent elastic constants—two Young's moduli (E and E'), two Poisson's ratios (ν and ν'), and one shear modulus G'.

The compliance has nine components:

$$S'_{ab} = \begin{pmatrix} 1/E & -\nu/E & -\nu'/E' & 0 & 0 & 0 \\ -\nu/E & 1/E & -\nu'/E' & 0 & 0 & 0 \\ -\nu'/E' & -\nu'/E' & 1/E' & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/G' & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/G' & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(1+\nu)/E \end{pmatrix}$$

so that

$$\varepsilon_1 = (1/E)[\sigma_1 - \nu\sigma_2 - (\nu' E/E')\sigma_3]$$
$$\varepsilon_2 = (1/E)[\sigma_2 - \nu\sigma_1 - (\nu' E/E')\sigma_3]$$
$$\varepsilon_3 = (1/E')[\sigma_3 - \nu'(\sigma_2 - \sigma_1)]$$
$$\varepsilon_4 = (1/G')\sigma_4$$
$$\varepsilon_5 = (1/G')\sigma_5$$
$$\varepsilon_6 = [2(1+\nu)/E']\sigma_6$$

and the stiffness is

$$C'_{ab} =$$
$$\begin{pmatrix} (1-\nu'^2 E/E')E/\Delta & (\nu+\nu'^2 E/E')E/\Delta & (1+\nu)\nu' E/\Delta & 0 & 0 & 0 \\ (\nu+\nu'^2 E/E')E/\Delta & (1-\nu'^2 E/E')E/\Delta & (1+\nu)\nu' E/\Delta & 0 & 0 & 0 \\ (1+\nu)\nu' E/\Delta & (1+\nu)\nu' E/\Delta & (1-\nu^2)E/\Delta & 0 & 0 & 0 \\ 0 & 0 & 0 & G' & 0 & 0 \\ 0 & 0 & 0 & 0 & G' & 0 \\ 0 & 0 & 0 & 0 & 0 & E/2(1+\nu) \end{pmatrix}$$

where $\Delta = (1+\nu)(1-\nu-2\nu'^2 E/E')$.

Note: $C'_{66} = E/2(1+\nu) = (C'_{11} - C'_{12})/2$.

Also, the engineering moduli are determined by the stiffnesses as follows:

$$E = (C_{11} - C_{12})[2(C_{11} + C_{12}) - C_{33}]/(C_{11} + C_{12})$$
$$E' = C_{13}^2(C_{11} - C_{13})[2(C_{11} + C_{12}) - C_{33}]/(C_{11} + C_{12})^2(C_{33} - C_{11} - C_{12})$$
$$\nu = (C_{11} + C_{12} - C_{33})/(C_{11} + C_{12})$$
$$\nu' = C_{13}/(C_{11} + C_{12})$$
$$G = (C_{11} - C_{12})/2.$$

Then

$$\sigma_1 = [(1-\nu'^2 E/E')E/\Delta][\varepsilon_1 + \nu(1+\nu'^2 E/\nu E')/(1-\nu'^2 E/E')\varepsilon_2$$
$$+\nu'(1+\nu)/(1-\nu'^2 E/E')\varepsilon_3] \text{ etc.}$$

Limitations

Energy considerations show that Poisson's ratio is restricted to the range

$$-1 < \nu < 2\nu'^2 E/E'.$$

Negative values are admissible, and not uncommon. The upper bound is no longer 1/2 as for isotropic solids.

5. Orthotropic Solid (Rhombohedral Symmetry)

An orthotropic solid has three orthogonal axes of symmetry. There are nine independent constants—three Young's moduli (E_1, E_2, E_3), three Poisson's ratios (ν_{12}, ν_{23}, ν_{31}), and three shear moduli (G_4, G_5, G_6).

Note that Poisson's ratio is not symmetric in the indices. Because the compliance is symmetric, it follows that

$$\nu_{ij}/E_j = \nu_{ji}/E_i \text{(indices not summed)}.$$

Caution: Some texts use a different notation with indices switched: $\mu_{ji} = \nu_{ij}$, so that

$$\mu_{ij}/E_i = \mu_{ji}/E_j.$$

The compliance is

$$S'_{ab} = \begin{pmatrix} 1/E_1 & -\nu_{12}/E_2 & -\nu_{13}/E_3 & 0 & 0 & 0 \\ -\nu_{21}/E_1 & 1/E_2 & -\nu_{23}/E_3 & 0 & 0 & 0 \\ -\nu_{31}/E_1 & -\nu_{32}/E_2 & 1/E_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/G_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/G_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/G_6 \end{pmatrix}$$

and the stiffness is

$$C'_{ab} =$$

$$\begin{pmatrix} (1-\nu_{23}\nu_{32})E_1/\Delta & (\nu_{12}+\nu_{13}\nu_{32})E_1/\Delta & (\nu_{13}+\nu_{12}\nu_{23})E_1/\Delta & 0 & 0 & 0 \\ (\nu_{21}+\nu_{23}\nu_{31})E_2/\Delta & (1-\nu_{13}\nu_{31})E_2/\Delta & (\nu_{23}+\nu_{21}\nu_{13})E_2/\Delta & 0 & 0 & 0 \\ (\nu_{31}+\nu_{21}\nu_{32})E_3/\Delta & (\nu_{32}+\nu_{12}\nu_{31})E_3/\Delta & (1-\nu_{12}\nu_{21})E_3/\Delta & 0 & 0 & 0 \\ 0 & 0 & 0 & G_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & G_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & G_6 \end{pmatrix}$$

where $\Delta = (1 - \nu_{12}\nu_{21} - \nu_{23}\nu_{32} - \nu_{31}\nu_{13} - 2\nu_{12}\nu_{23}\nu_{31})$.

The stiffnesses can be written directly in terms of the compliances using the relationships given above.

Limitations

Energy considerations show that Poisson's ratios are restricted as follows:

$$|\nu_{ij}| < (E_j/E_i)^{1/2} \text{ and } \Delta > 0.$$

APPENDIX 3

STATES OF STRESS OR STRAIN IN WAVES

A. STATES OF UNIAXIAL STRESS OR STRAIN

A wave carries a state of stress and strain which is uniform across a plane, or more generally a surface—the wave surface. There is no variation of the stress, strain, displacement, or velocity across the wave surface. These vary only along the normal to that surface, as illustrated in Fig. 142.

The displacement can be at some angle to the surface normal. When parallel to the normal, it gives rise to a normal strain, and when tangential, to a shear strain.

In a wave which extends over a wide surface in an unbounded medium, there is no lateral strain because that would imply an ever-increasing lateral displacement. This would generate ever-increasing lateral inertia which would inhibit the wave formation, a concept known as inertial confinement.

In this case, there can be only a normal strain and/or shear strains having one axis along the normal and the other transverse to the wavefront, i.e., ε_{nn}, ε_{nt1}, and ε_{nt2}, where n and t_1 or t_2 designate the normal and two perpendicular tangential directions respectively. These cases are called uniaxial strain states.

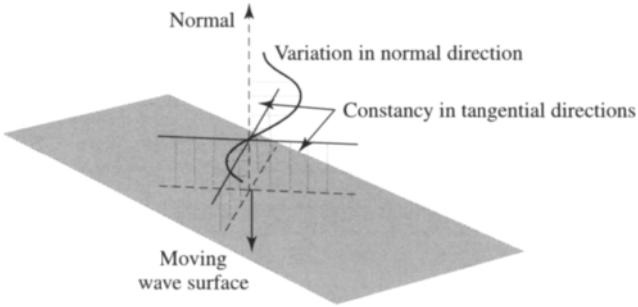


FIGURE 142 Uniform state over a wave surface

In a medium with free lateral surfaces, such as a rod, the free surface would not support lateral or shear stress. There can then be only a normal stress σ_n. This case is called a uniaxial stress state.

Uniaxial stress or strain are states particularly relevant to waves. In these cases, there is only one normal component of stress or of strain, the other two being zero. When stress or strain is uniaxial, the other, strain or stress, generally cannot be uniaxial, and is triaxial.

A1. Uniaxial Normal Strain: The Longitudinal Wave

The only strain in a wave with motion only along the normal, i.e., when the displacement has only a normal component, $u_n(n)$, where n denotes the normal direction and distance along the normal, is a normal strain $\varepsilon_{nn} = \partial u_n/\partial x_n$. All other components of displacement and strain are zero. The normal stress is $\sigma_{nn} = \mathrm{C}_{nnnn}\varepsilon_{nn}$ (n, t_1, or s are not summed indices), and there are lateral stresses $\sigma_{tt} = (\mathrm{C}_{ttnn}/\mathrm{C}_{nnnn})\sigma_{nn}$, and $\sigma_{ss} = (\mathrm{C}_{ssnn}/\mathrm{C}_{nnnn})\sigma_{nn}$, where s denotes a second tangential direction normal to t. For isotropy, these lateral stresses are equal since $(\mathrm{C}_{ttnn}/\mathrm{C}_{nnnn}) = (\mathrm{C}_{ssnn}/\mathrm{C}_{nnnn}) = \nu/(1-\nu)$. There are no shear stresses or strains in the wave coordinates, but there are in other directions (according to Mohr's circle, Appendix 1) as described later.

In an isotropic material this is the most common form of wave. It propagates in any direction. In an anisotropic material it only propagates along a material axis.

A typical configuration for this case is a thick slab of large area, with forces applied normally to the faces, but not to the edges, as illustrated in Fig. 143. The strain is uniaxial, because in-plane motion would be inhibited by size. The stress is triaxial.

The state of stress at the wavefront is triaxial with a normal component σ_n acting along the propagation direction, i.e., the wavefront normal, and, because of the Poisson effect, transverse components $\sigma_t = [\nu/(1-\nu)]\sigma_n$ in any two perpendicular directions (see formulas for stiffness and stresses in an isotropic material under one strain, in Appendix 2). There is no shear in the n–t plane,

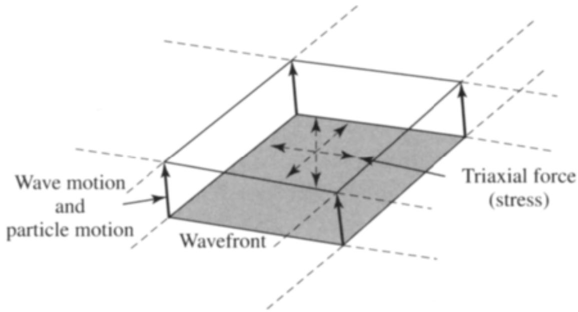


FIGURE 143 Uniaxial normal strain motions in a longitudinal wave

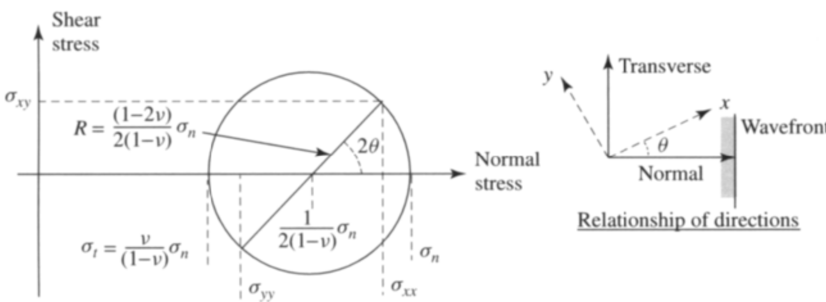


FIGURE 144 Mohr's circle for stresses in the coordinate directions derived from those at the wavefront

so these are the principal stresses. The components of stress in some other x–y coordinate directions, where the x-axis is at an angle α to the propagation direction, can then be found from Mohr's circle, as illustrated in Fig. 144.

Writing $m = \cos\theta$, so that $\cos 2\theta = 2m^2 - 1$, $\sin 2\theta = 2m(1 - m^2)^{1/2}$, the stresses are

$$\begin{aligned}
\sigma_{xx} &= [\sigma_n/2(1-\nu)][1 + (1-2\nu)(2m^2 - 1)] \\
&= [\sigma_n/(1-\nu)][(1-2\nu)m^2 + \nu] \\
\sigma_{yy} &= [\sigma_n/2(1-\nu)][1 - (1-2\nu)(2m^2 - 1)] \\
&= [\sigma_n/(1-\nu)][(1-\nu) - (1-2\nu)m^2] \\
\sigma_{xy} &= [\sigma_n/(1-\nu)](1-2\nu)m(1-m^2)^{1/2}.
\end{aligned}$$

A2. Uniaxial Transverse Shear Strain

In a transverse shear wave, the only displacement component $u_t(n)$ is in the plane of the wave and transverse to the wave normal. The variation of

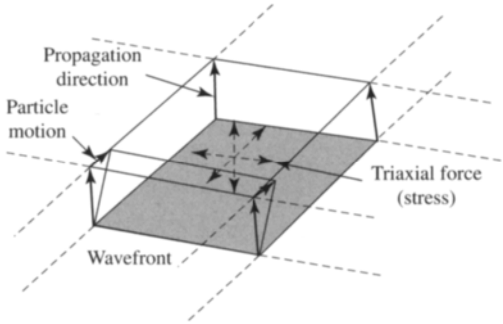


FIGURE 145 Motions in an in-plane transverse shear wave

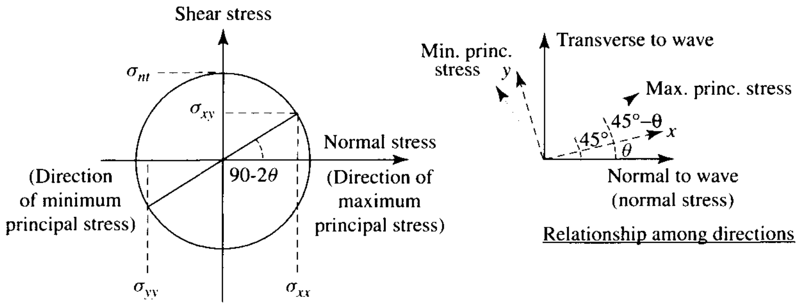


FIGURE 146 Mohr's circle for the stresses in a transverse wave

displacement is along the normal, and hence the shear strain lies in a plane containing the normal and a tangent as illustrated in Fig. 145.

Only a shear strain $\varepsilon_{nt} = \partial u_t/\partial x_n/2$ is created, where t is a direction parallel to the surface. There is again only one component of strain, and this can be called a transverse uniaxial state. The transverse stress is $\sigma_{nt} = C_{ntnt}\varepsilon_{nt}$ (n and t not summed), and there may be stresses in other directions, depending on the nature of the elasticity. Note that there is also a component of rotation $\omega_z = -\partial u_t/\partial x_n/2$.

The stresses for other directions can be determined from the appropriate Mohr's circle shown in Fig. 146. Here the direction of the maximum principal stress is at 45° to the maximum shear stress and thus to the wave normal. Hence an arbitrary direction x at an angle θ to the normal is at $45° - \theta$ to the maximum principal stress, as illustrated in the figure. The stresses are

$$\sigma_{xx} = \sigma_{nt}\sin 2\theta = \sigma_{nt}2p(1 - p^2)^{1/2}$$

$$\sigma_{yy} = -\sigma_{nt}\sin 2\theta = -\sigma_{nt}2p(1 - p^2)^{1/2}$$

$$\sigma_{xy} = \sigma_{nt}\cos 2\theta = \sigma_{nt}(2p^2 - 1)$$

where $p = \cos\theta$.

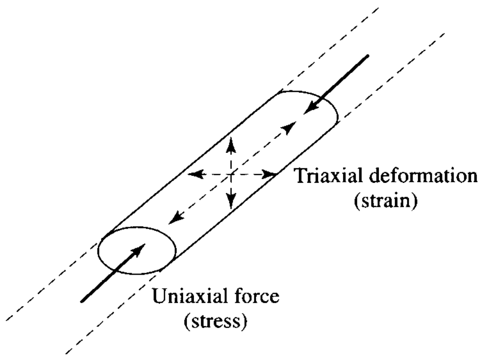


FIGURE 147 Illustration of uniaxial normal stress

A3. Uniaxial Normal Stress

In the case of a wave with only normal stress, the displacement may have components in all directions. The stress and displacement will vary only along the normal. Depending on the type of elasticity, all components of strain can be present.

An example of uniaxial normal stress is a thin rod with forces applied normally to the ends but not to the lateral surface, as illustrated in Fig. 147. The stress is uniaxial since there is no transverse stress. The strain is triaxial, with components along the rod and transverse to it, because of Poisson's ratio effects.

A4. Uniaxial Transverse Stress

In a wave with only transverse stress, the displacement may again have components in all directions. The stress and displacement will vary only along the normal. Depending on the elasticity, there can be all components of strain, because of Poisson's ratio effects. In an isotropic material, however, there will only be stresses in the plane of shear.

An example of uniaxial transverse stress is a thin rod with tangential forces (i.e., a torque) applied to the ends but not to the lateral surface, as illustrated in Fig. 148.

B. UNIAXIAL STRESS–STRAIN RELATIONSHIPS

The stress–strain relationship for these classes of elasticity reduce to the simple one-dimensional form

$$\sigma_{ax} = C\varepsilon_{ax} = C\frac{\partial u_{ax}}{\partial x}$$

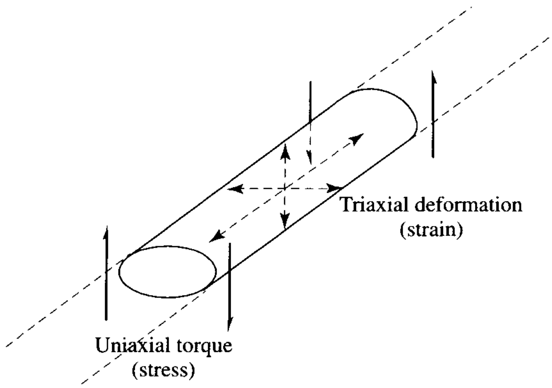


FIGURE 148 Illustration of uniaxial transverse stress

where the subscript ax refers to axial, C is an appropriate combination of elastic constants as presented later, and u is the displacement, whether normal or tangential to the wave surface. Other components of stress or strain may be present, but they are not of concern to the formation of waves.

This is the first fundamental equation for one-dimensional elastic wave propagation (as discussed in Appendix 5).

The coefficient C in this formula is given in the following tables for various material classes and propagation directions. For off-axis directions in materials other than isotropic there is no simple relationship, and displacements may lie in any direction.

(a) Normal Uniaxial Stress (Inverse of Compliance) C_σ

Material class	*1-Direction*	*2-Direction*	*3-Direction*
	$(\sigma_2 = \sigma_3 = 0)$	$(\sigma_2 = \sigma_3 = 0)$	$(\sigma_1 = \sigma_2 = 0)$
Isotropy	E	E	E
Cubic	E	E	E
Transverse isotropy (re 3-axis)	E	E	E'
Orthotropy	E_1	E_2	E_3

(b) Normal Uniaxial Strain (Stiffness) C_ε

Material class	*1-Direction*	*2-Direction*	*3-Direction*
	$(\varepsilon_2 = \varepsilon_3 = 0)$	$(\varepsilon_1 = \varepsilon_3 = 0)$	$(\varepsilon_1 = \varepsilon_2 = 0)$
Isotropy	$(1-\nu)E/(1+\nu)(1-2\nu)$	$(1-\nu)E/(1+\nu)(1-2\nu)$	$(1-\nu)E/(1+\nu)(1-2\nu)$
Cubic	$(1-\nu)E/(1+\nu)(1-2\nu)$	$(1-\nu)E/(1+\nu)(1-2\nu)$	$(1-\nu)E/(1+\nu)(1-2\nu)$
Transverse isotropy (re 3-axis)	$(1-\nu'^2E/E')E/\Delta_T$	$(1-\nu'^2E/E')E/\Delta_T$	$(1-\nu^2)E/\Delta_T$
Orthotropy	$(1-\nu_{23}\nu_{32})E_1/\Delta_O$	$(1-\nu_{31}\nu_{13})E_2/\Delta_O$	$(1-\nu_{12}\nu_{21})E_3/\Delta_O$

where

$$\Delta_T = (1+\nu)(1-\nu-2\nu'^2 E/E')$$

$$\Delta_O = (1-\nu_{12}\nu_{21}-\nu_{23}\nu_{32}-\nu_{31}\nu_{13}-2\nu_{12}\nu_{23}\nu_{31}).$$

Note that for isotropy, the coefficient C_ε for uniaxial strain can be written in several ways using the alternate elastic moduli as follows:

$$C_\varepsilon = E(1-\nu)/(1+\nu)(1-2\nu) = K+4G/3 = \lambda + 2\,\mu = 3(1-\nu)K/(1+\nu).$$

This is the longitudinal wave stiffness defined in Appendix 2.

(c) Transverse Uniaxial Stress (Inverse of Compliance) C_σ

Material class	*4-Direction*	*5-Direction*	*6-Direction*
	$(\sigma_5=\sigma_6=0)$	$(\sigma_6=\sigma_4=0)$	$(\sigma_4=\sigma_5=0)$
Isotropy	$E/2(1+\nu)$	$E/2(1+\nu)$	$E/2(1+\nu)$
Cubic	G	G	G
Transverse isotropy (re 3-axis)	G'	G'	$E/2(1+\nu)$
Orthotropy	G_4	G_5	G_6

(d) Transverse Uniaxial Strain (Stiffness) C_ε

Material class	*4-Direction*	*5-Direction*	*6-Direction*
	$(\varepsilon_5=\varepsilon_6=0)$	$(\varepsilon_6=\varepsilon_4=0)$	$(\varepsilon_4=\varepsilon_5=0)$
Isotropy	$E/2(1+\nu)$	$E/2(1+\nu)$	$E/2(1+\nu)$
Cubic	G	G	G
Transverse isotropy	G'	G'	$E/2(1+\nu)$
Orthotropy	G_4	G_5	G_6

Specific configurations of common concern in engineering are slender rods or beams, and thin plates. The stress–strain states for these configurations are influenced by their geometry as discussed in Appendices 11 and 12.

This Page Intentionally Left Blank

APPENDIX 4

BALANCE OF FORCES AND NEWTON'S LAW OF INERTIA

In 1687, Sir Isaac Newton published his laws of motion, one of which stated that a body at rest remains at rest, and one in motion remains in that motion, unless acted on by a force. The corollary to this is that force produces a change in motion, i.e., acceleration, and this is inversely proportional to the mass:

$$a = F/M, \quad \text{or } F = Ma.$$

Acceleration is the change in velocity:

$$a = dv/dt.$$

The units of force and mass are evidently not the same. Representing force units by [F], mass units by [M], length units by [L], and time units by [T], Newton's law requires that

$$[F] = [M][L]/[T]^2.$$

Absolute forces are defined as those which accelerate a unit mass by a unit acceleration. In common units, these forces are

$$\begin{aligned} 1 \text{ poundal} &= 1 \text{ lb}_M \cdot 1 \text{ fps}^2 \\ 1 \text{ dyne} &= 1 \text{ g} \cdot 1 \text{ cm/sec}^2 \\ 1 \text{ newton} &= 1 \text{ kg} \cdot 1 \text{ m/sec}^2 \end{aligned}$$

It is common to express force in the units of weight (equivalent to mass, e.g., pounds or grams), which are the gravitational forces exerted on unit mass, so that

$$1\ \mathrm{lb_F} = 1\ \mathrm{lb_M} \cdot g = 32.174\ \text{poundals}$$

$$1\ \mathrm{g_F} = 1\ \mathrm{g} \cdot g = 981\ \text{dynes}$$

$$1\ \mathrm{kg_F} = 1\ \mathrm{kg} \cdot g = 9.81\ \text{newtons}$$

where $g = 32.174\ \mathrm{fps}^2 = 980.665\ \mathrm{cm/sec}^2 = 9.8066\ \mathrm{km/sec}^2$ is the acceleration due to gravity.

A. ONE-DIMENSIONAL SYSTEM

Consider a cylindrical body (a rod) under a force which varies along the length, as illustrated in Fig. 149.

The net force in one direction, say along the x-axis, on a small piece of the rod of length dx is the differential between the force, $F(x)$, at one end and that, $F(x + dx)$, at the other:

$$\delta F = F(x + dx) - F(x) = (dF/dx)\delta x = A(d\sigma/dx)\delta x$$

with $\sigma = F/A$ where A is the area of cross-section. The mass of the piece is

$$\delta M = \rho A dx,$$

where ρ is the density, so that the acceleration is

$$dF/dM = \partial v/\partial t = (1/\rho)\partial\sigma/\partial x.$$

This is the second basic equation of one-dimensional wave propagation (as discussed in Appendix 5).

B. THREE-DIMENSIONAL SYSTEMS

In a 3-D system, normal and shear stresses contribute to the acceleration force, as illustrated in Fig. 150.

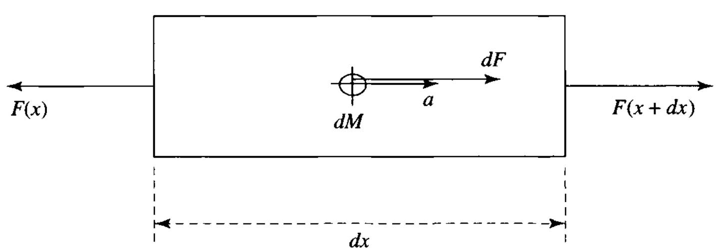


FIGURE 149 Increment of force and acceleration of mass increment

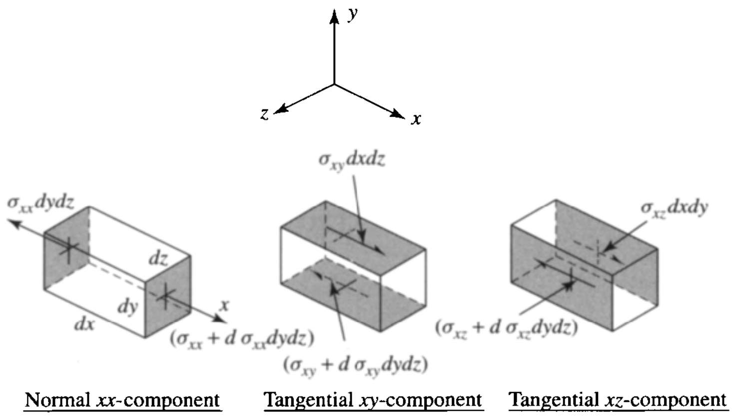


FIGURE 150 Illustration of force components in three dimensions

In the x-direction, for example, the x-variation of the normal stress component leads to a force increment $\delta_x \sigma_{xx} dy\,dz = (\partial\sigma_{xx}/\partial x)dx\,dy\,dz$.The variation of the x–y shear component leads to $\delta_y \sigma_{xy} dx\,dy = (\partial\sigma_{xy}/\partial y)dx\,dy\,dz$, and the x–z component leads to $\delta_z \sigma_{xz} dx\,dy = (\partial\sigma_{xz}/\partial z)dx\,dy\,dz$. Variations across the faces do not contribute as they are the same on both opposing faces. The net force increment is then

$$\delta F_x = (\partial\sigma_{xx}/\partial x + \partial\sigma_{xy}/\partial y + \partial\sigma_{xz}/\partial z)dx\,dy\,dz.$$

For all three directions, this can be written as

$$\delta F_i = (\partial\sigma_{ij}/\partial x_j)dV$$

where the indices i and j take values 1, 2, and 3, representing x, y, and z, and the repeated indices imply a sum over all three values.

The inertia force is

$$\delta F_i = \rho a_i dV,$$

so that the inertia equation can be generalized for triaxial stress to read

$$\partial\sigma_{ij}/\partial x_j = \rho dv_i/dt.$$

APPENDIX 5

THEORY OF WAVE PROPAGATION

The interaction between the forces of elastic deformation and of inertia is responsible for waves in solids and fluids. The equations which describe this interaction in terms of stress and velocity, and their consequences, are described below. There are two ways to analyze these equations:

- Eliminate one of the variables (stress or velocity), leaving a partial differential equation in the other
- Combine them into equations for which constant solutions can be found, called the characteristic equations

Specific solutions to the equations are described in Appendix 6.

Detailed presentations of the theory of elastic waves are given by Love (1944) and by Aki and Richards (1980). A simplified presentation of some aspects is given by Doyle (1989).

A. ONE-DIMENSIONAL SYSTEMS

In their simplest one-dimensional form, the elastic and inertial forces are represented by the following pair of equations (see Appendices 3 and 4):

Elasticity:

$$\sigma = C\partial u/\partial x$$

or, after differentiation with respect to time:

$$\partial\sigma/\partial t = C\partial v/\partial x,$$

and inertia:

$$\partial\sigma/\partial x = \rho\partial v/\partial t$$

where C depends on the material and on the direction chosen for the x-axis, and where $v = \partial u/\partial t$ is the particle velocity.

For any point in time and position, the two variables, stress and velocity, form a state vector (σ, v) which defines the state of the material.

This one-dimensional formulation only applies when uniaxial strain or stress states arise, i.e., for propagation in any direction in isotropic materials, or for propagation along a principal axis of an anisotropic material. The general case of propagation in any direction in anisotropic materials is discussed under three-dimensional waves in the next section.

A1. The Differential Equation

Differentiating the one-dimensional elasticity equation with respect to time, t, and the one-dimensional inertia equation with respect to position, x, leads to two equivalent expressions for the cross derivative of stress:

$$\partial^2\sigma/\partial x\,\partial t = C\partial^2 v/\partial x^2 = \rho\partial^2 v/\partial t^2$$

so that

$$\partial^2 v/\partial x^2 = (1/c^2)\partial^2 v/\partial t^2$$

where $c = (C/\rho)^{1/2}$.

This is a hyperbolic wave equation for velocity. [*Note*: A hyperbolic equation is one for which the solution cannot be extended from values known along certain curves (or surfaces) by using the second derivatives to determine the first derivatives.] Solutions for this differential equation are presented in Appendix 6.

By differentiating in the opposite sequence, two expressions for the cross derivative of velocity are obtained, leading to an identical equation for stress, so that it, too, has the same type of solutions.

Wavespeeds

The wavespeed $c = (C/\rho)^{1/2}$ depends, through the stiffness C, on the nature of the elasticity, the direction of propagation, and the nature (stress or strain state) of the wave. Formulas for C are given in Appendix 3.

A2. The Method of Characteristics

By adding to the temporal derivative of stress (from the one-dimensional elasticity equation) a multiple, α, of that for velocity (from the one-dimensional inertia equation), the following relationship is found:

$$\partial\sigma/\partial t + \alpha\rho\,\partial v/\partial t = \alpha\partial\sigma/\partial x + \alpha C\partial v/\partial x.$$

By taking $\alpha\rho = C/\alpha$, i.e., $\alpha = \pm(C/\rho)^{1/2} = \pm c$, where c is the wavespeed, so that $\alpha\rho = C/\alpha = \pm\rho c = \pm z$, where z is called the impedance, the equation can be written as

$$\partial(\sigma \pm zv)/\partial x \pm (1/c)\partial(\sigma \pm zv)/\partial t = 0.$$

This shows that the quantities $\sigma \pm zv$, called the Riemann invariants, $R_\pm$, satisfy the relationship

$$\partial R_\pm/\partial x \pm (1/c)\partial R_\pm/\partial t = 0.$$

Now, an increment of a Riemann invariant with increments of time and position is

$$\delta R_\pm = (\partial R_\pm/\partial x)\delta x + (\partial R_\pm/\partial t)\delta t$$

and because of the above relationships,

$$\delta R_\pm = (\partial R_\pm/\partial x)(\delta x \pm \mathrm{c}\,\delta t).$$

Hence, $\delta R_\pm = \delta\sigma \pm z\delta v = 0$ when $\delta x \pm c\delta t = 0$. The invariants are therefore constant, hence the name, as the wavefront moves along lines such that $x \pm ct =$ Const, called the characteristics.

The relationship also shows that any change in stress along a characteristic is accompanied by a change in velocity:

$$\delta\sigma = \pm z\delta v = 0 \text{ along } x \pm ct = \text{Const}.$$

These equations represent the characteristic jump equations in the state variables.

A3. *x–t* Wave Diagrams

Diagrams can be constructed in an x–t space, as shown in Fig. 151, to represent the propagation of one-dimensional waves, and to provide a convenient method for analyzing them.

A force applied to a point, x_0, at time zero, excites waves which propagate in two directions with increasing time—one toward decreasing x and one toward increasing x. These waves introduce changes in the state (stress and velocity) according to the characteristic jump equations.

Such x–t diagrams are used to analyze wave patterns in Figs. 47 and 52 of Section 4.

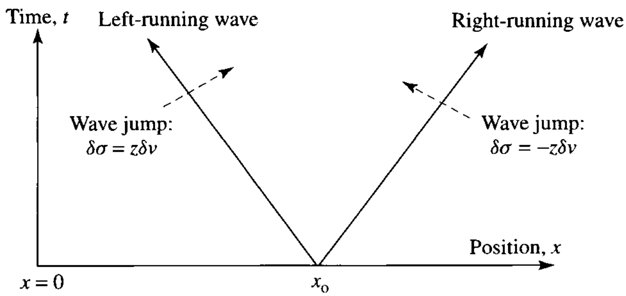


FIGURE 151 x–t diagram showing wave propagation and invariants

B. THREE-DIMENSIONAL SYSTEMS

The elasticity and inertia equations which govern three-dimensional waves are

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl},$$

which can be differentiated with respect to time:

$$\partial\sigma_{ij}/\partial t = C_{ijkl}\partial v_k/\partial x_l,$$

and

$$\rho\partial v_i/\partial t = \partial\sigma_{ij}/\partial x_j.$$

Here the state variables are the stress tensor and the velocity vector (σ_{ij}, v_i) forming a nine-parameter state vector (not rigorously a vector according to the coordinate transformation rule).

B1. The Differential Equation

By differentiating the elasticity equation with respect to position x_j, and the inertia equation with respect to time t, two forms of the cross derivative of stress are found which can be equated:

$$\partial^2\sigma_{ij}/\partial x_j\partial t = C_{ijkl}\partial v_k/\partial x_j\partial x_l = \rho\partial^2 v_i/\partial t^2.$$

This is a hyperbolic wave equation for velocity. A similar equation can be derived for the stresses.

B2. The Method of Characteristics

As in the one-dimensional case above, the method seeks functions, R, of the state variables σ_{ij} and v_i, i.e., $R(\sigma_{ij}, v_k)$, which are invariant on planes

$$\xi = \text{const}$$

where $\xi = n_i x_i \pm ct$, and n_i is the unit normal, $n_i n_i = 1$, to the plane. Then for any of the variables,

$$\partial(\ )/\partial x_i = [\partial(\ )/\partial\xi](\partial\xi/\partial x_i) = n_i[\partial(\ )/\partial\xi]$$
$$\partial(\ )/\partial t = [\partial(\ )/\partial\xi](\partial\xi/\partial t) = \pm c[\partial(\ )/\partial\xi]$$

so that

$$\partial(\ )/\partial x_i = \pm(1/c)n_i\partial(\ )/\partial t.$$

The governing differential equations can then be written as

$$\partial\sigma_{ij}/\partial t = \pm(1/c)n_l C_{ijkl}\partial v_k/\partial t$$
$$\rho\partial v_i/\partial t = \pm(1/c)n_j\partial\sigma_{ij}/\partial t$$

so that

$$n_l n_j C_{ijkl}\partial v_k/\partial t = \rho c^2\partial v_i/\partial t$$

and

$$z\delta v_i \pm n_j\delta\sigma_{ij} = 0$$

where $z = \rho c$ is the acoustic impedance as in the one-dimensional case. Evidently, the invariants are

$$R_i = \rho c v_i \pm n_j\sigma_{ij}.$$

The first equation is

$$[n_l n_j C_{ijkl} - \rho c^2]\partial v_k/\partial t = 0$$

and for there to be a nontrivial solution (i.e., for nonzero acceleration, $\partial v_k/\partial t \neq 0$), the determinant of the bracketed expression must be zero:

$$|\Gamma_{ik} - \rho c^2\delta_{ik}| = 0$$

where $\Gamma_{ik} = n_j n_l C_{ijkl}$ is the Christoffel stiffness in the direction of the normal n_i. These stiffnesses are symmetric: $\Gamma_{ik} = \Gamma_{ki}$.

This is the same equation as was given previously by the differential equation method. A numerical procedure for solving it is described in Appendix 9.

C. WAVE EXCITATIONS

UT waves can be transmitted into an object by a transducer excited by the piezoelectric response, or induced directly into the object by irradiation with a laser or electromagnetic waves, as discussed in Section 5.

C1. Piezoelectric Excitation

The mechanical stress, σ, created by an electric field, E, is

$$\sigma = eE$$

where e is the piezoelectric coefficient of the material. This stress is added to the mechanical stress $\sigma = K\varepsilon$, caused by deformation strain $\varepsilon = \partial u/\partial x$, where K is an elastic constant and u is the deformation:

$$\sigma = K\varepsilon + eE.$$

In conjunction with Newton's Law of Motion,

$$\partial\sigma/\partial x = \rho\partial^2 u/\partial t^2,$$

this leads to the nonhomogeneous wave equation (i.e., an equation having a forcing function)

$$K\partial^2 u/\partial x^2 - \rho\partial^2 u/\partial t^2 = e\partial E/\partial x.$$

The term on the right-hand side of this equation, the gradient of the electric field, is thus a source for waves.

C2. Laser Excitation

Incident light from a laser beam or other radiation, including microwaves and X-rays, is absorbed by the atoms in the surface layer of an object. The depth of this layer for a laser beam is a few microns, whereas it can be a centimeter or more for other, longer-wavelength radiation. The absorbed energy is thermalized (converted to atomic motion) in times of the order of 10^{-13} sec, essentially instantaneously. This leaves a distribution of energy which decays roughly exponentially with depth. For the shallow absorption depth of the laser light, the resulting energy density can be very high. According to the theory of solid-state physics this creates a pressure

$$p = \rho\Gamma e,$$

where ρ is the density, Γ is Gruneisen's coefficient (which is typically around 2 dynes·cm/erg for metals but less for polymers), and e is the energy density. The energy is equivalent to a temperature rise $\Delta T = e/c_v$, where c_v is the specific heat at constant volume.

Alternatively, according to the theory of thermoelasticity, the temperature rise would cause a volumetric strain $\Delta V/V = 3\alpha\Delta T = 3\alpha e/c_v$, where α is the coefficient of linear thermal expansion and 3α is the coefficient of volumetric expansion. Since expansion is constrained by inertia and only arises through wave propagation, a confining pressure is created at short times:

$$p = K\Delta V/V = (3\alpha K/c_v)e.$$

The two theories are consistent provided

$$\Gamma = 3\alpha K/\rho c_v.$$

The intensity of deposited energy can be high enough to produce permanent material damage, even vaporization, so that care must be taken in selecting the laser parameters.

The induced pressure is additive to the elastic stress

$$\sigma_{ij} = C_{ijkl}\partial u_k/\partial x_l - p\delta_{ij}$$

(using a negative sign since pressure is compressive). The differential equation then becomes

$$C_{ijkl}\partial^2 u_k/\partial x_j \partial x_l - \rho\partial^2 v_i/\partial t^2 = \rho\Gamma(\partial e/\partial t)\delta_{ij},$$

so that the rate of energy deposition $\partial e/\partial t$ is a source of elastic waves. Evidently high-power, short-pulsed lasers are required.

This Page Intentionally Left Blank

APPENDIX 6

SOLUTIONS TO THE WAVE EQUATIONS

The equations for elastic wave propagation are derived in Appendix 5. Solutions to these equations can be found as fundamental waves which propagate in unbounded regions independent of boundary conditions, or as boundary-dependent waves. Methods for evaluating these solutions are given here.

A. FUNDAMENTAL SOLUTION TYPES

A1. Spherical Waves: Unbounded Three-Dimensional Space with Point Source

This configuration (hard to realize in practice, but a useful concept) can be considered to arise from a force applied to an internal point in an isotropic body where the boundaries are remote. The solution is a spherical wave which decays with the inverse of radius, as illustrated in Fig. 152 and having the mathematical form

$$u(r, t) = r^{-1}U(\xi) \text{ where } \xi = (r \pm ct)$$

where r is the radius from the origin of the wave, and $U(\xi)$ is a function which describes the wave profile at some point and time. For example, at initial time

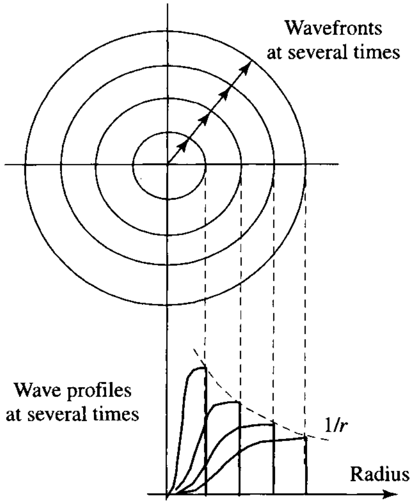


FIGURE 152 Fundamental spherical wave

$t = 0$, $U(\xi) = U(r)$, where $U(r)$ is the initial distribution of the wave, i.e., the initial condition for the solution.

A solution with the negative sign propagates outward with decreasing amplitude, and one with the positive sign propagates inward toward the origin with increasing amplitude.

This solution can be verified by direct differentiation as follows:

$$\partial u/\partial x_i = -r^{-2}(x_i/r)U + r^{-1}(dU/d\xi)(x_i/r)$$

where $r = (x_i x_i)^{1/2}$, so that $\partial r/\partial x_i = (1/2)2x_i/r$, and $\partial x_i/\partial x_i = 3$ in three dimensions.

After algebraic manipulation wherein many terms cancel,

$$\partial^2 u/\partial x_i \partial x_i = r^{-1} d^2 U/d\xi^2.$$

Similarly

$$\partial^2 u/\partial t^2 = c^2(1/r)d^2 U/d\xi^2.$$

These two terms satisfy the three-dimensional wave differential equation

$$\partial^2 u/\partial x_i \partial x_i = (1/c^2)\partial^2 u/\partial t^2.$$

Solutions for a distributed source within a body can be constructed by superposition: The distributed source is defined as an incremental function of volume, $S(r)dV$. This may represent a displacement, a velocity, or a stress. The solution for the entire distribution is then given as the sum of the contributions from all increments of the source:

$$u(r, t) = \int [S(r')/(r - r')]dV$$

where r is a distance to the solution point, and r' is the radius to a source point.

This is called Green's integral, and the kernel, $f(r, r') = 1/(r - r')$, is called Green's function.

A2. Cylindrical Waves: Unbounded Three-Dimensional Space with Line Source (or Two-Dimensional Space with Point Source)

This configuration is a line source in three dimensions and results in cylindrical waves with a logarithmic decay:

$$u(r', t) = \log(1/r')U(\xi) \text{ where } \xi = (r' \pm ct),$$

with r' being the two-dimensional radius, $r' = (x^2 + y^2)^{1/2}$. This solution can be verified by direct differentiation as follows:

$$\partial u/\partial x_i = r'^{-1}(x_i/r')U + \log(r'^{-1})(dU/d\xi)(x_i/r')$$

so that

$$\partial^2 u/\partial x_i \partial x_i = \log(r'^{-1})d^2U/d\xi^2$$

and

$$\partial^2 u/\partial t^2 = c^2 \log(1/r')d^2U/d\xi^2.$$

Two-dimensional indices are used here, so that $i = 1, 2$. Then as before, $r' = (x_i x_i)^{1/2}$, and $\partial r'/\partial x_i = (1/2)2x_i/r'$, but $\partial x_i/\partial x_i = 2$ in two dimensions. These two terms satisfy the differential equation.

A3. Step Waves: One-Dimensional Waves

A form of solution is the Heaviside step function, $H(t)$, where $H = 0$ for $t \leq 0$, and $H = 1$ for $t > 0$, as illustrated in Fig. 153.

The derivative of the step function is the Dirac delta function, also shown in Fig. 153, which is zero for all t but infinite for $t = 0$. Thus

$$dH(t)/dt = \delta(t).$$

A solution with two step waves propagating in opposite directions is then

$$v(x, t) = H(x \pm ct),$$

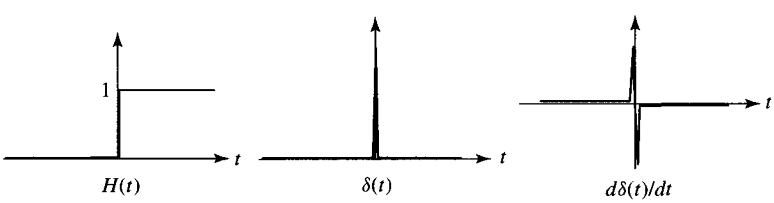


FIGURE 153 The Heaviside step function $H(t)$ and its first and second derivatives, the Dirac delta function, $\delta(t)$, and $d\delta(t)/dt$

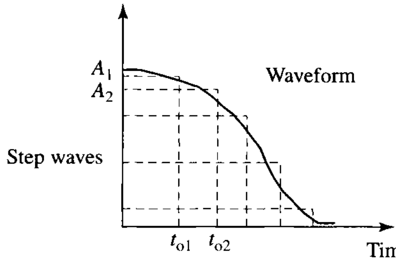


FIGURE 154 Waveform developed by superposition of step waves

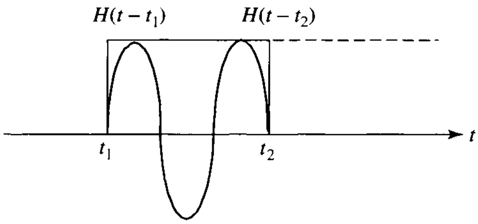


FIGURE 155 Illustration of tone burst

so that $\partial^2 v/\partial x^2 = \delta'(x \pm t)$ and $\partial^2 v/\partial t^2 = c^2\delta'(x \pm t)$. These satisfy the one-dimensional wave equation $\partial^2 v/\partial x^2 - (1/c^2)\partial^2 v/\partial t^2 = 0$.

Combinations (superpositions) of several such step waves with different initiating times can represent an arbitrary waveform as illustrated in Fig. 154:

$$v(x,t) = \int A_1(t_{o1})H(x+ct)dt_{o1} + \int A_2(t_{o2})H(x-ct)dt_{o2},$$

where A_1 and A_2 are functions of time and are defined from boundary and initial conditions.

Most waves are visualized as step waves with an additional structure, such as the tone burst illustrated in Fig. 155. This comprises an oscillatory, periodic motion which starts and stops abruptly and has the form

$$u = U_1[H(t-t_1) - H(t-t_2)]\sin\omega(t-t_1).$$

A4. Nonpropagating Solutions

There are also elementary nonpropagating linear solutions

$$u = A + Bx + Ct + Dxt$$

which can be used to satisfy certain boundary and initial conditions.

B. SEPARATION OF VARIABLES

B1. One-Dimensional Waves

Consider the one-dimensional wave equation for velocity $v(x, t)$:

$$\partial^2 v/\partial x^2 - (1/c^2)\partial^2 v/\partial t^2 = 0.$$

Solutions can be found by the method of separating variables. Assume that the dependencies of v on x and t are not coupled and can be represented by two independent functions:

$$v(x, t) = X(x)T(t).$$

The differential equation then becomes

$$d^2X/dx^2 T = (1/c^2)X d^2T/dt^2.$$

After division by XT, assuming neither is zero, the two terms are each dependent on only one distinct independent variable, x or t. These terms can then vary independently, and so each must be a constant, say k^2 (taken as positive):

$$(d^2X/dx^2)/X = (d^2T/dt^2)/c^2T = k^2.$$

There are therefore two separate ordinary differential equations for the two functions:

$$d^2X/dx^2 - k^2X = 0$$

$$d^2T/dt^2 - c^2k^2T = 0.$$

These have solutions

$$X = X_1 \sin kx + X_2 \cos kx$$

$$T = T_1 \sin \omega t + T_2 \cos \omega t,$$

with $\omega = kc$, and $X_1(k)$, $X_2(k)$, $T_1(k)$, and $T_2(k)$ are coefficients which can be functions of k or ω. The full solution, the product XT, has the form of sine and cosine waves running in each direction and summed over all possible frequencies:

$$v(x, t) = \int [A_1(\omega) \sin k(x - ct) + A_2(\omega) \cos k(x - ct)$$

$$+A_3(\omega) \sin k(x + ct) + A_4(\omega) \cos k(x + ct)]d\omega$$

where A_1–A_4 are functions of frequency which are determined by the boundary and excitation conditions, k is called the wave number, and $kc = \omega$, the coefficient of time, t, is the circular frequency.

C. PLANE WAVES IN TWO OR THREE DIMENSIONS

In two or three dimensions, a plane wave propagating with constant amplitude can be described by the phase function:

$$\phi = k_i x_i \pm \omega t = \kappa(n_i x_i \pm ct)$$

where $k_i = \kappa n_i$ is the propagation vector, $\kappa = (k_i k_i)^{1/2}$ is the wave number, and $c = \omega/\kappa$ is the wavespeed. Then the solution for a typical wave equation can be written as

$$u_i(x_j, t) = u_i(\phi),$$

because

$$\partial^2 u/\partial x_i \partial x_i = \kappa^2 \partial^2 u/\partial \phi^2 \text{ and } \partial^2 u/\partial t^2 = \omega^2 \partial^2 u/\partial \phi^2,$$

which satisfy the differential equation

$$\partial^2 u/\partial x_i \partial x_i = (1/c)^2 \partial^2 u/\partial t^2.$$

The equation for three-dimensional elastic waves in an anisotropic material (derived in Appendix 5) is

$$C_{ijkl} \partial v_k/\partial x_j \partial x_l = \rho \partial^2 v_i/\partial t^2.$$

Solutions can be sought in the form

$$v_i = v_i(\phi)$$

when the differential equation becomes

$$(C_{ijkl} n_j n_l - \rho \omega^2 \delta_{ik}) \partial^2 v_k/\partial \phi^2 = 0.$$

This requires that, for a nonzero velocity v_k,

$$|\Gamma_{ik} - \rho c^2 \delta_{ik}| = 0,$$

which is the result obtained by the Method of Characteristics in Appendix 5, where $\Gamma_{ik} = C_{ijkl} n_j n_l$ are the Christoffel stiffnesses, and $c = \omega/k$ is the wavespeed, with $k = (k_i k_i)^{1/2}$. This equation generally requires numerical solution, as described in Appendix 9.

C1. The Velocity Potentials for Isotropy

For isotropy, the equation reduces to the following:

$$(\lambda + 2\mu) \partial^2 v_j/\partial x_j \partial x_i + \mu \, \partial^2 v_i/\partial x_j \partial x_j = \rho \partial^2 v_i/\partial t^2.$$

According to the Helmholtz vector decomposition theorem (Wills, 1958), a vector such as the velocity vector can be represented by the derivatives of a scalar potential ϕ and a vector potential A_i with $A_{i,i} = 0$:

$$v_i = \partial \phi/\partial x_i + \varepsilon_{ijk} \partial A_j/\partial x_k$$

where ε_{ijk} is the alternating symbol, having values of 1 if the three indices take a cyclic sequence of values from 1 to 3 (i.e., 1, 2, 3; 2, 3, 1; or 3, 1, 2), -1 if an anticyclic sequence (i.e., 3, 2, 1; 2, 1,3; or 1, 3, 2), and 0 otherwise (i.e., if any two have the same value).

Note that $\partial v_i/\partial x_i$ (called the divergence of v_i) is $\partial^2 \phi/\partial x_i \partial x_i$ and that $\varepsilon_{ilm} \partial v_l/\partial x_m$ (called the curl of v_i) is $\varepsilon_{ijk} \varepsilon_{ilm} v_{l,m} = A_{j,k} - A_{k,j}$. When applied to the displacement vector, the divergence is the volume change represented by ϕ, and the curl is the rotation represented by A_i.

The Helmholz decomposition leads to the equation

$$\partial[(\lambda + 2\,\mu)\partial^2\phi/\partial x_j \partial x_j - \rho\partial^2\phi/\partial t^2]/\partial x_i$$
$$+\varepsilon_{ijk}\partial[\mu\ \partial^2 A_j/\partial x_l \partial x_l - \rho\partial A_j/\partial t^2]/\partial x_k = 0.$$

The parts of this equation for the two potentials can each be taken to be zero separately:

$$\partial^2\phi/\partial x_j \partial x_j - (1/c_l^2)\partial^2\phi/\partial t^2 = 0$$
$$\partial^2 A_j/\partial x_l \partial x_l - (1/c_s^2)\partial^2 A_j/\partial t^2 = 0$$

where

$$c_l^2 = (\lambda + 2\,\mu)/\rho = [(1-\nu)/(1+\nu)(1-2\nu)](E/\rho)$$
$$c_s^2 = \mu\,/\rho = [1/2(1+\nu)](E/\rho).$$

The scalar potential equation is identified with longitudinal waves because the motion is parallel to the propagation direction. The vector potential is associated with transverse waves, because the motion is perpendicular to the propagation.

Each of these two separate three-dimensional wave equations for the potential functions has solutions like those discussed earlier.

Note that the potential method is not useful when applied to the general anisotropy case, because the terms in the vector potential do not separate into simple wave equations.

The general case of three-dimensional waves, particularly in anisotropic media, requires numerical solution. Methods such as the transient dynamic finite-element or finite-difference procedures have been developed for this.

D. PLANE WAVE SOLUTIONS

The method of potentials described earlier is used for the two-dimensional case. There are two two-dimensional cases of interest, referred to as in-plane motion and transverse motion.

D1. In-Plane Motion

Consider a plane wave propagating along a normal, n, at an angle α to the x-axis in the x–y plane. Its plane lies along the z-axis as illustrated in Fig. 156. All motion is taken to lie in the x–y plane, and there is no variation in the z-direction. In a longitudinal wave the motion is along the normal, and in a transverse wave it is along the tangent.

There can be a scalar potential $\phi(x, y, t)$ for a longitudinal wave, and one component of the vector potential, $A_z(x, y, t)$ for a transverse wave. Note that

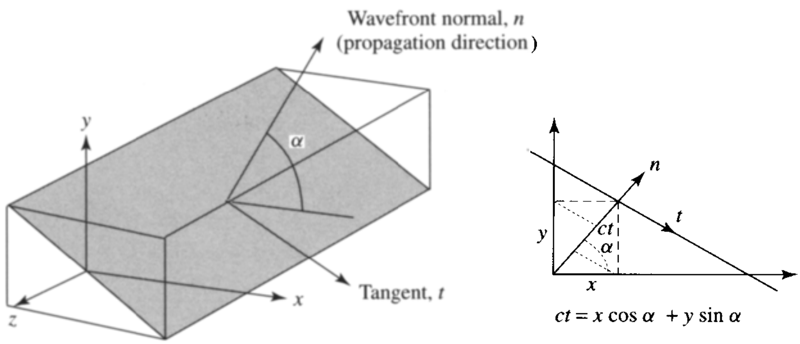


FIGURE 156 Coordinate system for plane waves

components A_x and A_y of the vector potential would give displacements in the z-direction and these are transverse motions as discussed later. The presence of a z-component of the potential does not imply motion or variation in the z-direction.

The differential equation requires that these potentials be of the form of waves:

$$\phi(x, y, t) = \phi(\xi) \text{ with } \xi = mx + ny - c_l t$$

and

$$A_z(x, y, t) = A_z(\eta), \text{ with } \eta = px + qy - c_s t$$

where (m, n) and (p, q) define the angles of the direction of propagation of the waves, with $m = \cos\alpha_l$, α_l being the angle between the propagation direction and the x-axis for the longitudinal wave, and $p = \cos\alpha_s$, α_s being the angle for transverse shear waves.

To satisfy the differential equations,

$$m^2 + n^2 = 1, \text{ so that } n = (1 - m^2)^{1/2}$$
$$p^2 + q^2 = 1, \text{ so that } q = (1 - p^2)^{1/2}.$$

The properties of each wave type are shown in the following table:

	Dilatational	Transverse
Displacements		
u_x	$\partial\phi/\partial x = m\, d\phi/d\xi$	$\partial A_z/\partial y = q\, dA_z/d\eta$
u_y	$\partial\phi/\partial y = n\, d\phi/d\xi$	$-\partial A_z/\partial x = -p\, dA_z/d\eta$
u_z	0	0

	Dilatational	Transverse
Velocities		
v_x	$\partial^2\phi/\partial x\partial t = mc_l d^2\phi/d\xi^2$	$\partial^2 A_z/\partial y\partial t = qc_s d^2 A_z/d\eta^2$
v_y	$\partial^2\phi/\partial y\partial t = nc_l d^2\phi/d\xi^2$	$-\partial^2 A_z/\partial x\partial t = -pc_s d^2 A_z/d\eta^2$
v_z	0	0
Strains		
$\epsilon_{xx} = \partial u_x/\partial x$	$m^2 d^2\phi/d\xi^2$	$pqd^2 A_z/d\eta^2$
$\epsilon_{yy} = \partial u_y/\partial y$	$n^2 d^2\phi/d\xi^2$	$-pqd^2 A_z/d\eta^2$
$\epsilon_{xy} = (\partial u_x/\partial y + \partial u_y/\partial x)/2$	$2mnd^2\phi/d\xi^2$	$(q^2 - p^2)d^2 A_z/d\eta^2$
Stresses (from Appendix 2)		
$\sigma_{xx} = C_l\{\varepsilon_{xx} + [\nu/(1-\nu)]\varepsilon_{yy}\}$	$C_l\{m^2 + [\nu/(1-\nu)]n^2\}d^2\phi/d\xi^2 = [C_l/(1-\nu)][(1-2\nu)m^2 + \nu]d^2\phi/d\xi^2$	$[C_l(1-2\nu)/2(1-\nu)]2pqd^2 A_z/d\eta^2$
$\sigma_{yy} = C_l\{\varepsilon_{yy} + [\nu/(1-\nu)]\varepsilon_{xx}\}$	$C_l\{n^2 + [\nu/(1-\nu)]m^2\}d^2\phi/d\xi^2 = [C_l/(1-\nu)][(1-\nu) - (1-2\nu)m^2]d^2\phi/d\xi^2$	$-[C_l(1-2\nu)/2(1-\nu)]2pq \times d^2 A_z/d\eta^2$
$\sigma_{xy} = G\varepsilon_{xy}$	$C_l[(1-2\nu)/2(1-\nu)]2mnd^2\phi/d\xi^2$	$[C_l(1-2\nu)/2(1-\nu)](q^2 - p^2) \times d^2 A_z/d\eta^2$

Note that $C_l = E(1-\nu)/(1+\nu)(1-2\nu)$, and $G = E/2(1+\nu) = C_l[(1-2\nu)/2(1-\nu)]$ are elastic constants appropriate to the stress–strain conditions.

These results are the same as given for stress states in Appendix 3, provided that

$$d^2\phi/d\xi^2 = \sigma_n/C_l$$

$$d^2 A_z/d\eta^2 = [2(1-\nu)/(1-2\nu)]\sigma_{nt}/C_l = \sigma_{nt}/G.$$

Boundary conditions on excitation at the surface determine the values of σ_n and σ_{nt} and hence of $d^2\phi/d\xi^2$ and $d^2 A_z/d\eta^2$.

It can be shown that

For a longitudinal wave, $\rho c_l v_x = \sigma_{xx}m + \sigma_{xy}n$, and $\rho c_l v_y = \sigma_{xy}m + \sigma_{yy}n$

For a transverse wave, $\rho c_s v_x = \sigma_{xx}p + \sigma_{xy}q$, and $\rho c_s v_y = \sigma_{xy}p + \sigma_{yy}q$

which are the characteristic or impedance equations $zv_i = \sigma_{ij}m_j$ as developed in the Method of Characteristics (Part C of Appendix 5).

D2. Transverse Motion

Consider the same wave configuration as in Fig. 156, but with motions in the z-direction, and thus transverse to the x–y plane. There can be two transverse waves described by two wave functions, $A_x(x, y, t)$ and $A_y(x, y, t)$, but there is no dilatational wave. The wave functions satisfy the wave equation and can be taken in the forms

$$A_x(x, y, t) = A_x(\eta),\ A_y(x, y, t) = A_y(\eta), \text{ with } \eta = px + qy - c_s t.$$

There is one component of displacement and velocity in each wave, denoted by the second index, x or y:

$$\begin{aligned} u_{z|x} &= \partial A_x/\partial y = p\,dA_x/d\eta \\ u_{z|y} &= -\partial A_y/\partial x = -q\,dA_x/d\xi, \\ v_{z|x} &= \partial^2 A_x/\partial y \partial t = p\,d^2 A_x/d\eta^2 \\ v_{z|y} &= -\partial^2 A_y/\partial x \partial t = -q\,d^2 A_y/d\eta^2 \end{aligned}$$

where the symbols $|x$ and $|y$ designate the relevant potential. There are two strains from the combined potentials:

$$\begin{aligned} \epsilon_{xz} &= \partial u_z/\partial x = p^2 d^2 A_x/d\eta^2 - pq d^2 A_y/d\eta^2 \\ \epsilon_{yz} &= \partial u_z/\partial_y = pq\, d^2 A_x/d\eta^2 - q^2\, d^2 A_y/d\eta^2 \end{aligned}$$

and two stresses:

$$\begin{aligned} \sigma_{xz} &= G(p^2\, d^2 A_x/d\eta^2 - pq\, d^2 A_y/d\eta^2) \\ \sigma_{yz} &= G(pq\, d^2 A_x/d\eta^2 - q^2\, d^2 A_y/d\eta^2) \end{aligned}$$

Boundary conditions prescribed on surface forces will determine which of these waves, or what combination of them, is excited. The wave deriving from A_x has a stress resultant in the x–z plane, while that from A_y lies in the y–z plane.

$$\begin{aligned} \sigma_{zn} &= \sigma_{xz} n_x + \sigma_{yz} n_y = p\sigma_{xz} + q\sigma_{yz} \\ &= G[pd^2 A_x/d\eta^2 - qd^2 A_y/d\eta^2] = \rho c_s v_{z|x} \\ \sigma_{zt} &= \sigma_{xz} n_y - \sigma_{yz} n_x = q\sigma_{xz} - p\sigma_{yz} \\ &= G[qd^2 A_x/d\eta^2 - pd^2 A_y/d\eta^2] = \rho c_s v_{z|y}. \end{aligned}$$

APPENDIX 7

DISPERSION, GROUP VELOCITY

The speed of a wave propagating in an unbounded uniform elastic medium is constant—it depends on elasticity and density, both of which are constant for the small range of stresses used in UT. In nonuniform materials, of course, this may not be true. Wood and some man-made materials are nonuniform, as discussed in Section 5. For nonelastic (e.g., viscoelastic) media, for high-amplitude waves, or for waves near the boundaries of a body, the wavespeed may depend on frequency and other parameters such as dimensions. Waves whose speed depend on frequency are called dispersive. Increasing frequency may decrease the wavespeed (as in rod waves), or increase it (as in bending) (see Appendices 11 and 12).

Consider the phase of a dispersive wave, in which the wavespeed depends on frequency:

$$\phi(\omega) = kx - \omega t = k(x - ct)$$

with $\omega = kc$, and $c = c(\omega)$ is the phase velocity, or wavespeed.

In a wave which has components with a range of frequencies, those components which have the same phase will superimpose and interfere constructively to form a wave of increased amplitude. At a fixed position and time, the phase of different components will be the same, called a stationary phase, when

$$d\phi/d\omega = (dk/d\omega)x - t = 0.$$

This represents a wave on which $x - c_g t = 0$, where $c_g = d\omega/dk$ is called the group velocity. Then, from $\omega = kc$,

$$c_g = d\omega/dk = c + k(dc/d\omega)(d\omega/dk) = c + k(dc/d\omega)c_g$$

so that

$$c_g = c/[1 - (\omega/c)(dc/d\omega)].$$

When the wavespeed is independent of frequency, i.e., when $dc/d\omega = 0$, the group velocity is the same as the wavespeed. If the wavespeed decreases with frequency so that $dc/d\omega < 0$, then the group velocity is less than the wavespeed, and conversely.

A. WAVE PROPAGATION IN VISCOELASTIC MEDIA

This is a large subject which is introduced only briefly here. In a viscoelastic material, forces arise in response to both deformation, i.e., strain ε, and to rate of deformation, i.e., strain rate, $\partial\varepsilon/\partial t$. Thus the elasticity relationship for linear elasticity, $\sigma = \sigma(\varepsilon) = C\varepsilon$, is replaced in the linear viscoelastic case by the constitutive relationship

$$\sigma = \sigma(\varepsilon, \partial\varepsilon/\partial t) = C\varepsilon + D\partial\varepsilon/\partial t$$

where C is an elasticity coefficient, and D is a viscosity coefficient. In a simple one-dimensional wave propagation configuration, there are then the two equations

Constitutive relationship: $\sigma = C\partial u/\partial x + D\partial^2 u/\partial x\partial t$

Newton's Law: $\rho\partial^2 u/\partial t^2 = \partial\sigma/\partial x.$

Combining these two equations to eliminate stress gives

$$\rho\partial^2 u/\partial t^2 = C\partial^2 u/\partial x^2 + D\partial^3 u/\partial x^2\partial t.$$

For a harmonic wave of the type

$$u = Ue^{i(k_i x_i - \omega t)},$$

this gives the complex frequency–wave number relationship

$$\rho\omega^2 = Ck_i k_i + iDk_i k_i\omega$$

and the wavespeed is given by

$$c^2 = \omega^2/k_i k_i = c_o^2(1 + iD/C\omega)$$

where $c_o = (C/\rho)^{1/2}$.

The wavespeed has a frequency dependence which leads to group behavior. At very high frequency this is the same as the elastic case, $c \to c_o$, and is infinite at low frequency.

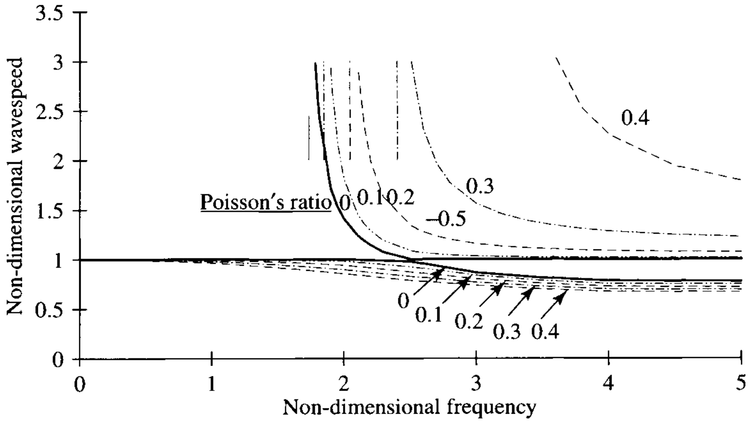


FIGURE 157 Approximate nondimensional dispersion curves for a circular rod

B. WAVE PROPAGATION IN A THICK ROD

The Mindlin–Herrmann Approximation (Doyle, 1989)

The stress–strain state and the balance of elastic and inertia forces which act on a rod when lateral motion are considered is discussed in Appendix 11. These forces combine to produce propagating waves whose speed depends on the frequency, the diameter of the rod, and its elasticity.

As developed in an approximate method in Appendix 11, the analysis leads to two types of wave. The first of the two modes is essentially the bar wave of simple theory, but it slows with increasing frequency. In an exact analysis, it can be shown that these waves are asymptotic at high frequency to Rayleigh waves on the free surface of the rod. The second mode is a high-frequency mode, nonpropagating at low frequency, whose speed decreases asymptotically toward the bar velocity as frequency increases.

A graph of the dependence of β, the nondimensional wavespeed, on γ, the nondimensional frequency, representing the approximate nondimensional dispersion curves for a circular rod, is shown for a range of Poisson's ratios in Fig. 157. Here $\beta = c/c_0$ and $\gamma = a\omega/c_0$. This is 2π times the ratio of the time for a wave to cross the radius, a, i.e., $2\pi a/c_0$, to the time period of the wave $t_p = 2\pi/\omega$.

The corresponding group velocities are shown in Fig. 158. The group velocity of the bar mode is lower than the wavespeed, but is the same at very low frequency. The group velocity for the second mode starts at zero at the cutoff frequency when the wavespeed is infinite, and increases with frequency.

For zero Poisson's ratio, the group velocities for both modes have a discontinuity at a frequency above the cutoff frequency, and their continuations

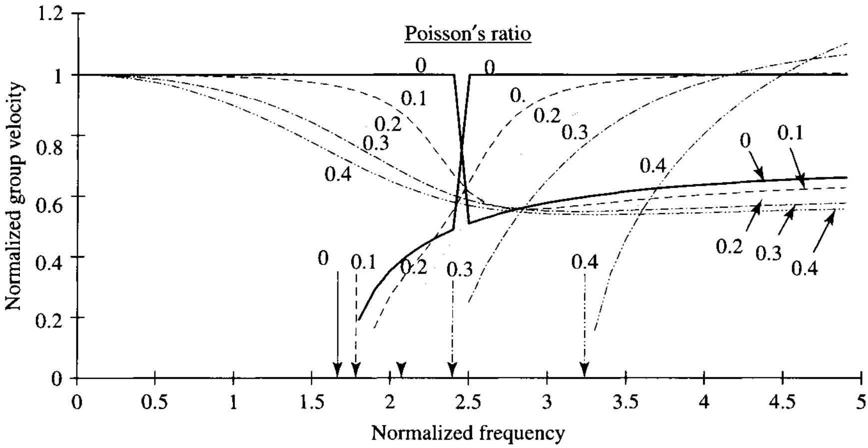


FIGURE 158 Approximate normalized group velocity for a circular rod

switch with each other. At other Poisson's ratios, the curves exhibit a gradual transition.

C. BENDING WAVES

Bending is dispersive at all frequencies. The simplest theory of bending, the Engineers' Bending Theory (EBT), applies to low frequencies, but at high frequencies, the more advanced Timoshenko theory is more satisfactory. In the EBT, the shear force is indeterminate because rotation of the cross-section is taken to be the same as the slope of the deformation. This implies that a normal cross-section remains normal, but in the Timoshenko theory the cross-section is allowed to rotate.

The elasticity relationships and inertia forces are discussed in Appendix 12. The EBT leads to the dispersion equation

$$EI\kappa^4 - \mu\omega^2 = 0,$$

so that the wavespeed is

$$c(\omega) = \omega/k = (EI/\mu)^{1/4}\omega^{1/2}.$$

In nondimensional form this is

$$\beta = \gamma^{1/2}$$

where $\beta = c/c_0$ is the nondimensional wavespeed, with $c_0 = (E/\rho)^{1/2}$, and $\gamma = \omega k/c_0$ is the nondimensional frequency.

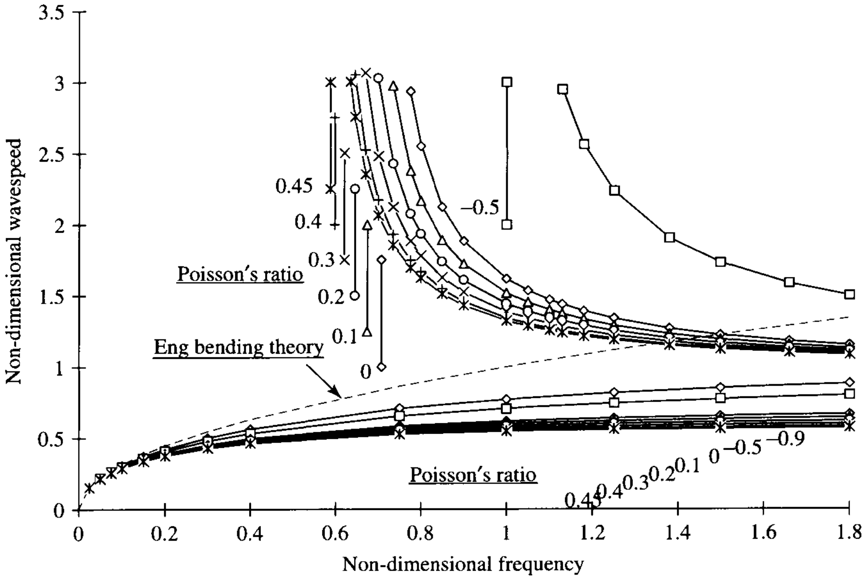


FIGURE 159 Approximate dispersion relationships for bending of a Timoshenko beam

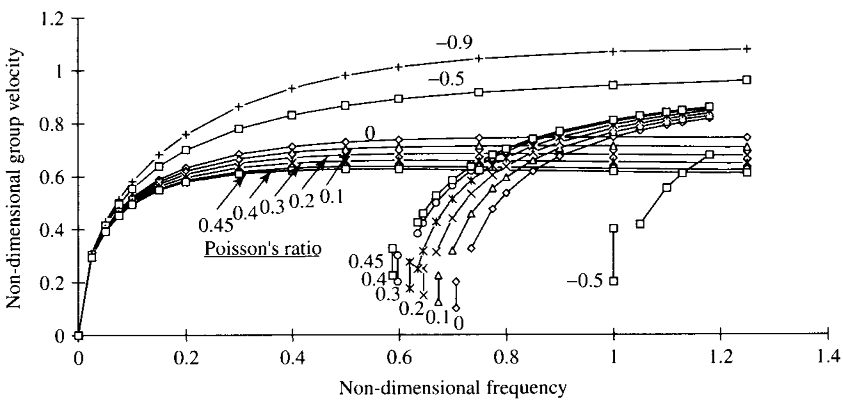


FIGURE 160 Approximate normalized group velocity for a timoshenko Beam

This is a dispersive behavior. The Timoshenko theory leads to a complicated dispersion relationship developed in Appendix 12, giving two dispersive wavespeeds, which are plotted as functions of Poisson's ratio in Fig. 159. The corresponding group velocities are shown in Fig. 160.

This Page Intentionally Left Blank

APPENDIX 8

TRANSDUCER BEAM FORMING

A transducer is a distributed source, which can be considered to emit spherical waves from all points of its face, with amplitude decreasing as the inverse of the radius. Near the transducer, these waves interfere destructively forming a complex wave pattern of peaks and nulls, called the near field, but at further distances they interfere constructively forming plane waves in a conical region called the far field, which is the UT beam.

A. SIMPLIFIED ANALYSIS OF BEAM FORMING

A simplified analysis is given next which illustrates the essential aspects of beam forming and provides approximate formulae for the near-field distance and the beam cone angle. A more complete yet still approximate analysis is given later.

Consider the interference, at the near-field distance N, between a wave emitted from the center of the transducer and one emitted from the edge. When these two waves meet and are out of phase, they cancel forming a null which is the edge of the beam cone at a point P illustrated in Fig. 161.

The distance to P from the edge of the transducer of diameter $D = 2R$ is N, and from the center O it is $(N^2 + R^2)^{1/2}$. To be out of phase at the near-field

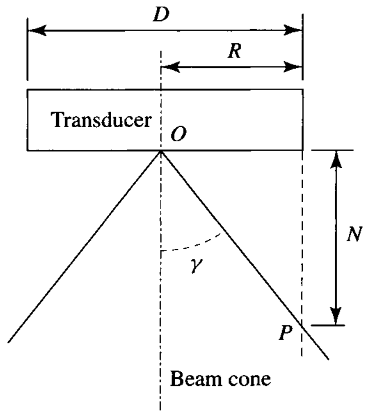


FIGURE 161 Simplified representation of beam cone

distance, N, the difference in the phases $\phi = kx - \omega t$ of the two waves is $\delta\phi = \pi$, where $k = 2\pi/\lambda$ is the wave number, and λ is the wavelength.

Then

$$\delta\phi = k\delta x = k[(N^2 + R^2)^{1/2} - N] = \pi.$$

This requires that

$$N = (R^2 - \lambda^2/4)/\lambda = (D^2 - \lambda^2)/4\lambda \simeq D^2/4\lambda$$

(since the wavelength is usually smaller than the transducer diameter).

The beam cone half-angle is then

$$\gamma = \text{atan}(R/N) \simeq \text{atan}(2\lambda/D).$$

A small wavelength or a large transducer produce a narrow beam. In the limit of zero diameter, the beam covers the whole space as in a spherical wave from a point source.

B. ANALYSIS OF RADIAL VARIATION OVER BEAM

The method followed below is essentially one developed in fluid mechanics for the disturbance propagated in a fluid or gas from a piston. It is also similar to the analysis of wave radiation from an antenna in electrical engineering. The analysis does not consider the directional aspects of stress and velocity in a solid: it considers pressure produced at a point by a wave as independent of the direction of the wave, so that the pressure from various waves can be added arithmetically. In a solid, the components of stress or velocity should be resolved in direction before addition. This approximation is followed here to render the analysis amenable to solution.

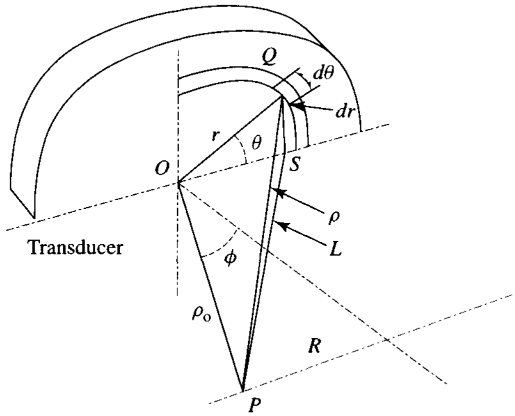


FIGURE 162 Geometry of beam forming

Consider a wave propagating from a small area $rdrd\theta$ of a circular ring at a point Q at (r, θ) on the surface of the transducer, to a far field point P at a distance L from the face and R radially outward, as illustrated in Fig. 162.

The incremental wave amplitude (signal) at P due to a source on an area increment $dA = rdrd\theta$ is

$$dS = \sigma_o(e^{-ik\rho}/k\rho)rdrd\theta$$

where σ_o represents the strength of the wave emitted per unit area by the transducer, and ρ is the distance PQ from source to field point. The time factor $e^{i\omega t}$ is omitted for clarity.

The total signal at Q from all points P on the transducer is then the integral

$$S = \sigma_o \int\int e^{-ik\rho} r\, dr\, d\theta / k\rho$$

where ρ is the radial distance from the incremental source point Q to the far-field point P. The following distances are evident from the geometry configurations:

$$OP = \rho_o = (L^2 + R^2)^{1/2},\ OS = r\cos\theta,\ \text{and}\ SQ = r\sin\theta, \phi = \text{atan}(R/L)$$

so that SP is

$$\begin{aligned} SP &= [OP^2 + OS^2 - 2OPOS\cos(\pi/2 + \phi)]^{1/2} \\ &= [\rho_o^2 + r^2\cos^2\theta + 2\rho_o r\cos\theta\sin\phi]^{1/2}. \end{aligned}$$

The source–point distance, PQ, is then

$$\begin{aligned} \rho &= [SP^2 + SQ^2]^{1/2} \\ &= [\rho_o^2 + r^2 + 2\rho_o r\cos\theta\sin\phi]^{1/2}. \end{aligned}$$

By neglecting the term r^2, this is

$$\rho \simeq \rho_o[1 + 2r\cos\theta\sin\phi/\rho_o]^{1/2}$$

and using the first two terms of a binomial expansion for the square root,

$$\rho \simeq \rho_o + r\cos\theta\sin\phi.$$

Neglecting the term $r\cos\theta\sin\phi$ in ρ in the denominator, the integral becomes

$$S \simeq \sigma_o e^{-ik\rho_o}/k\rho_o \int\int e^{-ikr\cos\theta\sin\phi} r\,dr\,d\theta.$$

The integral over θ is of the form of a standard Bessel function integral:

$$\int_0^{2\pi} e^{-ia\cos\theta}d\theta = 2\pi J_o(a)$$

so that taking $a = kr\sin\phi$ leads to the integral

$$S \simeq 2\pi\sigma_o e^{-ik\rho_o}/k\rho_o \int J_o(kr\sin\phi) r\,dr.$$

This integral over r is of the form of another standard Bessel function integral

$$\int_0^{d/2} J_o(ar) r\,dr = (d/2a) J_1(ad/2).$$

Putting $a = k\sin\phi$ results in

$$S \simeq 2\pi\sigma_o(e^{-ik\rho_o}/k\rho_o)(d/2k\sin\phi)J_1[(kd/2)\sin\phi]$$
$$\simeq 2\, S_o J_1(x)/x$$

where

$S_o = \sigma_o(\pi d^2/4)/(k\rho_o)$ is the nominal total signal at the distance ρ_o

$x = (kd/2)\sin\phi = \pi(d/\lambda)\sin[\text{atan}(R/L)]$

and the factor $e^{-ik\rho_o}$ combines with the suppressed factor $e^{i\omega t}$ to represent the propagating wave. The distribution $J_1(x)/x$ is illustrated in Fig. 163. The distribution shows that there is a sequence of side lobes of low intensity outside the primary lobe, which is the beam. The cone edge of the beam is at $x \simeq 3.8317\ldots$, which is the first zero of the Bessel function and leads to the formula

$$(kd/2)\sin\phi = 3.8317\ldots$$

so that the cone angle for the main lobe is

$$\phi = \text{asin}(3.8317\ldots\lambda/\pi d) = \text{asin}(1.2197\ldots\lambda/d).$$

The first side lobe falls between the first and second zeroes ($x = 3.8317\ldots$ to $7.0156\ldots$) of the Bessel function, for which $\phi = \text{asin}(2.2331\ldots\lambda/d)$.

These two angles are plotted vs d/λ, the ratio of transducer diameter to wavelength, in Fig. 41 in Section 3. The effective transducer diameter, d_{eff}, can

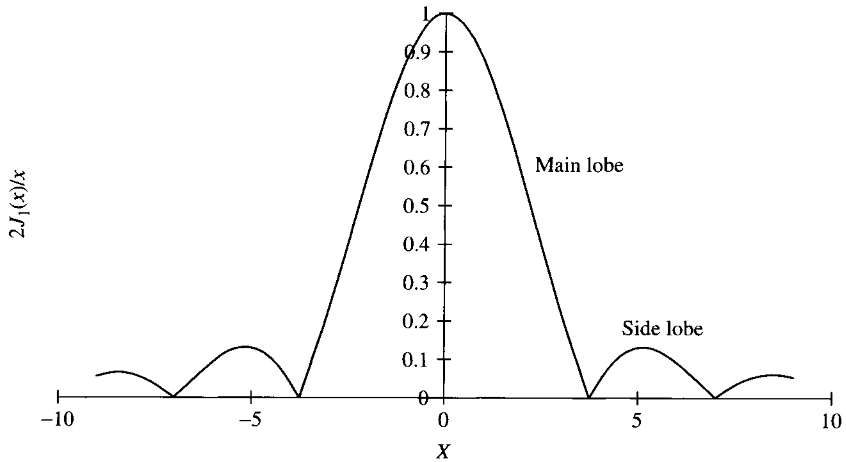


FIGURE 163 Distribution $J_1(x)/x$

be determined by estimating the beam angle and reading off the diameter from the graph. This method does not give accurate results because the beam edge is not easy to identify, but it gives a useful first approximation. A refined result is obtained by fitting a curve to the radial distribution given by the expansion formula

$$S(x)/S_0 = 2J_1(x)/x = 1 - (1/8)x^2 + (1/192)x^4 \ldots .$$

Now x can be approximated by the expression

$$x \sim (\pi d_{\text{eff}}/\lambda)(R/L),$$

so that the radial distribution can be fitted to the formula

$$S(x)/S_0 \sim 1 - (1/8)(\pi d_{\text{eff}}/\lambda)^2(R/L)^2.$$

The parameter x of the beam distribution depends on the two ratios R/L, which is the radial off-axis distance normalized to axial distance from the transducer, and d/λ, which is the transducer diameter normalized to wavelength. The beam distribution is plotted as a function of R/L for several values of d/λ, in Fig. 40 of Section 3.

A larger transducer produces a narrower beam as does a shorter wavelength. Since $\lambda = fc$, where f is frequency and c is wavespeed, a lower frequency produces a narrower beam. Furthermore, the beam profile changes with material, and one with a lower wavespeed produces a narrower beam. A beam changes when it is transmitted from one material to another.

C. ANALYSIS OF AXIAL VARIATION OF BEAM STRENGTH

The approximations made in the foregoing analysis negate the axial variation. An alternative analysis is made here by considering only points on the axis and using less restrictive approximations.

The analysis is based on the geometry illustrated in Fig. 164, which differs from Fig. 162 in that the far-field point P lies on the transducer axis. Here, the distance from the source element at Q to P is

$$\begin{aligned}\rho = QP = (SP^2 + SQ^2)^{1/2} &= (OP^2 + OS^2 + SQ^2)^{1/2} \\ &= [L^2 + r^2 \cos^2\theta + r^2 \sin^2\theta]^{1/2} \\ &= [L^2 + r^2]^{1/2}.\end{aligned}$$

Since there is no angular dependence, the integral over θ leads to a factor of 2π, and the signal is

$$S = 2\pi\sigma_o \int_0^{d/2} e^{-ik\rho} r dr / k\rho.$$

Now, because $\rho^2 = L^2 + r^2$, it follows that

$$r dr = \rho d\rho$$

so that

$$S = 2\pi(\sigma_o/k) \int_L^{\rho_e} e^{-ik\rho} d\rho = 2\pi(\sigma_o/k)[e^{-ik\rho_e} - e^{-ikL}]/(-ik)$$

where $\rho_e = (L^2 + d^2/4)^{1/2}$ is the distance from the edge of the transducer. Then

$$S = (2\pi\sigma_o/k^2)e^{-ikL}[e^{-ik(\rho_e - L)} - 1]/i.$$

Noting that $e^{-ix} = \cos x - i \sin x$, and dismissing the complex result, this gives

$$\begin{aligned}S &= (2\pi\sigma_o/k^2)e^{-ikL} \sin k(\rho_e - L) \\ &= (2\pi\sigma_o/k^2)e^{-ikL} \sin\{(2\pi L/\lambda)[(1 + d^2/4L^2)^{1/2} - 1]\}.\end{aligned}$$

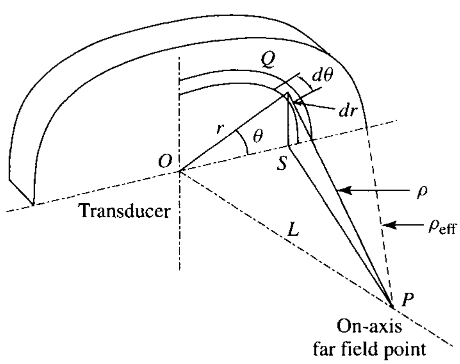


FIGURE 164 Geometry for on-axis far-field point

The exponential factor, e^{-ikL}, combines with the suppressed time factor $e^{i\omega t}$ to represent the propagating wave. The variation of signal along the axis depends on the transducer diameter and the wavelength.

The magnitude of the nondimensional signal can be represented by the nondimensional formula

$$|S|/S_o = 2(\lambda/\pi d)^2|\sin\{[2\pi(L/\lambda)][1 + (d/\lambda)^2/4(L/\lambda)^2]^{1/2} - 1\}|$$

where $S_o = (\pi d^2/4)\sigma_o$ is the nominal total signal. This function is plotted without the factor $2(\lambda/\pi d)^2$ for several values of d/λ in Fig. 42 of Section 3.

The near-field limit is determined by the largest distance $L = N$ at which the signal is 0, i.e., the argument of the sine function is π, so that the signal exhibits no further oscillation in strength at points further out. Then

$$(N^2 + d^2/4)^{1/2} - N = \lambda/2,$$

which leads to the same result as in the approximate analysis:

$$N = (d^2/4 - \lambda^2/4)/\lambda = (d^2 - \lambda^2)/4\lambda \simeq d^2/4\lambda.$$

For distances much larger than the diameter, the two distances ρ_e and L are almost the same, so that the argument of the sine is small, and the sine is approximated by its argument. The signal amplitude is then

$$[\pi^2(d/\lambda)^2]|S|/S_o = 2\pi(L/\lambda)[(1 + d^2/4L^2)^{1/2} - 1].$$

Now, for $L > d$, using the approximation $(1 + d^2/4L^2)^{1/2} \simeq (1 + d^2/8L^2 \ldots)$ allows the signal strength to be written as

$$|S|/S_o \simeq (1/2\pi)(\lambda/L)$$

showing that in the far field, the signal varies inversely with distance, L, as for a point source. From Fig. 42 it appears that this approximation gives results for $|S|$ which are less than S_o provided $L = L_{min} > 4\lambda$ when $d/\lambda = 2$, i.e., when $N/\lambda = 1$, or $>\lambda$ when $d/\lambda = 1$ and $N/\lambda = 1/4$. Hence $L_{min} > 4N$.

This Page Intentionally Left Blank

APPENDIX 9

SOLUTIONS FOR ANISOTROPY

Considerations of a plane wave with unit normal n_i propagating in an anisotropic medium lead to the following equation (Appendix 5):

$$(n_j n_l C_{ijkl} - \rho c^2 \delta_{ik}) \partial v_k / \partial t = 0.$$

This is a principal- (or eigen-) value problem for determining principal directions p_i such that there are nontrivial solutions for acceleration, $\partial v_k / \partial t \neq 0$, which can be written as

$$(\Gamma_{ik} - \rho c^2 \delta_{ik}) p_k = 0$$

where $\Gamma_{ik} = n_j n_l C_{ijkl}$ are the Christoffel stiffnesses in the direction of the normal n_i. These stiffnesses are symmetric: $\Gamma_{ik} = \Gamma_{ki}$.

The determinant of the bracketed expression must be zero, giving the determinantal equation

$$|\Gamma_{ik} - \rho c^2 \delta_{ik}| = 0.$$

In general there are three principal (eigen) values for ρc^2, and three principal (eigen) vectors.

When written out in terms of the stiffnesses for general anisotropy with 21 constants using the six-dimensional notation (see Appendices 1 and 2) the

Christoffel stiffnesses are as follows:

$$\Gamma_{11} = l^2C_{11} + m^2C_{66} + n^2C_{55} + 2mnC_{56} + 2nlC_{15} + 2lmC_{16}$$
$$\Gamma_{22} = l^2C_{66} + m^2C_{22} + n^2C_{44} + 2mnC_{24} + 2nlC_{46} + 2lmC_{26}$$
$$\Gamma_{33} = l^2C_{55} + m^2C_{44} + n^2C_{33} + 2mnC_{34} + 2nlC_{35} + 2lmC_{45}$$
$$\Gamma_{12} = l^2C_{16} + m^2C_{26} + n^2C_{45} + mn(C_{46} + C_{25}) + nl(C_{14} + C_{56}) + lm(C_{12} + C_{66})$$
$$\Gamma_{13} = l^2C_{15} + m^2C_{46} + n^2C_{35} + mn(C_{45} + C_{36}) + nl(C_{13} + C_{55}) + lm(C_{14} + C_{56})$$
$$\Gamma_{23} = l^2C_{56} + m^2C_{24} + n^2C_{34} + mn(C_{44} + C_{23}) + nl(C_{36} + C_{45}) + lm(C_{25} + C_{46})$$

where (l, m, n) are the components of n_i.

For restricted classes of material of symmetry equal to or greater than orthotropic, and with nine distinct elastic constants, i.e, excluding rhombohedral symmetries (those whose axes of symmetry are not orthogonal), these expressions reduce as follows:

$$\Gamma_{11} = l^2C_{11} + m^2C_{66} + n^2C_{55}$$
$$\Gamma_{22} = l^2C_{66} + m^2C_{22} + n^2C_{44}$$
$$\Gamma_{33} = l^2C_{55} + m^2C_{44} + n^2C_{33}$$
$$\Gamma_{12} = lm(C_{12} + C_{66})$$
$$\Gamma_{13} = nl(C_{13} + C_{55})$$
$$\Gamma_{23} = mn(C_{44} + C_{23}).$$

A. WAVESPEED

The determinantal equation is a cubic in (ρc^2) which can be written explicitly as the characteristic equation

$$(\rho c^2)^3 - \Gamma_1(\rho c^2) + \Gamma_2(\rho c^2) - \Gamma_3 = 0$$

where the coefficients Γ_1, Γ_2, and Γ_3 are called the invariants of the matrix Γ_{ij}:

$$\Gamma_1 = \Gamma_{ii} = \Gamma_{11} + \Gamma_{22} + \Gamma_{33}$$
$$\Gamma_2 = (\Gamma_1^2 - \Gamma_{ij}\Gamma_{ij})/2 = \Gamma_{11}\Gamma_{22} + \Gamma_{22}\Gamma_{33} + \Gamma_{33}\Gamma_{11} - \Gamma_{12}^2 - \Gamma_{23}^2 - \Gamma_{31}^2$$
$$\Gamma_3 = |\Gamma_{ij}| = \Gamma_{11}\Gamma_{22}\Gamma_{33} - \Gamma_{11}\Gamma_{23}^2 - \Gamma_{22}\Gamma_{31}^2 - \Gamma_{33}\Gamma_{12}^2 + 2\Gamma_{12}\Gamma_{23}\Gamma_{31}.$$

This equation has three solutions for (ρc^2), showing there are three possible wave types, called the modes, each with a principal vector corresponding to the direction of material motion, v_i, for that mode.

Note that in most systems of units the elastic constants are large numbers, typically of the order of 10^6 in English units (psi), and 10^{11} in cgs units (dynes/cm^2). The terms in the three invariants are then of the order of 10^6, 10^{12}, and 10^{18}, or 10^{11}, 10^{22}, and 10^{33}. These disparate magnitudes cause numerical difficulties which can be ameliorated by dividing all the elastic constants by a suitable number such as the largest of them, and then multiplying the roots, ρc^2, resulting from solving the characteristic equation by that number.

Solutions of the characteristic equation determine the wavespeeds and are called the principal values or eigenvalues of the matrix; they are generally all real for elastic waves. They can be found by numerical methods using an available mathematical program or spreadsheet, or by writing a program.

The simplest method finds the smallest root by direct substitution of values of (ρc^2) into the cubic function, starting from the value $(\Gamma_3/\Gamma_2)/10$. Note that (Γ_3/Γ_2) would approximate the solution if (ρc^2) were small so that the higher order terms are not significant.

Values of (ρc^2) are incremented upward from this initial value until the cubic function changes sign, i.e., passes through zero. The value is then reduced by the last step, the step is reduced by half, and the process repeated. This procedure is continued, progressively reducing the step, until the change in the value of (ρc^2) is of no further significance.

The two other roots are found by algebraic solution of the quadratic

$$(\rho c^2)^2 + A(\rho c^2) + B = 0$$

which results from factorization determined by artificial division, so that

$$y^3 - \Gamma_1 y^2 + \Gamma_2 y - \Gamma_3 = y(y^2 + Ay + B) = 0$$

where

$$y = \rho c^2 - (\rho c^2)_1$$

and

$$A = 3(\rho c^2)_1 - \Gamma_1$$

$$B = 3(\rho c^2)_1^2 - 2\Gamma_1(\rho c^2)_1 + \Gamma_2.$$

The remaining two roots are then solutions of the quadratic equation

$$y^2 + Ay + B = 0$$

and these are

$$y_{2,3} = (\rho c^2)_{2,3} - (\rho c^2)_1 = -[A \pm (A^2 - 4B)^{1/2}]/2.$$

Evidently $A < 0$ or else at least one of the roots would be negative, and this would not be physically reasonable.

The trial-and-error method just described gives the smallest root, $(\rho c^2)_1$, and the positive sign in $(\rho c^2)_2$ makes this the next smallest, with $(\rho c^2)_3$ being the largest root. In calculations starting from one of the axes, say $\theta = 0$, the modes can be identified as quasi-transverse modes, 1 and 2, in this order, followed by the quasi-longitudinal mode, 3. However, the sequence changes when B changes sign at $B = 0$, since then $y = 0$, and $(\rho c^2)_2 = (\rho c^2)_1$. Examples of calculated wavespeeds and polarization angles are shown in Figs. 29 and 30 in Section 3.

B. MATERIAL (PARTICLE) MOTION AND THE PRINCIPAL DIRECTIONS (VECTORS)

Nontrivial solutions for the acceleration vector $\partial v_k/\partial t$ which satisfy the equations

$$(\Gamma_{ik} - \rho c^2 \delta_{ik})\partial v_k/\partial t = 0$$

are determined only to the extent of three orthogonal unit vectors $p_{\alpha i}$ (the principal vectors or eigenvectors of the matrix, one for each mode, designated by $\alpha = 1, 2$, or 3) which give the directions of the accelerations. The magnitude of the acceleration for each mode is determined by the boundary or initial conditions.

The principal vectors are called the polarization vectors. Note that in general, material motion, i.e., polarization, is not parallel to the wave velocity, which for a plane wave, is along the normal to the wavefront.

The directions of the principal vectors are determined by the set of three equations for each root α of the equation:

$$[\Gamma_{ik} - (\rho c^2)_\alpha \delta_{ik})]p_{\alpha k} = 0.$$

They can be found by defining the following parameters (according to Musgrave, 1954):

$$G_1 = \pm(\Gamma_{21}\Gamma_{13}/\Gamma_{23})^{1/2}$$

$$G_2 = \pm(\Gamma_{32}\Gamma_{21}/\Gamma_{31})^{1/2}$$

$$G_3 = \pm(\Gamma_{13}\Gamma_{32}/\Gamma_{12})^{1/2}$$

so that

$$\Gamma_{12} = G_1 G_2$$

$$\Gamma_{23} = G_2 G_3$$

$$\Gamma_{31} = G_3 G_1.$$

Note that this definition cannot be made when any of the three off-diagonal stiffnesses is zero, as discussed in the special cases later. In those cases the principal vectors are found by a simpler procedure.

The signs of the G's must be selected so as to preserve the signs of the Γ's. It is not possible to have one of the Γ_{ij} negative at a time, only none or two, because the elastic constants are positive and the Christoffel stiffnesses depend on the product of two components of the wave vector. If one or two of the components is negative, then two of the stiffnesses are negative, i.e., when

$$\Gamma_{12} < 0 \text{ and } \Gamma_{23} < 0, \text{ then } G_2 < 0, \text{ and } G_1 \text{ and } G_3 > 0.$$

$$\Gamma_{23} < 0 \text{ and } \Gamma_{31} < 0, \text{ then } G_3 < 0, \text{ and } G_1 \text{ and } G_2 > 0.$$

$$\Gamma_{31} < 0 \text{ and } \Gamma_{12} < 0, \text{ then } G_1 < 0, \text{ and } G_2 \text{ and } G_3 > 0.$$

The first of the three equations which define the principal vectors $p_{\alpha i}$ can now be written as

$$[\Gamma_{11} - (\rho c^2)_\alpha]p_{\alpha 1} + G_1 G_2 p_{\alpha 2} + G_1 G_3 p_{\alpha 3} = 0,$$

and by adding and subtracting the term $G_1^2 p_{\alpha 1}$, this becomes

$$[\Gamma_{11} - (\rho c^2)_\alpha - G_1^2]p_{\alpha 1} + G_1(G_1 p_{\alpha 1} + G_2 p_{\alpha 2} + G_3 p_{\alpha 3}) = 0$$

so that

$$p_{\alpha 1} = G_1(G_i p_{\alpha i})/[(\rho c^2)_\alpha - \Gamma_{11} + G_1^2].$$

The other two components $p_{\alpha 2}$, and $p_{\alpha 3}$ follow the same form, so that for each root $(\rho c^2)_\alpha$, the components can be written as

$$p_{\alpha i} = G_i G_j p_{\alpha j}/[(\rho c^2)_\alpha - \Gamma_{ii} + G_i^2) \quad (i \text{ not summed}).$$

The division in this expression for $p_{\alpha i}$ is generally admissible because the denominator cannot be zero when the determinantal equation is satisfied.

Because the term $(G_j p_{\alpha j})$ is unknown, the vectors must be normalized to unit length by requiring the product

$$p_{\alpha i} p_{\beta i} = \delta_{\alpha\beta},$$

where α and β denote two roots.

The acceleration vector for each mode can then be written as

$$\partial v_{\alpha i}/\partial t = A p_{\alpha i}$$

where A is a scalar amplitude factor determined by boundary or initial conditions.

The material motion in each mode is then at an angle, the polarization angle, to the wave normal given by the formula

$$\delta = \mathrm{acos}(p_{\alpha i} n_i).$$

C. SPECIAL CASES

In the following cases of special directions, the method of finding the roots is different, so the mode sequence must be determined from their magnitudes.

C1. Propagation Vector (Wave Normal) Lies in a Plane of Two Axes

If the wave normal has no component along one of the axes (say the x-axis so that $l = 0$), then only one of the off-diagonal components of the stiffness matrix is nonzero and $\Gamma_{12} = \Gamma_{13} = 0$, and $\Gamma_{23} = mn(C_{44} + C_{23})$. The principal value equations are then

$$(\Gamma_{11} - \rho c^2)p_1 = 0$$
$$(\Gamma_{22} - \rho c^2)p_2 + \Gamma_{23}p_3 = 0$$
$$\Gamma_{23}p_2 + (\Gamma_{33} - \rho c^2)p_3 = 0$$

The first of these equations requires that either

$$(\Gamma_{11} - \rho c^2) = 0 \text{ or } p_1 = 0$$

and the other two require that either their determinant is zero,

$$\begin{vmatrix} (\Gamma_{22} - \rho c^2) & \Gamma_{23} \\ \Gamma_{23} & (\Gamma_{33} - \rho c^2) \end{vmatrix} = 0$$

or that $p_2 = p_3 = 0$.

For the first case of the first equation, i.e., the mode $\rho c^2 = \Gamma_{11}$, the determinant of the second and third equations is not zero, so that $p_2 = p_3 = 0$, and hence $p_1 = 1$. For the second case, the determinant is

$$(\Gamma_{22} - \rho c^2)(\Gamma_{33} - \rho c^2) - \Gamma_{23}^2 = 0,$$

which is solved as a quadratic equation to give

$$\rho c^2 = \{\Gamma_{22} + \Gamma_{33}) \pm [(\Gamma_{22} - \Gamma_{33})^2 + 4\Gamma_{23}^2]^{1/2}\}/2.$$

The first equation is not satisfied by the term $(\Gamma_{11} - \rho c^2)$, so that $p_1 = 0$. The second of the defining equations (and equivalently through the determinantal equation, the third) requires that the components of the principal vector be related by

$$(\Gamma_{22} - \rho c^2)p_2 = -\Gamma_{23}p_3,$$

so that

$$p_2 = A\Gamma_{23} \text{ and } p_3 = A(\Gamma_{22} - \rho c^2)$$

where A is a constant determined by requiring $p_2^2 + p_3^2 = 1$:

$$A = 1/[(\Gamma_{22} - \rho c^2)^2 + \Gamma_{23}^2]^{1/2}.$$

The results for the other cases when the wave normal has no component along the y-axis, $m = 0$, or along the z-axis, $n = 0$, can be written down by interchanging the indices 2 to 3 or 1 and 3 to 1 or 2, respectively.

Subcase: Transverse Isotropy

(i) The x_1-axis is the symmetry axis, so that axes 2 and 3 are equivalent.
Then

$$\rho c^2 = \Gamma_{11}, \text{ or } p_1 = 0$$

and

$$\rho c^2 = \Gamma_{22} + \Gamma_{23}, \text{ or } p_2 = p_3 = 0.$$

The three modes (wavespeeds and principal vectors) are then

$$\rho c_1^2 = \Gamma_{11}, \text{ and } p_1 = (1, 0, 0)$$
$$\rho c_2^2 = \rho c_3^2 = \Gamma_{22} + \Gamma_{23}, \text{ and } p_2 = (0, 1, 0), p_3 = (0, 0, 1).$$

(ii) The x_2 (or x_3) axis is the symmetry axis, so that x_1 and x_3 are equivalent, so that $\Gamma_{33} = \Gamma_{11}$ and $\Gamma_{23} = \Gamma_{13} = 0$.
Then

$$\rho c^2 = \Gamma_{11}, \text{ or } p_1 = 0, \text{ or } p_3 = 0$$

and

$$\rho c^2 = \Gamma_{22}, \text{ or } p_2 = 0.$$

The three modes (wavespeeds and principal vectors) are then

$$\rho c_1^2 = \Gamma_{11}, \text{ and } p_1 = (1/\sqrt{2}, 0, 1/\sqrt{2})$$
$$\rho c_2^2 = \Gamma_{22}, \text{ and } p_2 = (0, 1/\sqrt{2}, 1/\sqrt{2})$$
$$\rho c_2^2 = \Gamma_{11}, \text{ and } p_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0).$$

C2. Wave Normal Lies along One Axis

If the wave normal lies along one of the axes, then two components of the normal are zero, and all three off-diagonal stiffnesses are zero: In these cases, the normal is $(l, m, n) = (1, 0, 0)$, $(0, 1, 0)$, or $(0, 0, 1)$, and the principal value equations simplify to the following:

$$(\Gamma_{11} - \rho c^2)p_1 = 0$$
$$(\Gamma_{22} - \rho c^2)p2 = 0$$
$$(\Gamma_{33} - \rho c^2)p_3 = 0.$$

The three roots are then

$$\rho c^2 = \Gamma_{11}, \Gamma_{22}, \text{ or } \Gamma_{33}$$

and $p_2 = p_3 = 0$, or $p_3 = p_1 = 0$, or $p_1 = p_2 = 0$, respectively, so that the principal vectors are (1, 0, 0), (0, 1, 0), and (0, 0, 1).

C3. The Isotropic Case

In the isotropic case, the elastic coefficients are (see Appendix 2)

$$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu\ (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$$

so that the Christoffel stiffness is

$$\Gamma_{ik} = (\lambda + \mu)n_i n_k + \mu\ \delta_{ik}$$

The invariants are

$$\Gamma_1 = \lambda + 4\ \mu,\ \Gamma_2 = \mu\ (2\lambda + 5\ \mu),\ \text{ and } \Gamma_3 = \mu^2\ (\lambda + 2\ \mu).$$

Since Γ_1 is the sum of the three roots, and Γ_3 is their product, the solution can readily be seen to be

$$\rho c^2 = \lambda + 2\ \mu,\ \rho c^2 = \mu,\ \rho c^2 = \mu$$

so that

$$c = [(\lambda + 2\ \mu)/\rho]^{1/2},\ \text{ or } c = (\mu/\rho)^{1/2}.$$

These speeds are independent of direction, as they should be in an isotropic material, and are the longitudinal and transverse wavespeeds.

Alternatively, because of the form of Γ_{ik}, the determinantal equation can be written as

$$|n_i n_k - N\delta_{ik}| = 0$$

where $N = (\rho c^2 - \mu)/(\lambda + \mu)$.

On expanding the determinant, many terms cancel leaving only the following:

$$N^2(N - 1) = 0.$$

The solutions are $N = 0$ (i.e., $\rho c^2 = \mu$), or $N = 1$ (i.e., $\rho c^2 = \lambda + 2\ \mu$), as before.

The G parameters are

$$G_i = (\lambda + \mu)^{1/2} n_i.$$

a. Longitudinal Waves

For the root $\rho c^2 = \lambda + 2\ \mu$, the denominator of $p_{\alpha i}$ is $\lambda + \mu$, so that the normalized principal vectors are

$$p_i = n_i.$$

The acceleration vector, and therefore the velocity, is along the wave normal, i.e., the propagation direction,

$$v_i = v_N n_i,$$

where v_N is the normal excitation velocity for the wave. This wave is then called a longitudinal wave as discussed in Appendix 3.

b. Transverse Waves

When the roots $\rho c^2 = \mu$ are considered, the principal vectors are undefined because the denominator of $p_{\alpha i}$ is 0. According to matrix theory, they must be orthogonal to the longitudinal vector. Thus there are two perpendicular vectors lying in the plane of the wavefront:

$$a_i(\text{with } a_i n_i = 0), \text{ and } b_i = \varepsilon_{ijk} n_j a_k, \text{ so that } a_i b_i = n_i b_i = 0$$

where a_i lies in the x–y plane, and b_i lies in the N–z plane and both are determined by the excitation. There are thus two transverse waves.

The velocity in each of these waves is

$$v_i = v_T a_i \text{ or } v_T b_i$$

where v_T is the transverse excitation velocity.

D. GROUP VELOCITY

The wavespeed of a plane wave differs with direction, so those plane waves in a small cone of directions which have the same phase, $\phi = \kappa(k_i x_i - ct) = 0$, will combine to form a wavefront. As illustrated in Fig. 165, the plane waves propagating in two neighboring directions will intersect at a point which lies on a surface forming the wavefront. The distance to this point represents the distance propagated by the group of plane waves at the group velocity. The direction to the point represents the direction in which energy propagates. The following analysis is due to Musgrave (1954).

The phases of neighboring plane waves are given by the equations

$$k_i x_i - ct = 0, \text{ and } (k_i + dk_i)x_i' - (c + dc)t = 0$$

where x_i is a point on the first plane and x_i' one on the second. A group wave develops at time t when these two points coincide:

$$x_i = x_i'$$

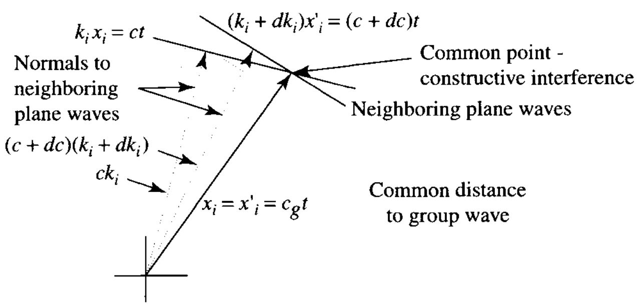


FIGURE 165 Illustration of group velocity as formed by intersection of plane waves

and the phases of the two waves are the same, so that

$$x_i dk_i - t dc = 0.$$

The propagation direction vector is a unit vector, $k_i k_i = 1$, so that

$$k_i dk_i = 0.$$

The wavespeed and wave number are related through the determinantal equation $\Delta(k_i, c) = 0$, so that their increments are related by the following:

$$(\partial\Delta/\partial k_i)dk_i + (\partial\Delta/\partial c)dc = 0.$$

Combining the three equations in the increments dk_i and dc using undetermined multipliers α and β gives the equation

$$(x_i dk_i - t dc) + \alpha k_i dk_i + \beta[(\partial\Delta/\partial k_i)dk_i + (\partial\Delta/\partial c)dc] = 0.$$

Collecting terms in each increment and requiring them to satisfy the equation separately, leads to the conditions

$$x_i + \alpha k_i + \beta(\partial\Delta/\partial k_i) = 0$$

$$-t + \beta(\partial\Delta/\partial c) = 0.$$

Then

$$\beta = t/(\partial\Delta/\partial c)$$

$$\alpha k_i = -x_i - \beta(\partial\Delta/\partial k_i).$$

Multiplying the second equation by k_i and using the conditions $k_i k_i = 1$ and $x_i k_i = ct$ leads to a solution for α:

$$\alpha = -[c + (\partial\Delta/\partial k_i)k_i/(\partial\Delta/\partial c)]t.$$

Inserting these into the equation for x_i gives the group velocity

$$c_{gi} = x_i/t = [ck_i + \{(\partial\Delta/\partial k_j)/(\partial\Delta/\partial c)\}(\delta_{ij} - k_i k_j)].$$

This consists of two terms, one having the magnitude and direction of the plane wave speed, and the other perpendicular to it, with the magnitude $dc/dk_i = (\partial\Delta/\partial k_i)/(\partial\Delta/\partial c)$. Evidently if $\partial\Delta/\partial c = 0$ the group velocity is indeterminate, but since that is the case for isotropy, the group velocity can be taken to be the wavespeed, so that $c_{gi} = ck_i$. The same result arises if $\partial\Delta/\partial k_i = 0$.

The group velocity has the magnitude

$$c_g = (c_{gi}c_{gi})^{1/2}$$

and is directed at an angle ε to the propagation vector k_i such that

$$\varepsilon = \mathrm{acos}(c_{gi}k_i/c_g) = \mathrm{acos}(c/c_g).$$

The determinantal equation connecting the increments dc and dk_i, illustrated in Fig. 166, is

$$\Delta(c, k_i) = (\rho c^2)^3 - \Gamma_1(\rho c^2)^2 + \Gamma_2(\rho c^2) - \Gamma_3 = 0,$$

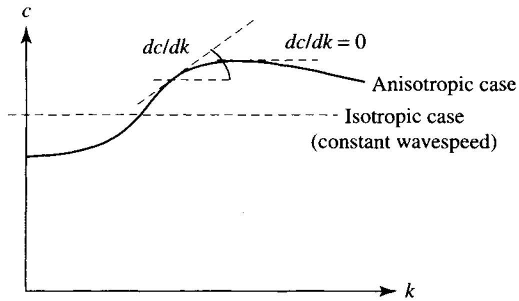


FIGURE 166 Illustration of wavespeed gradient

with $\Gamma_i = \Gamma_i(k_j)$, so that

$$(\partial\Delta/\partial c) = [3(\rho c^2)^2 - 2\Gamma_1(\rho c^2) + \Gamma_2]2\rho c,$$

and

$$(\partial\Delta/\partial k_i) = -(d\Gamma_1/dk_i)(\rho c^2)^2 + (d\Gamma_2/dk_i)(\rho c^2) - d\Gamma_3/dk_i.$$

The derivatives $(d\Gamma_i/dk_j)$ can be evaluated from the definitions of the invariants Γ_i and of the Christoffel stiffnesses Γ_{ij}. This evaluation is tedious but straightforward, but does not lend itself to direct algebraic manipulation.

This Page Intentionally Left Blank

APPENDIX 10

OBLIQUE INTERACTIONS BETWEEN WAVES AND BOUNDARIES

A. OBLIQUE FREE-SURFACE REFLECTION OF A LONGITUDINAL WAVE

Consider a longitudinal wave with potential function $\phi_I(\xi)$, where $\xi = mx + ny - c_l t$, m and n are the direction cosines of the wave normal, with $m^2 + n^2 = 1$, so that, taking θ as the angle between the wave normal and the x-axis, $m = \cos\theta$ and $n = \sin\theta$. Its wavespeed is c_l. This wave is incident on a free surface of an isotropic material at $y = 0$, as sketched in Fig. 167.

The phase created by the wavefront of the incident wave resolved along the surface $y = 0$ is $\xi' = mx - c_l t$. Any wave generated in reflection must have this same resolved phase along the surface, in order for the motions and stresses to be continuous. Hence any longitudinal reflection must have the same phase, $\xi = mx + ny - c_l t$.

It is customary to define the angle of incidence of a wave onto a surface as the angle between the wave normal (i.e., the propagation direction) and the surface normal:

$$\alpha = \pi/2 - \theta,$$

so that

$$m = \sin\alpha, n = \cos\alpha.$$

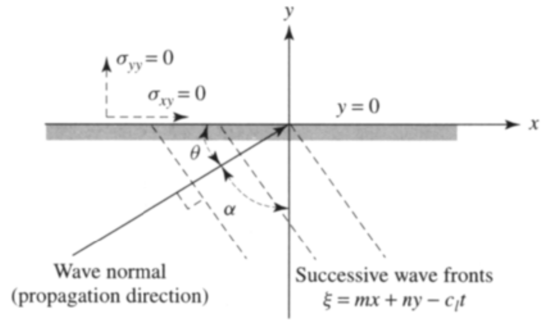


FIGURE 167 Oblique longitudinal wave incident on a free surface at y = 0

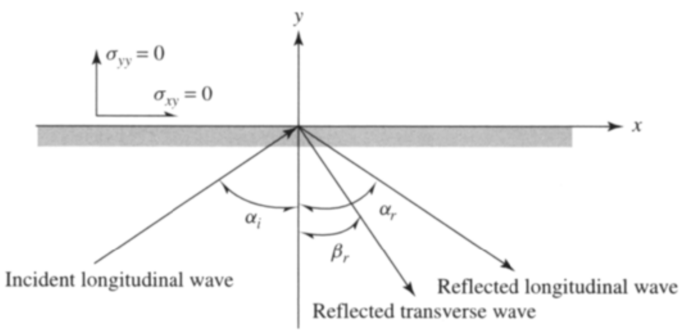


FIGURE 168 Definition of angles of incidence in oblique reflection of longitudinal wave

The free surface requires that the lateral stress σ_{yy} and the shear stress σ_{xy} at the surface $y = 0$ not be balanced by external forces; hence, both are zero. The interaction of the incident wave with the surface then imposes two conditions on the reflection process, and these can only be met by two waves. Thus reflection must induce both a longitudinal and a transverse wave as shown in Fig. 168. Note that in an anisotropic material there can be three components of stress, requiring three reflected waves.

A1. Snell's Law

The phase-matching constraint among the waves imposes a relationship, called Snell's Law as reviewed in Section 4, between the angles of reflection and incidence:

$$\sin\alpha_i/c_l = \sin\alpha_r/c_l = \sin\beta_r/c_s$$

where c_i is the wavespeed for the incident wave, and c_r is the speed for the reflection.

Thus for transverse reflection of an incident longitudinal wave,

$$\sin\beta_r = (c_s/c_l)\sin\alpha_i$$

so that $\beta_r = \text{asin}[(c_s/c_l)\sin\alpha_i] = \text{asin}\{[(1-2\nu)/2(1-\nu)]^{1/2}\sin\alpha_i\}$. This relationship is plotted in Fig. 59 in Section 4.

The incident and reflected longitudinal waves are at the same angle relative to the normal, i.e., $\alpha_i = \alpha_r$, because they have the same speed. The angle of the reflected transverse wave is less than that of the longitudinal wave (i.e., closer to the normal) because $c_s < c_l$.

The direction cosines n, p, q can all be expressed, for an isotropic solid, in terms of m:

$$n = (1-m^2)^{1/2}$$

$$p = (c_s/c_l)m = [(1-2\nu)/2(1-\nu)]^{1/2}m$$

$$q = (1-p^2)^{1/2} = [1-\{(1-2\nu)/2(1-\nu)\}m^2]^{1/2}.$$

A2. The Stress Potentials and Reflection Coefficients

The incident longitudinal wave is described by a wave potential function ϕ_i, and the longitudinal and transverse reflection waves are described by wave potential functions ϕ_r and A_{zr}, respectively, as reviewed in Appendix 6. In the reflections, the y-components of the wave vectors, i.e., n and q, are negative. The normal and shear stresses on the surface are

$$\sigma_{yy} = [C_l/(1-\nu)]\{[(1-\nu)n^2 + \nu m^2][\partial^2\phi_i/\partial\xi^2 + \partial^2\phi_r/\partial\xi^2]$$
$$+(1-2\nu)pq\partial^2 A_{zr}/\partial\eta^2\} = 0$$

$$\sigma_{xy} = [C_l(1-2\nu)/2(1-\nu)]\{2mn(\partial^2\phi_i/\partial\xi^2 - \partial^2\phi_r/\partial\xi^2)$$
$$+(q^2-p^2)\partial^2 A_{zr}/\partial\eta^2\} = 0.$$

For the surface to be stress-free, each of these expressions must be zero, leading to a pair of equations for the derivatives of the reflection potential functions in terms of the incident:

$$[(1-\nu)n^2 + \nu m^2]\partial^2\phi_r/\partial\xi^2 + (1-2\nu)pq\partial^2 A_{zr}/\partial\eta^2$$
$$= -[(1-\nu)n^2 + \nu m^2]\partial^2\phi_i/\partial\xi^2$$

$$2mn\partial^2\phi_r/\partial\xi^2 + (p^2-q^2)\partial^2 A_{zr}/\partial\eta^2 = 2mn\partial^2\phi_i/\partial\xi^2.$$

Note that if there were no transverse reflection (i.e., if $A_{zr} = 0$) and only a longitudinal reflection, no solution would be possible, since then the equations reduce to

$$\partial^2\phi_r/\partial\xi^2 = -\partial^2\phi_i/\partial\xi^2$$

$$\partial^2\phi_r/\partial\xi^2 = \partial^2\phi_i/\partial\xi^2.$$

Hence the reflection of an incident longitudinal wave must induce both a longitudinal and a transverse reflection.

The equations have the solutions

$$(\partial^2\phi_r/\partial\xi^2)/(\partial^2\phi_i/\partial\xi^2) = -\{(p^2 - q^2)[(1-\nu)n^2 + \nu m^2] + 2(1-2\nu)mnpq\}/\Delta$$

$$(\partial^2 A_{zr}/\partial\eta^2)/(\partial^2\phi_i/\partial\xi^2) = 4mn[(1-\nu)n^2 + \nu m^2]/\Delta$$

where $\Delta = [(1-\nu)n^2 + \nu m^2](p^2 - q^2) - 2(1-2\nu)mnpq$.

According to Appendix 6, the driving normal stress in a longitudinal wave is

$$\sigma_n = C_l \partial^2\phi/\partial\xi^2$$

and the driving shear stress for a transverse wave is

$$\sigma_{nt} = G\partial^2 A_z/\partial\eta^2$$

The solutions can then be written as the longitudinal-to-longitudinal wave normal stress reflection coefficient

$$C_{n,ll} = \sigma_{nr}/\sigma_{ni} = (\partial^2\phi_r/\partial\xi^2)/(\partial^2\phi_i/\partial\xi^2)$$

and as the longitudinal-to-transverse wave shear stress conversion coefficient

$$C_{nt,lx} = \sigma_{ntr}/\sigma_{ni} = [(1-2\nu)/2(1-\nu)][(\partial^2 A_{zr}/\partial\eta^2)/(\partial^2\phi_i/\partial\xi^2)].$$

These coefficients are plotted in Figs. 64 and 65 of Section 4.

Note that the reflection coefficient at 0^o is $C_n = -1$, independent of Poisson's ratio, ν. For $\nu = 0.5$ (equivalent to incompressibility) it is always -1, but switches to $+1$ as $\nu \to 0$. For $\nu < 0$, the behavior is different.

For normal incidence, when $m = 0$ (i.e., $\alpha_i = 0$), $n = 1$, $p = 0$, $q = 1$, then $C_{nl} = -1$ and $C_{ntl} = 0$, so that a longitudinal wave reflects as a longitudinal wave of opposite sign with no transverse wave, as discussed above.

B. OBLIQUE FREE SURFACE REFLECTION OF A TRANSVERSE WAVE

Consider a transverse wave with potential function $A_{zi}(\eta)$, where $\eta = px + qy - c_s t$ and $p^2 + q^2 = 1$, incident on a free surface of an isotropic material at $y = 0$. This wave can reflect as a transverse wave with potential $A_{zr}(\eta)$, where $\eta = px + qy - c_s t$ and a longitudinal wave with potential $\phi_r(\xi)$, where $\xi = mx + ny - c_l t$. This system of waves is illustrated in Fig 169.

A transverse wave with motion transverse to the plane of propagation, as discussed in Appendix 6, is not analyzed here, but the procedure is similar to the following.

As before, Snell's law applies between the angles of incidence of these waves so that

$$\sin\alpha_r = (c_l/c_s)\sin\beta_i$$

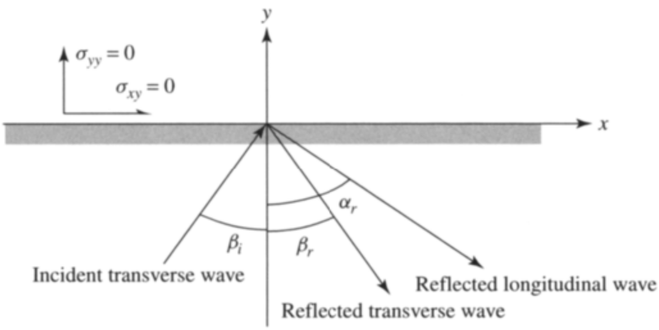


FIGURE 169 Definition of angles of incidence in oblique reflection of transverse wave

and $\alpha_r = \text{asin}[(c_{sl}/c_s)\sin\beta_i] = \text{asin}\{[(1-2\nu)/2(1-\nu)]^{-1/2}\sin\beta_i\}$. This relationship, which is essentially a rotation of that of Fig. 59 in Section 4, is plotted in Fig. 60 in Section 4. The critical angles beyond which no reflection is possible are shown in Fig. 61 in Section 4.

The direction cosines can be expressed in terms of the direction cosine of the incident transverse wave, p:

$$q = (1-p^2)^{1/2},$$
$$m = (c_l/c_s)p = [2(1-\nu)/(1-2\nu)]^{1/2}p$$
$$n = (1-m^2)^{1/2} = [1-2(1-\nu)/(1-2\nu)p^2]^{1/2}.$$

As before, the free surface signifies that the lateral stress σ_{yy} and the shear stress σ_{xy} at the surface $y = 0$ are not balanced by external forces, hence both are zero.

$$\sigma_{yy} = [C_l/(1-\nu)]\{-(1-2\nu)pq[\partial^2 A_{zi}/\partial\eta^2 - \partial^2 A_{zr}/\partial\eta^2] + [(1-\nu)n^2 + \nu m^2]\partial^2\phi_r/\partial\xi^2\} = 0$$
$$\sigma_{xy} = [C_l(1-2\nu)/2(1-\nu)]\{(q^2-p^2)[\partial^2 A_{zi}/\partial\eta^2 + \partial^2 A_{zr}/\partial\eta^2] - 2mn\partial^2\phi_r/\partial\xi^2\} = 0.$$

Then,

$$[(1-\nu)n^2 + \nu m^2]\partial^2\phi_r/\partial\xi^2 + (1-2\nu)pq\,\partial^2 A_{zr}/\partial\eta^2 = (1-2\nu)pq\,\partial^2 A_{zi}/\partial\eta^2$$
$$2mn\,\partial^2\phi_r/\partial\xi^2 + (p^2-q^2)\partial^2 A_{zr}/\partial\eta^2 = -(p^2-q^2)\partial^2 A_{zi}/\partial\eta^2.$$

Again, note that if no longitudinal reflection were present (i.e., if $\phi_r = 0$) but only a transverse reflection, no solution would be possible. Hence the reflection of an incident transverse wave must induce both a longitudinal and a transverse reflection.

The solutions to these equations define the reflection and conversion coefficients

$$C_{n,xl} = \sigma_{nr}/\sigma_{nti} = [2(1-\nu)/(1-2\nu)][(\partial^2\phi_r/\partial\xi^2)/(\partial^2 A_{zi}/\partial\eta^2)]$$
$$= 4(1-\nu)(p^2-q^2)pq/\Delta$$
$$C_{nt,xx} = \sigma_{ntr}/\sigma_{nti} = (\partial^2 A_{zr}/\partial\eta^2)/(\partial^2 A_{zi}/\partial\eta^2)$$
$$= -\{[(1-\nu)n^2+\nu m^2](p^2-q^2)+2(1-2\nu)mnpq\}/\Delta$$

where $\Delta = [(1-\nu)n^2+\nu m^2](p^2-q^2)-2(1-2\nu)mnpq$.

These solutions represent reflection coefficients relative to the magnitude of the incident wave, as plotted in Figs. 66 and 67 of Section 4.

C. REFRACTION (OBLIQUE TRANSMISSION) OF A LONGITUDINAL OR TRANSVERSE WAVE

Transmission of a wave across an interface between two materials induces two or three transmissions, and two or three reflections. This arises because there are four or six conditions to be met:

- A balance of the two or three stress components acting across the interface
- A match of the two or three displacements or velocity components across the interface so that there is no separation or interpenetration of the two materials

For an interface $y = 0$, these equations are

$$\sigma_{yyi} = \sigma_{yyr} + \sigma_{yyt}$$
$$\sigma_{xyi} = \sigma_{xyr} + \sigma_{xyt}$$
$$\sigma_{xzi} = \sigma_{xzr} + \sigma_{xzt}$$
$$v_{xi} = v_{xr} + v_{xt}$$
$$v_{yi} = v_{yr} + v_{yt}$$
$$v_{zi} = v_{zr} + v_{zt}.$$

In an isotropic material, when the waves lie in the x–y plane, the z-components are all zero, reducing the set of equations to four.

The transmitted waves are generally at different angles than the reflected, being in a different material, and are called refracted waves, as indicated in Fig. 68 of Section 4.

The incident wave may be a longitudinal wave with potential $\phi_i(\xi_1)$ where $\xi_1 = m_1x + n_1y - c_{l1}t$, or a transverse one with potential $A_{zi}(\eta_1)$, where $\eta_1 = p_1x + q_1y - c_{s1}t$, and the subscript 1 denotes the properties of material 1, that with the incident wave.

The phases for the longitudinal reflected and transmitted waves are $\phi_m(\xi_m)$, where $\xi_m = m_m x + n_m y - c_{lm} t$, and for the transverse waves they are $A_{zm}(\eta_m)$, where $\eta_m = p_m x + q_m y - c_{sm} t$, where $m = 1$ or 2 to distinguish the two materials. There are four or six applications of Snell's law to relate the direction cosines for the reflected and refracted waves to the incident, together with the four differential equation conditions $m_m^2 + n_m^2 = 1$ and $p_m^2 + q_m^2 = 1$.

For the isotropic case, the resulting four equations can be expressed in terms of the potential derivatives as in the free-surface reflection case, using the expressions for stress and velocity (given in Appendix 6). The result can be written in the form of a 4 × 4 matrix equation:

$$[M](X) = (Y)$$

where $[M]$ is the 4 × 4 matrix

$$\begin{pmatrix} -m_1/z_{l1} & -q_1/z_{s1} & q_2/z_{s2} & m_2/z_{l2} \\ -n_1/z_{l1} & -p_1/z_{s1} & p_2/z_{s2} & n_2/z_{l2} \\ n_1^2 + N_1 m_1^2 & n_2^2 + N_2 m_1^2 & 2p_1 q_1 & 2p_2 q_2 \\ 2M_1 m_1 n_1 & 2M_2 m_2 n_2 & (q_1^2 - p_1^2) & (q_2^2 - p_2^2) \end{pmatrix}$$

with

$$N_1 = [\nu_1/(1-\nu_1)],\ N_2 = [\nu_2/(1-\nu_2)],\ M_1 = [(1-2\nu_1)/2(1-\nu_1)],\ M_2 = [(1-2\nu_2)/2(1-\nu_2)]$$

and

$$(X) = \{[(\partial^2\phi_1/\partial\xi^2)/\rho_1 c_{l1}^2], [(\partial^2\phi_2/\partial\xi^2)/\rho_2 c_{l2}^2], [(\partial^2 A_1/\partial\eta^2)/\rho_1 c_{s1}^2], [(\partial^2 A_2/\partial\eta^2)/\rho_2 c_{s2}^2]\},$$

which is a vector (really, a grouping) comprising the four normalized potentials, and

$$(Y) = (v_{xi}, v_{yi}, \sigma_{yyi}, \sigma_{xyi}).$$

This vector specifies the incident wave state, whether on a transverse or a longitudinal wave.

A major complication is that the direction cosines for the waves differ because the wave speeds in the two materials are, in general, different. An algebraic solution is not feasible. Numerical methods must be used with specific properties, particularly the two Poisson's ratios.

D. RAYLEIGH WAVES

A Rayleigh surface wave runs along a free surface with a wavefront perpendicular to the surface. The motions decrease exponentially with depth from a maximum at the surface, whereas the stresses are zero at the surface, growing to a maximum at some depth, and then decreasing to zero exponentially. The motion is two-dimensional in a vertical plane.

The wave is a combination of longitudinal and transverse motions, both propagating at a wavespeed which is different from their speeds in an unbounded region.

The potential functions can be taken as follows:

$$\phi(x, y, t) = e^{-ay} f(\xi)$$
$$A_z(x, y, t) = e^{-by} g(\xi)$$

where $\xi = k(x - c_r t)$ is the phase of both, with a wavespeed c_r, which is the Rayleigh wavespeed and a and b are to be determined. Both potentials have the same dependence on x and t in order to have a coherent wave in which the stress components from each potential interact to meet the stress-free condition at the surface. The dependence on y need not be the same, as indicated by the indices a and b, nor are the amplitudes of the functions f and g.

The differential equations for the potentials given in Appendix 6 require that

$$k^2 d^2 f/d\xi^2 + a^2 f = (k^2 c_r^2/c_l^2)\, d^2 f/d\xi^2$$
$$k^2 d^2 g/d\xi^2 + b^2 g = (k^2 c_r^2/c_s^2)\, d^2 g/d\xi^2$$

where c_l is the longitudinal wavespeed, and c_s is the transverse wavespeed. These are two ordinary differential equations,

$$d^2 f/d\xi^2 + \alpha^2 f = 0$$
$$d^2 g/d\xi^2 + \beta^2 g = 0$$

with $\alpha^2 = a^2/k^2(1 - \gamma_l^2)$, $\beta^2 = b^2/k^2(1 - \gamma_s^2)$, and $\gamma_l = c_r/c_l$, $\gamma_S = c_r/c_s$. For the solutions to have the same dependence on ξ, $\alpha = \beta$, so that

$$a^2/(1 - \gamma_l^2) = b^2/(1 - \gamma_s^2),$$

and then

$$f = A \cos \alpha\xi + B \sin \alpha\xi$$
$$g = C \cos \alpha\xi + D \sin \alpha\xi$$

The displacements are

$$u_x = \partial f/\partial x + \partial A_z/\partial y = k e^{-ay}\, df/d\xi - b e^{-by} g$$
$$u_y = \partial f/\partial y - \partial A_z/\partial x = -a e^{-ay} f - k e^{-by}\, dg/d\xi$$

and the strains are

$$\varepsilon_{xx} = \partial u_x/\partial x = k^2 e^{-ay}\, d^2 f/d\xi^2 - bk e^{-by}\, dg/d\xi$$
$$\varepsilon_{yy} = \partial u_y/\partial y = a^2 e^{-ay} f + bk e^{-by}\, dg/d\xi$$
$$\varepsilon_{xy} = \partial u_x/\partial y + \partial u_y/\partial x = -2ak\, e^{-ay}\, df/d\xi + (b^2 g - k^2\, d^2 g/d\xi^2) e^{-by}.$$

The normal and shear stresses at the surface $y = 0$ are (from Appendix 3)

$$\sigma_{yy} = C_l\{a^2\phi + bkdg/d\xi + [\nu/(1-\nu)](k^2d^2f/d\xi^2 - bkdg/d\xi)\} = 0$$

$$\sigma_{xy} = C_l[(1-2\nu)/2(1-\nu)]\{-2ak\,df/d\xi + b^2g - k^2d^2g/d\xi^2\} = 0.$$

The freedom from stress at the surface $y = 0$ and the relationships between a, b, α, and k give four equations for the coefficients A to D of the functions f and g. The equations separate into identical pairs in A and D, and in B and -C:

$$\{(1-\gamma_l^2) - [\nu/(1-\nu)]\}A - [(1-2\nu)/(1-\nu)](1-\gamma_s^2)^{1/2}D = 0$$

$$2(1-\gamma_l^2)^{1/2}A + (2-\gamma_s^2)D = 0.$$

These two homogeneous equations in the two unknowns A and D can have nonzero solutions only if the determinant of the coefficients of A and D is zero. Using the definitions $\rho c_l^2/E = (1-\nu)/(1+\nu)(1-2\nu)$ and $\rho c_s^2/E = 1/2(1+\nu)$, so that $\gamma_l^2/\gamma_s^2 = c_s^2/c_l^2 = (1-2\nu)/2(1-\nu)$, allows the determinant to be written as

$$(2-\gamma_s^2)^2 = 4(1-\gamma_s^2)^{1/2}(1-\gamma_l^2)^{1/2}.$$

This can be expressed as a cubic equation for γ_s^2:

$$\gamma_s^6 - 8\gamma_s^4 + 8[(2-\nu)/(1-\nu)]\gamma_s^2 - 8/(1-\nu) = 0.$$

Solutions to this equation are obtained by a trial-and-error spreadsheet method increasing γ_s to determine when the function changes sign. One real root for the Rayleigh speed as a fraction of the shear speed, i.e., for γ_s, is found as a function of Poisson's ratio, as shown in Fig. 170. Also shown on this figure is

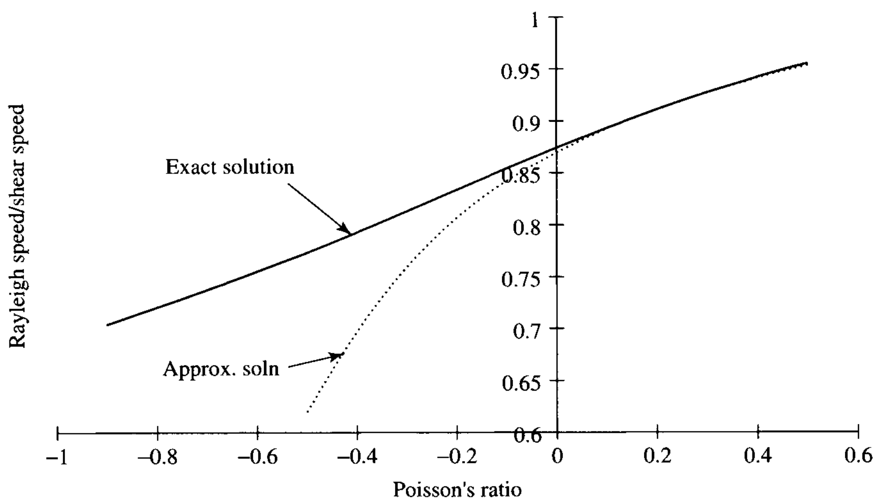


FIGURE 170 The dependence of the Rayleigh wavespeed on Poisson's ratio

the solution according to the approximate formula (Doyle, 1989)

$$\gamma_{s.approx} = (0.87 - 1.12\nu)/(1 - \nu).$$

Evidently, the approximation is only useful for positive values of ν. For this range, the Rayleigh wavespeed is around 0.87 to 0.95 of the shear wavespeed.

APPENDIX 11

LATERAL STRESS AND STRAIN IN RODS UNDER AXIAL LOADS

Slender rods are of concern in engineering design. Their behavior is influenced by their geometry, as discussed hereafter. A slender rod is one in which motion arises along the axis of the rod. Transverse deformations are called bending and are discussed in Appendix 12.

A. SIMPLE APPROXIMATION FOR THIN RODS

Simple wave propagation in a thin rod assumes that the rod can freely expand or contract laterally to relieve lateral stresses, but ignores inertia of this lateral motion. This leads to a uniaxial stress state with the simple wave equation

$$c_0^2 \partial^2 u/\partial x^2 = \partial^2 u/\partial t^2$$

where $c_0 = (E/\rho)^{1/2}$ is the rod or bar wavespeed, E is Young's modulus of elasticity, ρ is density, x is distance along the rod; and t is time. This approximation leads to waves which propagate at constant speed:

$$u(x, t) = u(x - c_0 t).$$

This is adequate for thin rods excited at low frequency, but thick rods and high frequencies demand a more detailed analysis.

B. THE MINDLIN–HERRMANN APPROXIMATION

There is no lateral motion near the central axis of a rod, so in a thick rod, the Poisson effect induces transverse stress at the axis. At the surface, lateral motion develops to relieve transverse stresses. The stress state and lateral motion then vary over the cross-section. An exact analysis of these effects has been made for a circular cylindrical rod by Pochhammer and Chree (discussed by Love, 1944, Article 201). Several approximations have been developed, such as the Mindlin–Herrmann approximation (see Doyle, 1989) which is discussed here.

Consider a rod loaded at one end so that a wave propagates along the axis, inducing an axial deformation with displacement $u_x(x, t)$, as illustrated in Fig. 171.

There is also a radial displacement $u_r(x, r, t)$, which is assumed to vary linearly with radius, $u_r(x, r, t) = rU_r(x, t)$. The strains are then

Axial:	$\varepsilon_{xx} = \partial u_x/\partial x$
Radial:	$\varepsilon_{rr} = \partial u_r/\partial r = U_r$
Circumferential (hoop):	$\varepsilon_{\theta\theta} = u_r/r = U_r = \varepsilon_{rr}$
Radial–axial shear:	$\varepsilon_{rx} = (\partial u_r/\partial x + \partial u_x/\partial r)/2 = (1/2)r\partial U_r/\partial x.$

The consequent stresses are

$$\sigma_{xx} = [E(1-\nu)/(1+\nu)(1-2\nu)]\{\partial u_x/\partial x + [2\nu/(1-\nu)]U_r\}$$

$$\sigma_{rr} = \sigma_{\theta\theta} = [E/(1+\nu)(1-2\nu)]\{U_r + \nu\partial u_x/\partial x\}$$

$$\sigma_{rx} = E[1/2(1+\nu)]r\partial U_r/\partial x$$

The radial stress should be zero at the free surface $r = a$, although in this approximate theory this condition is not met. Note that if there is no lateral (radial) stress (i.e., $\sigma_{rr} = 0$), then

$$U_r = -\nu\varepsilon_{xx}, \text{ so that } \sigma_{xx} = E\varepsilon_{xx},$$

which is the elementary rod theory. This is also the result if Poisson's ratio is zero.

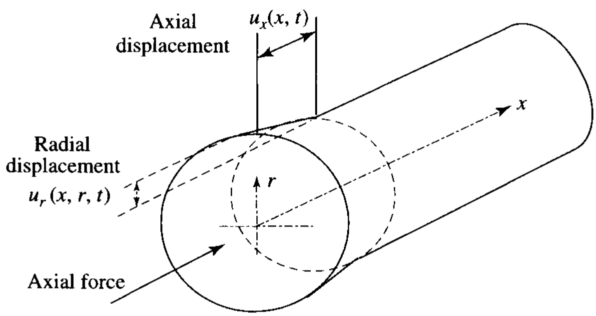


FIGURE 171 Slender rod with axial load

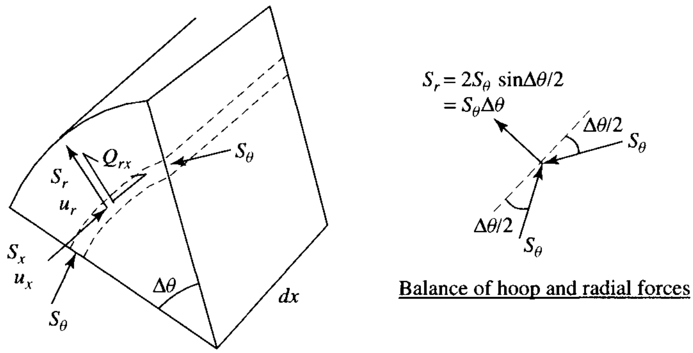


FIGURE 172 Radial and circumferential forces in a rod

The resultants of these stresses over the faces of a wedge, as illustrated in Fig. 172, are

$$S_x = \Delta\theta \int \sigma_{xx} r\, dr = [E(1-\nu)/(1+\nu)(1-2\nu)]\{\partial u_x/\partial x$$
$$+[2\nu/(1-\nu)]U_r\}a^2\Delta\theta/2$$
$$S_\theta = dx \int \sigma_{\theta\theta}\, dr = [E/(1+\nu)(1-2\nu)]\{U_r + \nu\partial u_x/\partial x\}a\, dx$$
$$Q_{rx} = k_1\Delta\theta \int \sigma_{rx} r\, dr = k_1 E[1/2(1+\nu)](\partial U_r/\partial x)a^3\Delta\theta/3$$

where k_1 is a correction factor which must be determined by experiment. Its default value, assuming the theory matches experiment, is unity:

$$k_1 = 1.$$

The total axial inertia force is

$$F_{xI} = \rho A\, dx\partial^2 u_x/\partial t^2 = \Delta\theta\, dx\rho a^2(\partial^2 u_x/\partial t^2)/2$$

so that the balance of axial forces is

$$(\partial F_x/\partial x)\, dx = F_{xI},$$

which leads to the equation

$$[E(1-\nu)/(1+\nu)(1-2\nu)]\{\partial^2 u_x/\partial x^2 + [2\nu/(1-\nu)]\partial U_r/\partial x\} = \rho\partial^2 u_x/\partial t^2.$$

The balance of circumferential forces, illustrated in Fig. 172, is

$$S_r = 2k_2 S_\theta \sin\Delta\theta/2 = k_2 S_\theta\Delta\theta$$

where k_2 is another correction factor.

The total radial (lateral) inertia force is

$$F_{rI} = \rho\, dx \int (\partial^2 U_r/\partial t^2) r\, dA = \Delta\theta \quad \rho\, dx(\partial^2 U_r/\partial t^2) \int r^2\, dr$$
$$= \Delta\theta\rho\, dx(\partial^2 U_r/\partial t^2)a^3/3$$

so that the balance of radial forces is

$$(\partial Q_{rx}/\partial x)\, dx - S_r = F_{rI}$$

i.e.,

$$(1/3)Ek_1a^2[1/2(1+\nu)]\partial^2 U_r/\partial x^2 - k_2[E/(1+\nu)(1-2\nu)]$$
$$\{U_r + \nu\partial u_x/\partial x\} = (1/3)\rho a^2\partial^2 U_r/\partial t^2.$$

For steady oscillations, assumed to be periodic in both time and position,

$$u_x = Ae^{-i(\kappa x - \omega t)}, \quad \text{and } u_r = Be^{-i(\kappa x - \omega t)}.$$

The axial equation then reduces to the following:

$$[E(1-\nu)/(1+\nu)(1-2\nu)]\{-\kappa^2 u_x - [2\nu/(1-\nu)]i\kappa U_r\} = -\rho\omega^2 u_x$$

so that

$$i\kappa U_r = [(1-\nu)/2\nu](1/c_l^2 - 1/c^2)\omega^2 u_x$$

where $c_l^2 = [(1-\nu)/(1+\nu)(1-2\nu)](E/\rho)$ gives the wavespeed for a longitudinal plane wave, and $\kappa = \omega/c$, where $c(\omega)$ is the wavespeed for the rod.

The radial equation becomes

$$(a^2\omega^2/3c_o^2)[1 - k_1(c_s^2/c^2)] - k_2(c_l^2/c_o^2)/(1-\nu)i\kappa U_r$$
$$= k_2(c_l^2/c_o^2)[\nu/(1-\nu)](\omega^2/c^2)u_x$$

where $c_s^2 = [1/2(1+\nu)](E/\rho)$ gives the shear wavespeed.

Note that if $k_1 = k_2 = 0$, then the radial equation reduces to $U_r = 0$, and the axial equation reduces to the simple case.

After substituting the expression for the lateral displacement U_r from the axial equation into the radial equation, this becomes a relationship between the normalized nondimensional wavespeed, $\beta = c/c_o$, where $c_o = (E/\rho)^{1/2}$ is the bar wavespeed, and the nondimensional frequency parameter, γ, where

$$\gamma = (a\omega/c_o).$$

This is 2π times the ratio of the time for a wave to cross the radius, a, i.e., $t_r = a/c_o$, to the time period of the wave, $t_p = 2\pi/\omega$.

The result is a quadratic equation in β^2:

$$a(\beta^2)^2 + b\beta^2 + c = 0$$

with

$$a(\gamma) = \{k_2 - (1+\nu)(1-2\nu)\gamma^2/3\}$$
$$b(\gamma) = -\{k_2 - [k_1(1-2\nu) + 2(1-\nu)]\gamma^2/6\}$$
$$c(\gamma) = -k_1(1-\nu)\gamma^2/6(1+\nu).$$

There are two solutions for β^2,

$$\beta^2 = [-b \pm (b^2 - 4ac)^{1/2}]/2a,$$

so that β has four solutions, of which pairs are equal and opposite, representing waves propagating in opposite directions. Solutions can be found using specific values of Poisson's ratio, or values can be computed numerically for varying Poisson's ratio. This requires estimates of the correction factors k_1 and k_2, which can be set initially at 1, and changed to produce accord with experimental data.

Solutions for β^2 are always real because it can be shown that the term $(b^2 - 4ac)$ reduces to the sum of two squares, i.e., the radical is always positive.

Solutions for β are real if $\beta^2 > 0$, and this is always true for the positive radical. For the negative radical, it is true provided $a < 0$, i.e.,

$$\gamma^2 > 3k_2/(1+\nu)(1-2\nu).$$

There is thus a cutoff circular frequency

$$\omega_c = 3k_2 c_o/a(1+\nu)(1-2\nu)$$

below which the negative radical produces an imaginary solution.

Note that at low frequencies as $\omega \to 0$, $\gamma \to 0$, so that $a \to k_2$, $b \to -k_2$, and $c \to 0$, and the solutions simplify to

$$\beta^2 \to 1 \text{ or } 0, \quad \text{i.e., } c \to c_o \text{ or } 0.$$

The factors k_1 and k_2 do not enter the low-frequency response.

When $k_1 = k_2 = 1$,

$$a = 1 - (1+\nu)(1-2\nu)\gamma^2/3$$
$$b = -[1 - (3-4\nu)\gamma^2/6]$$
$$c = -(1-\nu)\gamma^2/6(1+\nu).$$

A graph of the dependence of β (i.e., of c/c_o) on γ (i.e., on $a\omega/c_o$) for these assumptions, which represents the approximate nondimensional dispersion curves for a circular rod, is shown for a range of Poisson's ratios in Fig. 157 in Appendix 7.

There are two modes of response, corresponding to the two solutions. The first of the two modes represents waves which, for low frequency, propagate at the bar velocity as in the simple theory. At high frequency, exact solutions (not discussed here) show that the wavespeed is the Rayleigh speed. The other mode is complex for low frequencies, representing an attenuating wave, but at high frequencies above the cutoff it has real wavespeed which starts at infinity and drops to the bar velocity at infinite frequency.

The group velocity (see Appendix 7) is given by the formula:

$$c_g = c/[1 - (\omega/c)(dc/d\omega)].$$

It follows that this can also be written as

$$c_g = c/[1 - (\omega^2/c^2)(dc^2/d\omega^2)]$$

and then in the nondimensional form

$$\beta_g = \beta/[1 - (\gamma^2/\beta^2)(d\beta^2/d\gamma^2)].$$

Since

$$a(\beta^2)^2 + b\beta^2 + c = 0,$$

differentiation gives

$$(da/d\gamma^2)(\beta^2)^2 + (db/d\gamma^2)\beta^2 + (dc/d\gamma^2) + [2a\beta^2 + b](d\beta^2/d\gamma^2) = 0$$

and then

$$(d\beta^2/d\gamma^2) = -[(da/d\gamma^2)(\beta^2)^2 + (db/d\gamma^2)\beta^2 + (dc/d\gamma^2)]/[2a\beta^2 + b]$$

with

$$da/d\gamma^2 = -(1+\nu)(1-2\nu)/3$$
$$db/d\gamma^2 = [k_1(1-2\nu) + 2(1-\nu)]/6$$
$$dc/d\gamma^2 = -k_1(1-\nu)/6(1+\nu).$$

The group velocities are shown in Fig. 158 in Appendix 7.

APPENDIX 12

BENDING WAVES IN BEAMS AND PLATES

Beams and thin plates are of concern in engineering design. Their behavior is influenced by their geometry, as discussed later. A beam is a slender bar in which motion occurs transverse to the axis, called bending. A plate is a thin sheet in which deformation can occur in the direction of the sheet or transverse to it. A beam is essentially one-dimensional. A plate is two-dimensional. Motion in the plane of a plate is called axial or membrane motion.

A. BENDING OF BEAMS

The simplest analysis of bending at low frequency is the Engineers' Bending Theory (EBT), also known as the Euler–Bernoulli Theory (EBT). More detailed analyses which are appropriate for high frequencies are the Timoshenko and the Mindlin–Herrmann theories (see Doyle, 1989). An exact analysis is the Pochhammer–Chree theory for rods (Love, 1944).

A1. The EBT

In the EBT, the shear force is indeterminate because rotation of the cross-section is taken to be the same as the slope of the deformation. This is usually stated as

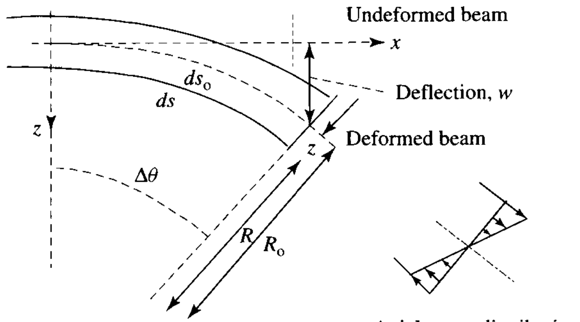


FIGURE 173 Illustration of strain–curvature relationship

"plane sections remain plane and normal," as depicted in Fig. 173: a line across the beam normal to the center line becomes a radial line.

The cross-section does not rotate relative to the axis in this model and there is no shear deformation. The lateral displacement of the beam, $w(x, t)$, depends only on distance, x, along the beam. It is taken here as positive when in the direction of increasing z. An axial displacement $u(x, z, t)$ arises which varies across the depth and depends on x. It is taken as positive when in the direction of increasing x.

Referring to Fig. 173, the length of a short arc at depth z (taken positive downward) is $ds = R\Delta\theta$, and assuming that the center line does not change so that its undeformed length is $ds_o = R_o\Delta\theta$. The increment of axial displacement is

$$du(x, z) = ds - ds_o = (R - R_o)/\Delta\theta = -z/\Delta\theta$$

and the strain is

$$\varepsilon(x, z) = \partial u/\partial s_o = -z/R_o = -\kappa z$$

where $\kappa = 1/R_o = \partial^2 w/\partial x^2$ is the curvature along the axis, taken positive for downward curvature along increasing x. (This relationship can be derived through differential geometry.) The strain is thus linear over the cross-section and is negative (i.e., compressive) on the underside.

In the absence of lateral stress in a thin beam, the stress state is uniaxial. The axial stress is then

$$\sigma(x, z) = E\varepsilon(x, z)$$

where E is Young's modulus of elasticity. The resultant axial force on the cross section is

$$F_x = \int \sigma(x, z) dA = E\kappa \int z dA.$$

This is zero when the $z = 0$ axis is taken to be along the centroidal axis of the cross-section.

The stress has a resultant bending moment obtained by integrating the torques provided by the axial stress over the area of the cross-section,

$$M(x) = -\int \sigma(x, z) z dA = EI\kappa = EI\partial^2 w/\partial x^2.$$

Positive moment is defined to produce positive curvature, or, in the vernacular, hogging.

$I = \int z^2 dA = \int z^2 b(z) dz$ is the second moment of area of the cross-section, where $b(z)$ is the breadth of the cross-section at depth z. This is often called the moment of inertia, but since density is not involved, it is a purely geometric construct, also called the second moment of area. I can also be written in terms of the radius of gyration, k, so that $I = Ak^2$. The product EI is called the elastic bending stiffness.

In a rectangular bar the breadth $b(z)$ is constant, and the depth spans from $-d/2$ to $d/2$, where d is the depth of the beam, so that

$$I = \int_{-d/2}^{d/2} z^2 b(z) dz = bd^3/12.$$

Then $k = d/(2\sqrt{3})$.

Consider an element of a beam with external forces (loadings) as illustrated in Fig. 174. These include a bending moment, M, a shear force, Q, at each end, and a distributed force, p, over the length, as well as the inertia force resisting deformation. Although there is no shear strain, a shear force is required to balance the bending moment in cases where there is no transverse pressure loading.

A balance of the downward transverse forces normal to the beam requires that

$$dQ + \int p\, dx = \rho A\, dx \partial^2 w/\partial t^2,$$

so that

$$\partial Q/\partial x = \rho A d^2 \partial^2 w/\partial t^2 - p,$$

where the limit of $(\int_x^{x+dx} p dx)/dx \to p$.

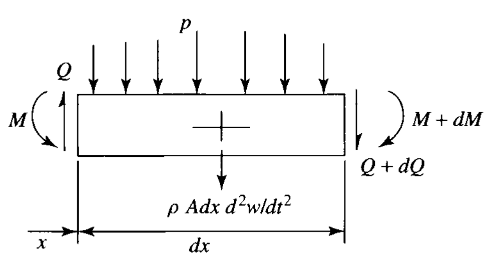


FIGURE 174 External forces on an element of a beam

A balance of the bending moments about the right-hand end of the element requires that

$$dM + Q\,dx - \int px\,dx = 0,$$

when rotatory inertia is neglected, so that

$$\partial M/\partial x + Q = 0,$$

since the limit of $(\int_x^{x+dx} pxdx)/dx = O(dx) \to 0$.

In combination with the transverse equation above, this leads to an equation of lateral motion in which the undetermined shear force, Q, plays no role:

$$\partial^2 M/\partial x^2 + \mu\partial^2 w/\partial t^2 = p,$$

where $\mu = \rho A$ is the areal mass, or mass per unit length. Using the elasticity relationship for bending moment given earlier, i.e., $M = EI\partial^2 w/\partial x^2$, this leads to a fourth-order equation for displacement:

$$EI\partial^4 w/\partial x^4 + \mu\partial^2 w/\partial t^2 = \partial^2 p/\partial x^2.$$

Without lateral forces, harmonic waves of the type

$$w(x,t) = We^{-i(\alpha x - \omega t)}$$

obey the dispersion equation

$$\alpha^4 - (\mu/EI)\omega^2 = 0, \text{ or } \omega^2 = (\mu/EI)c^4,$$

where $\alpha = \omega/c$, so that

$$c = (EI/\mu)^{1/4}\omega^{1/2} = (\omega c_0/k)^{1/2},$$

where $c_0 = (E/\rho)^{1/2}$ is the bar wavespeed, k is the radius of gyration so that $I = Ak^2$, and $\mu = \rho A$, where A is the cross-sectional area and ρ is the density.

In nondimensional form, this can be written as

$$\beta = \gamma^{1/2}$$

where $\beta = c/c_0$ and $\gamma = \omega k/c_0$ is the nondimensional frequency. Evidently the bending wave is always dependent on frequency and for low frequency approaches zero.

A2. The Timoshenko Bending Theory

In this model, both shear deformation and rotation of the cross-section are considered, as illustrated in Fig. 175. The equations for simple bending are now modified by the addition of the relationship for transverse shear and by the inclusion of the inertial effects of rotation of the cross-section.

The sides of a rectangular element $dx\,dz$ of the beam at x are deflected downward (positive deflection) at the right-hand end by a displacement $w(x,t)$, and to the right (positive) at the bottom by $u(x,z) = -\phi z$, where $\phi(x,t)$ is

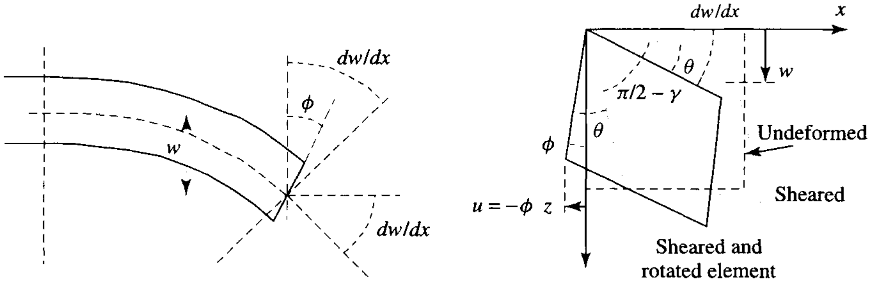


FIGURE 175 Shear deformation and cross-section rotation in a timoshenko beam

the rotation of a cross-section relative to the original normal cross-section. The element is sheared by an angle γ and rotated by an angle θ. This assumes that the plane cross-section remains plane but not normal to the axis as in the case of simple bending, but it rotates by an angle ϕ.

The axial strain is then

$$\varepsilon_x = \partial u/\partial x = -z\partial\phi/\partial x.$$

The axial stress is $E\varepsilon_x = -Ez\partial\phi/\partial x$, so that the bending moment, positive to increase deflection, is

$$M = -\int \sigma(x, z) z dA = E\,(\partial\phi/\partial x)\int z^2 dA = EI\partial\phi/\partial x.$$

The upper face of the element is rotated through an angle $(\partial w/\partial x - \theta)$ relative to its original horizontal direction, and the side is rotated through an angle $(\theta - \phi)$ relative to the normal. The element is thus sheared by the angle

$$\gamma = (\partial w/\partial x - \theta) + (\theta - \phi) = (\partial w/\partial x - \phi).$$

The shear stress is

$$\tau = G\varepsilon_{xz} = G(\partial w/\partial x - \phi),$$

where G is the shear modulus. This has a resultant shear force on the right-hand end, positive to increase deflection, of

$$Q = \int \tau dA = k_s GA(\partial w/\partial x - \phi),$$

where k_s is a correction factor introduced to account for approximations and which is found from experimental data. The factor GA is called the shear stiffness.

Rotation of the cross-section induces a rotatory inertia torque

$$T_\phi = k_\phi \rho I \partial^2\phi/\partial t^2$$

where k_ϕ is a correction factor accounting for approximations. The balance of moments, or the equation of rotatory motion, is then

$$\partial M/\partial x + Q = k_\phi \rho I \partial^2\phi/\partial t^2.$$

Together with the equation for the balance of lateral forces, which is the same as in the EBT;

$$\partial Q/\partial x = \mu \partial^2 w/\partial t^2 - p,$$

there are now four equations for the variables M, Q, w, ϕ. These can be combined to give two in the deformation unknowns, w and ϕ:

$$EI\partial^2\phi/\partial x^2 + k_s GA(\partial w/\partial x - \phi) = k_\phi \rho I \partial^2\phi/\partial t^2$$

$$k_s GA(\partial^2 w/\partial x^2 - \partial\phi/\partial x) = \mu\partial^2 w/\partial t^2 - p.$$

Considering a harmonic-wave-like motion in which

$$w = We^{-i(\alpha x - \omega t)} = We^{-i\alpha(x-ct)}$$

$$\phi = \Phi e^{-i(\alpha x - \omega t)} = \Phi e^{-i\alpha(x-ct)},$$

with $c = \omega/\alpha$, the equations become

$$-EI\alpha^2\Phi + k_s GA(-i\alpha W - \Phi) = -k_\phi \rho I \omega^2 \Phi$$

$$k_s GA(-\alpha^2 W + i\alpha\Phi) = -\mu\omega^2 W + p.$$

When there is no distributed applied force ($p = 0$), the second equation gives

$$i\alpha\Phi = (\alpha^2 - \mu\omega^2/k_s GA)W.$$

After substitution into the first equation this gives the nondimensional dispersion (frequency–wavespeed) relationship

$$EI\alpha^4 - \mu\omega^2 - [k_\phi \rho I + (EI\mu)/(k_s GA)]\alpha^2\omega^2 + (k_\phi \rho I \mu)/(k_s GA)\omega^4 = 0,$$

which is of the form

$$a(\beta^2)^2 + b\beta^2 + c = 0,$$

with

$$a = 1 - 2(1+\nu)(k_\phi/k_s)\gamma^2$$

$$b = [2(1+\nu)/k_s + k_\phi]\gamma^2$$

$$c = -\gamma^2,$$

where $\beta = c/c_o$ is a nondimensional wavespeed, $\gamma = \omega k/c_o$ is a nondimensional frequency, $c_o = (E/\rho)^{1/2}$ is the bar speed, and ν is Poisson's ratio.

There are two modes of propagation, given (when the unknown correction factors are taken to be unity, $k_s = k_\phi = 1$) by the solutions

$$\beta^2 = [-b \pm (b^2 - 4ac)^{1/2}]/2a,$$

with

$$a = 1 - 2(1+\nu)\gamma^2$$

$$b = (3 + 2\nu)\gamma^2$$

$$c = -\gamma^2.$$

Since the term $(b^2 - 4ac)$ in the radical can be combined to give terms which are squared, the radical is real for both roots. The solution for β^2 with the positive radical is always positive, so that the solution for β is always real. The negative radical gives negative β^2 and thus imaginary (nonpropagating) wavespeed, unless the second term in the radical is negative:

$$\gamma > [2(1+\nu)]^{-1/2} \text{ i.e., } \omega > c_s/k$$

where $c_s = (G/\rho)^{1/2} = c_o/[2(1+\nu)]^{1/2}$ is the shear speed.

That is, the cutoff frequency represents the time taken for a shear wave to travel across the radius of gyration, a fraction of the thickness. For frequencies above this limit, this mode gives real wavespeeds, so this is a low-frequency cutoff.

The wavespeed is plotted for several values of Poisson's ratio in Fig. 159 in Appendix 7.

The group velocity, calculated by the formulae of Appendix 11, is plotted in Fig. 160 of Appendix 7.

B. THIN PLATES

A thin plate can experience either in-plane deformations or transverse deformations. The former are referred to as membrane motions, and the latter as bending.

Associated with these deformations are in-plane forces applied to the edges of the plate for membrane motions, and lateral surface forces, or shear forces and moments applied to the edges for bending.

A plate is thin in the lateral or transverse direction, having significant size in the other two dimensions. It is essentially two-dimensional, as illustrated in Fig. 176. The simplest analysis of a plate is the classical plate theory (CPT). More complex analyses, e.g., the Mindlin–Herrmann theory (Doyle, 1989), given earlier in simplified form for a beam, are available for plates, but are not discussed here.

B1. The Classical Plate Theory (CPT)

In a plate the displacement, $w(x, y, t)$, depends only on in-plane position, (x, y) and time, t. Axial displacements, $u(x, y, z, t)$ and $v(x, y, z, t)$, arise which vary across the depth and depend on position, (x, y) and time, t.

The equations which describe plate response have the same basis as for a beam, but include two-dimensional features. In addition, when plates are made up of anisotropic laminates, as is common in reinforced plastic design, the axial and bending behavior are not separable, as has been done for rods and beams. Thus there are in-plane axial and shear forces, $\mathbf{N} = (N_x, N_y, N_{xy})$ and bending moments, $\mathbf{M} = (M_x, M_y, M_{xy})$, which are related to axial and shear strains,

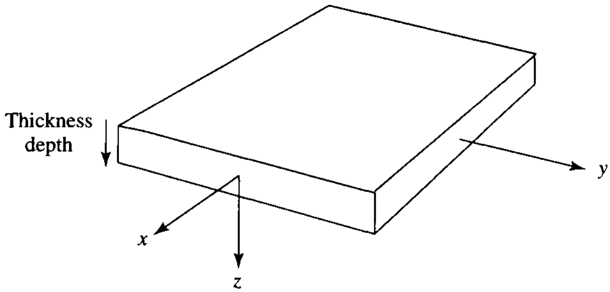


FIGURE 176 Geometric description of a plate

$\varepsilon = (\varepsilon_x, \varepsilon_y, \gamma_{xy})$ and to curvatures $\kappa = (\kappa_x, \kappa_y, \kappa_{xy})$:

$$\mathbf{N} = [A]\varepsilon + [B]\kappa$$

$$\mathbf{M} = [B]\varepsilon + [D]\kappa$$

where $[A]$, $[B]$, and $[D]$ are 3×3 matrices of elastic constants, involving integrals over the thickness, and therefore including thickness, and first and second moments of depth. The arrangements of the individual laminae in a laminate is therefore significant. Accordingly changes in any layer should be detectable by their influence on wavespeed. Note that in general, the in-plane and bending responses are coupled through the coefficients [B]. In a laminate with a symmetric arrangement of layers through the thickness, the coefficients [B] are zero, and the in-plane and bending responses are then uncoupled.

For a laminate of n layers,

$$[A] = \sum_{k=0}^{n} \mathbf{Q}_k(z_k - z_{k-1}), [B] = \sum_{k=0}^{n} \mathbf{Q}_k(z_k^2 - z_{k-1}^2), [D] = \sum_{k=0}^{n} \mathbf{Q}_k(z_k^3 - z_{k-1}^3),$$

and $\mathbf{Q}$ is the 3×3 matrix of plane-stress elastic stiffnesses rotated to the appropriate axes for each layer.

Transverse shear forces, as in the EBT, are not related to deformation in CPT. The advanced theory by Mindlin and Herrmann includes such a relationship.

The strains are defined from the axial displacements (u, v) as

$$\varepsilon_x = \partial u/\partial x, \varepsilon_y = \partial v/\partial y, \text{ and } \gamma_{xy} = (\partial u/\partial y + \partial v/\partial x),$$

and the curvatures are defined from the lateral displacement w:

$$\kappa_x = -\partial^2 w/\partial x^2, \kappa_y = -\partial^2 w/\partial y^2, \kappa_{xy} = -\partial^2 w/\partial x \partial y.$$

The equations of motion for the in-plane and transverse motions are

$$\partial N_x/\partial x + \partial N_{xy}/\partial y = \mu \partial^2 u/\partial t^2,$$

$$\partial N_y/\partial y + \partial N_{xy}/\partial x = \mu \partial^2 v/\partial t^2$$

$$\partial Q_x/\partial x + \partial Q_y/\partial y = \mu \partial^2 w/\partial t^2.$$

The balance of moments is described by the equations

$$\partial M_x/\partial x + \partial M_{xy}/\partial y + Q_x = 0$$

$$\partial M_{xy}/\partial x + \partial M_y/\partial y + Q_y = 0,$$

and these lead to the equation for lateral motion:

$$\partial^2 M_x/\partial x^2 + 2\partial^2 M_{xy}/\partial x \partial y + \partial^2 M_y/\partial y^2 = \mu \partial^2 w/\partial t^2.$$

For a simple plate of one material, the in-plane equations lead to a set of equations for the displacements, similar to rod waves:

$$A_{ij}\partial^2 u_k/\partial x_i \partial x_j + \mu \partial^2 u_k/\partial t^2 = 0,$$

and the bending equation becomes a two-dimensional extension of simple beam theory:

$$D_{ijkl}\partial^4 w/\partial x_i \partial x_j \partial x_k \partial x_l + \mu \partial^2 w/\partial t^2 = 0.$$

This equation requires numerical solution similar to that used for waves in anisotropic materials, discussed in Appendix 9.

This Page Intentionally Left Blank

APPENDIX 13

TIME-DOMAIN ANALYSIS

Details of mathematical concepts and algorithms for some of the data analysis procedures used in the time-domain software described in Section 5 are presented below.

A. THE HILBERT TRANSFORM AND THE ANALYTIC ENVELOPE AND PHASE

An envelope of the waveform, which provides a visual interpretation of the amplitude variation in the waveform, can be created through a numerical procedure based on the Hilbert Transform. The procedure also creates a time-domain phase function.

A1. The Hilbert Transform

The Hilbert Transform is an integration operator which can be perceived as performing a differentiation. The Hilbert transform of $\sin(x)$ is $\cos(x)$, and conversely, that of $\cos(x)$ is $-\sin(x)$.

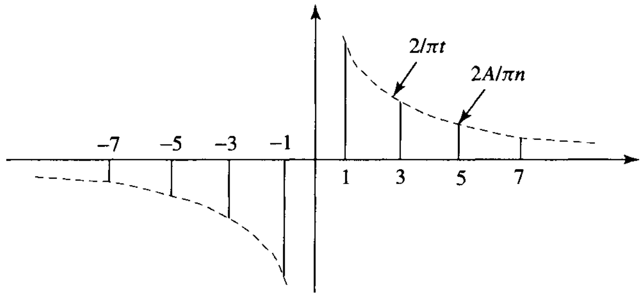


FIGURE 177 Numerical kernel for Hilbert transform

To generate the Hilbert transform, the points of a waveform are multiplied by a function of time which is antisymmetric about 0, and the result is integrated over a wide range of time. This function, called the kernel, is $k(t) = 1/t$, so that the transform at time t of a waveform $f(t)$ is

$$H(t) = (1/\pi) \int_{-\infty}^{\infty} [f(\tau)/(t-\tau)]d\tau.$$

The kernel $k(t) = 1/t$ is infinite at $t = 0$ so that in numerical procedures this point is eliminated and the kernel is approximated by a set of discrete values as sketched in Fig. 177.

$$\begin{aligned} k_n &= 2A/\pi n && \text{for } n \text{ odd integers} \\ &\;\; 0 && \text{for } n \text{ even integers} \end{aligned}$$

where A is an amplitude correction factor which has been tested for typical UT waveforms and found to be around 1.2.

The transform of a waveform with discrete sampling is then

$$H_m = \sum_{n=-N}^{N} k_n w_{n-m}$$

where the summation is typically taken over 15 points ($N = 7$).

A2. The Analytic Envelope

The analytic envelope is based on the concept that the total energy of an elastic system, $E = U + K$, is constant where U is the potential energy of elastic deformation (typically, $U \sim u^2$, where u is the displacement) and K is the kinetic energy (typically, $K \sim v^2$, where v is the velocity, $v = du/dt$). The analytic envelope is similar to the energy at a point of the waveform and is given by

$$E_m = (w_m^2 + H_m^2)^{1/2}.$$

An analytic envelope provides a clear display of the waveform, as illustrated in Fig. 10 in Section 2.

A3. The Analytic Phase

The analytic phase can be defined as

$$\phi_{\text{an}}(t) = \text{atan}(H_n/w_n).$$

Because the atan function only provides values from $-\pi$ to π, this phase wraps between these values if it exceeds them. An unwrapping procedure is described in Section 6B.

B. TIME-DOMAIN SIGNAL CONDITIONING

Several techniques are used to condition waveforms so that they are smoother and better suited for further analysis, particularly spectral analysis.

B1. Data Averaging

Spurious random effects in a waveform, such as might be caused by electrical or environmental variations, can be minimized by repeated acquisition of the signal from the test configuration. The assumption is made that these variations differ from waveform to waveform, whereas the signal does not. The signal is then improved by averaging the waveforms. It is neither desirable or necessary to store all the waveforms.

A sequence of nominally identical waveforms is shown in Fig. 178, together with an average waveform. The average $S_{\text{av}.N}$ of the N values S_n of the waveform at each time in N tests is calculated using the formula for the average:

$$S_{\text{av}}.N = \left(\sum_{n=1}^{N} S_n\right) / N$$

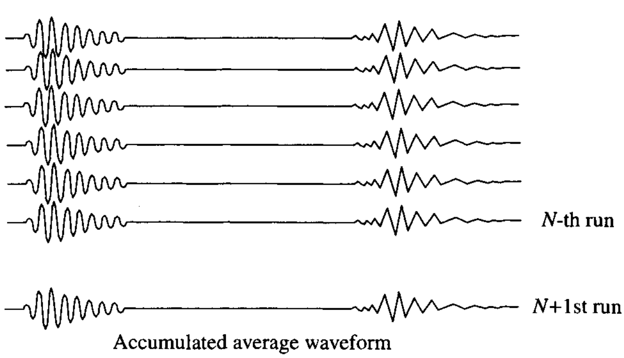


FIGURE 178 A sequence of nominally identical waveforms

where n denotes the nth run of the series of tests. When a new waveform is acquired, the average is updated for $N + 1$ points, so that

$$S_{av.N+1} = \left(\sum_{n=1}^{N+1} S_n\right)/(N+1) = \left(\sum_{n=1}^{N} S_n\right)/(N+1) + S_N/(N+1)$$
$$= [N/(N+1)]S_{av.N} + S_{N+1}/(N+1).$$

Thus the average is updated for each new waveform by multiplying the values of the previous average by the factor $N/(N+1)$, and adding the latest value divided by $(N + 1)$. This process requires storing only one waveform. The process is repeated, starting from $N = 1$, the first waveform, typically 100 to 1000 times, until the waveform settles to a stable waveform.

B2. Time-Domain Smoothing

A waveform with a large number of points which exhibits noise, i.e., erratic variation from point to point, can be smoothed by taking averages over a group of neighboring points (3 to 15 is common). Typically, the smoothed value $S_{N.\mathrm{smooth}}$ for the center point N of the group is replaced by the average:

$$S_{N.\mathrm{smooth}} = \left(\sum_{n=n_{\min}}^{n_{\max}} S_{N+n}\right)/n_{\mathrm{pts}}$$

where n denotes the point number within a group of n_{pts} (an odd number) spread about N, $n_{\min} = N - (n_{\mathrm{pts}} - 1)/2$, $n_{\max} = N + (n_{\mathrm{pts}} - 1)/2$, and S_{N+n} is the original value at $N + n$, as illustrated in Fig. 179.

Thus a "comb" which selects a group of points is swept across the waveform to produce a new one. Note that the points at the start and end of the waveform cannot be smoothed. They may be deleted if obviously noisy.

This procedure introduces distortion because it reduces peaks and flattens slopes as illustrated in Fig. 115 in Section 6.

The extent to which this occurs depends on the number of points in the comb in relation to the number in a wave of the signal. Too few points does not smooth sufficiently, and too many produces distortion. Experimentation with the number of points must be used.

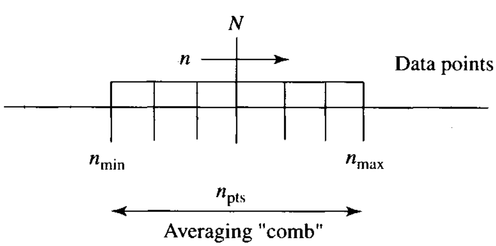


FIGURE 179 Illustration of data points used in averaging

B3. Detrending (Offset and Drift)

When the baseline (zero value) of a waveform does not pass through the average of all values in the waveform, it is said to have an offset, as illustrated in Fig. 112 of Section 6. This is equivalent to a zero frequency component, which is large and inconvenient in spectral analysis.

When the baseline has a gradual trend upward or downward, it is said to have drift, which causes an inconveniently large low frequency component. Both these effects are likely to represent some effect outside the test being made, but this point must be considered carefully.

The method of least squares is applied to all points of a waveform to determine the best fit to a straight line,

$$S_{\text{fit}} = a + bt,$$

where a is the axis offset and b is the rate of drift. The deviation from this line, $\delta S = S(t) - S_{\text{fit}}(t)$, of each digital data point of the waveform is squared, to allow equal contributions of positive and negative deviation, and summed over all points of the waveform, $\Delta = \sum \delta^2$. The total deviation is minimized through variations of the coefficients a and b:

$$\partial\Delta/\partial a = \partial\Delta/\partial b = 0,$$

where

$$\Delta = \sum_{n=1}^{N} (S_n - a - bt_n)^2$$

and n denotes the nth point in the waveform.

This leads to two equations for a and b with the solutions

$$a = \left[\left(\sum S_n\right) - b\left(\sum t_n\right)\right]/N = \mu_S - b\mu_t$$

$$b = \left(\sum S_n t_n - N\mu_S\mu_t\right) / \left(\sum t_n^2 - N\mu_t^2\right)$$

where $\mu_S = \left(\sum S_n\right)/N$ and $\mu_t = \left(\sum t_n\right)/N$ are the means of the waveform and time points.

Four summations are made, $\sum S_n$, $\sum t_n$, $\sum S_n t_n$, and $\sum t_n^2$, followed by application of the formulas, first for b and then for a. The waveform points are then corrected to eliminate offset and drift by using the formula

$$S_{n.\text{corr}} = S_n - a - bt_n.$$

It may be desirable to precondition each waveform before averaging if the repetition rate is low or the electrical and environmental conditions change fast.

B4. Convolution

A convolution is an integral over time of a shifting product of two time functions. The response of a system to an excitation can usually be expressed as a convolution of the fundamental system response, $f(t)$, and the excitation, $F(t)$:

$$R(t) = F(0)f(t) + \int_0^t [dF(\tau)/dt]f(t-\tau)d\tau$$
$$= F(t)f(0) + \int_0^t F(t)[df/dt]_{(t-\tau)}d\tau.$$

Convolution is used when the response function of a system is known, such as a filter. It is then referred to as a transfer function.

The significance of the convolution can be demonstrated by considering the response of a mechanical damped mass-spring system, or of an electrical inductance–resistance–capacitance (LRC) circuit, illustrated in Fig. 180.

The response $u(t)$ of such simple systems, called single degree-of-freedom (SDOF) systems is governed by a differential equation of the type

$$Ku + \mu du/dt + M d^2u/dt^2 = F(t),$$

which, for a unit step input, has the solution, called the indicial response:

$$h(t) = \{1 - e^{-\xi\omega_n t}\cos[(1-\xi^2)^{1/2}\omega_n t - \phi]\}/K,$$

where $\xi = c/c_c$, $c_c = 2(km)^{1/2}$ is the critical damping, $\omega_n = (K/M)^{1/2}$ is the natural frequency, and $\phi = \pi/2 - \text{asin}(1-\xi^2)^{1/2}$.

To develop the solution for an arbitrary input forcing function $F(t)$, consider the response as a series of steps, each excited by a portion of the forcing function, as illustrated in Fig. 131 in Section 6.

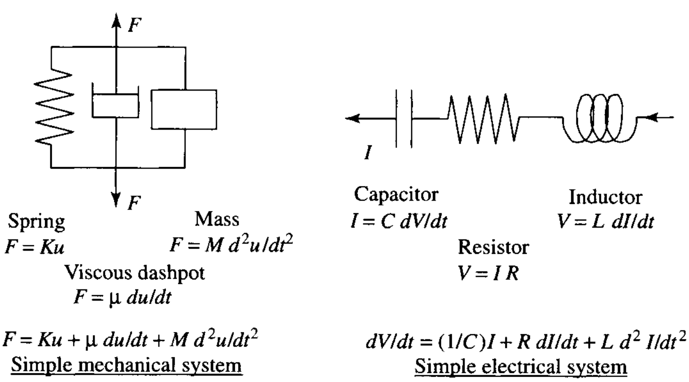


FIGURE 180 Simple mechanical and electrical systems for illustrating response

The solution is given by either of the convolution integrals

$$u(t) = F(0)h(t) + \int_0^t [dF/dt]_{(\tau)}h(t-\tau)d\tau$$

$$= F(t)h(0) + \int_0^t F(t)[dh/dt]_{(t-\tau)}d\tau$$

It is usually convenient to evaluate convolution integrals through Fourier transforms, as discussed in Section 6.

This Page Intentionally Left Blank

REFERENCES

Achenbach, J. D., "Wave Propagation in Elastic Solids." North Holland Press, Amsterdam, 1973.

Aki, K., and Richards, P. G., "Quantitative Seismology, Theory and Methods." W. H. Freeman and Co., San Francisco, 1980.

Doyle, J. F., "Wave Propagation in Structures, an FFT-based Spectral Analysis Methodology." Springer-Verlag, New York, 1989.

Ensminger, D., "Ultrasonics: Fundamentals, Technology, and Applications," 2nd ed. M. Dekker, New York, 1998.

Ishimaru, A., "Wave Propagation and Scattering in Random Media." Academic Press, New York, 1978.

Krautkramer, J., and Krautkramer, H., "Ultrasonic Testing of Materials," 4th ed. Springer Verlag, Berlin, 1990.

Kutruff, H., "Ultrasonics: Fundamentals and Applications." Elsevier, Amsterdam, 1991.

Love, A. E. H., "A Treatise on the Mathematical Theory of Elasticity," 4th ed. Dover Publications, New York, 1944.

Mason, W. P., "Physical Acoustics and the Properties of Solids," van Nostrand, Princeton, NJ, 1958.

Musgrave, M. J. P., On the propagation of elastic waves in aeolotropic media, I—General principles." *Proc. Roy. Soc.* **A226**, 339 (1954).

Povey, M. J. W., "Ultrasonic Techniques for Fluid Characterization." Academic Press, San Diego, 1997a.

Povey, M. J. W., and Mason, T. G., "Ultrasonics in Food Processing." Chapman and Hall, London, 1997b.

Sahay, S. K., and Kline, R. A., Three dimensional representation surfaces for anisotropic media, *Rev. Prog. QNDE* **11**, 37 (1991).

Sokolnikoff, I. S., and Specht, R. D., "Mathematical Theory of Elasticity." McGraw-Hill, New York, 1946.

Thurston, R. N., and Pierce, A. D., "Ultrasonic Measurement Methods," Vol XIX, Physical Acoustics Academic Press, London 1990.

Wells, P. N. T., "Physical Principles of Ultrasound Diagnosis," Academic Press (1969).

Wills, A. P., "Vector Analysis with an Introduction to Tensor Analysis." Dover Publications, New York, 1958.

INDEX

This Page Intentionally Left Blank