Part I

Waves:

The Foundation of Seismology

1

General Properties of Waves

Reflection seismology is the science of examining the earth's interior through the analysis of mechanical waves. Seismic rays are bent, reflected, refracted, diffracted, and scattered. The emitted signal is weakened by these effects as well as geometric spreading and attenuation.

The waves are generated by controlled sources and travel through fluids, solids, and porous solids. In this chapter, we consider those properties of waves, which do not depend on the kind of material supporting the wave propagation. For example, Snell's law is a general property of wave motion but differs in detail between reflection of acoustic and elastic waves. When considering a general property like Snell's law, in this chapter, we will discuss the scalar wave case.

When considered from this classification scheme (for that is what it is), all acoustic wave phenomena are also present in elastic wave propagation as well as a relatively small number of new, uniquely elastic effects. Similarly, the passage from elastic to poroelastic introduces just a few new ideas and observable effects.

This chapter emphasizes conceptual discussion and visualization of these concepts. Mathematics enters where necessary and useful, but not gratuitously.

1.1 Mechanical waves

Mechanical wave motion involves a disturbance propagating through a compressible or elastic medium thereby transferring energy from one region to another without transfer of mass [107]. The energy transport is via elastic properties of the material, whether elastic in the sense of fluids characterized by mass density and bulk modulus, or solids described by mass density and stiffness parameters. Such parameters form a numerical model of the material and a primary goal of reflection seismology is to estimate these parameters from observed seismic data.

When considering a fluid or solid medium for wave propagation problems, we distinguish between the interior (or body) and the free surface. By free surface, we mean a boundary between the material and a vacuum, or at least a region with relatively small elastic parameters and/or density. For example, in many problems, the earth surface behaves very much like a free surface because bulk modulus and density of the lower atmosphere are orders of magnitude smaller than these parameters are in surface rocks.

Even the ocean/air interface is a good approximation to a free surface. In general, the body can be bounded on all sides by free surfaces, as would be the case in studying seismic waves in an entire planet. For the problems we will encounter, there will be one free surface representing the surface of the solid earth or ocean.

A broad distinction of mechanical waves is made between body and surface waves. Body waves are those that can propagate through the interior of a fluid or solid, while surface waves exist only in the vicinity of the free surface.

Sound is the only kind of mechanical body wave in a fluid. For an isotropic elastic solid, two kinds are possible, termed P and S. In an anisotropic solid, there are three possible mechanical waves. The technically correct designations are quasi P-wave (qP), quasi S-wave one (qS1), and quasi S-wave two (qS2). In practice, however, people quickly tire of saying “quasi” all the time and these waves are called P, S1, and S2. The shear waves travel at different speeds and are therefore also referred to as fast and slow S-waves, with S1 being the faster of the two. Finally, one additional kind of body wave is theoretically possible when the solid is porous. This new wave is called a Type II P-wave and only exists at frequencies well above surface seismic band of 10-100 Hz. However, this wave can play a role in sonic logs operating at 10 to 15 thousand Hertz, and especially in core measurements (one megahertz or higher).

1.2 Particle motion

As a mechanical wave progresses, it is said to be propagating in a particular direction (the direction of energy transport). This wave causes the material to vibrate and the orientation of this vibration is called the direction of particle motion. Of fundamental importance are the relative directions of propagation and particle motion, (see Fig. 1.1).

picture 01-01 scaled 900

Fig. 1.1 Mechanical waves are distinguished by the vibration patterns they cause in the medium. By definition, a longitudinal wave (upper right) has particle motion parallel to the direction of propagation. In transverse wave motion (lower right), the particles vibrate perpendicular to the propagation direction.

In general, there are two kinds of pure waves—longitudinal and transverse. A longitudinal wave is one in which the particle motion is parallel to the propagation direction. Examples include sound and seismic P-waves. In a transverse wave, the particle motion is perpendicular to the propagation direction. A seismic S-wave is of this type.

Waves of a mixed type are possible, in which the particle motion is some combination of parallel and perpendicular vibration relative to the direction of propagation. Examples include surface gravity waves on water (prograde ellipse), Rayleigh wave (retrograde ellipse shallow to prograde ellipse deep), anisotropic P-waves, and near field elastic waves.

1.3 Polarization

Polarization is not possible for longitudinal wave motion but occurs in a transverse wave when particle motion is confined to one direction. In Figure 1.2 we see a 3D view of a grid of points that lie in a plane (left plots). A transverse wave propagating in the x-direction can have particle motion in the y-direction (upper right) or the z-direction (lower right). These are pure mode transverse waves; a general transverse wave can have particle motion in any direction perpendicular to the propagation direction.

picture 01-02 scaled 900

Fig. 1.2 Polarization is a property of transverse waves. A horizontally polarized wave (upper right) has particle motion confined to the horizontal plane. Vertical polarization occurs when the particles vibrate only in the vertical direction (lower right).

For seismic shear waves, it is common to distinguish a shear wave with horizontal polarization (SH) from one with vertical polarization (SV). Another example involves fast (S1) and slow (S2) shear waves associated with propagation in anisotropic rock.

1.4 Elastic properties

Whether the material under consideration is a fluid, solid, or porous solid, it is characterized by a certain number of elastic parameters. Mass density is a parameter universal to all media and is not generally counted among the elastic parameters. In a slight abuse of terminology, we use the term, *elastic parameter*, to describe the bulk modulus of a fluid even though a fluid is not an elastic solid. At any rate, the general scheme is enumerated this way: a fluid is characterized by one elastic parameter, an isotropic solid by two, and the simplest useful anisotropic solid by four. This kind of anisotropy is the representation most often used in petroleum seismic work, but a generally anisotropic material could have up to 21 elastic parameters.

1.5 Wavespeed

As a waveform travels through a solid or fluid, it is possible to track a particular feature and associate a speed with its movement. This is the phase velocity of the wave, and, as with any velocity, it has both magnitude and direction (that is to say, it is a vector). However, in seismic studies, the velocity referred to is seldom the vector velocity of a specific traveling wave, but the scalar wavespeed, which is a material parameter associated with a point in space. Clear understanding of this distinction avoids a vast amount of confusion. In accordance with standard practice, throughout this, work we will use the terms *velocity* and *wavespeed* interchangeably.

In general terms, a wave propagating through a fluid or solid has a distinct velocity that depends on four quantities [77],

$$v(\underline{x}, \hat{n}, \_, A) \qquad (1)$$

where $\underline{x}$ is the observation point in 3D space, $\hat{n}$ is a unit vector pointing in the direction of wave travel, _ is the wavelength, and A is the amplitude of the wave.

If wavespeed depends on the observation point, the material is said to be heterogenous. This is the common case with seismic body waves since rock properties vary from point to point in the earth. For an acoustic wave example, consider the speed of sound in a lecture hall. Near a warm object, the sound speed is relatively high, while it is relatively low near a cool object.

One can imagine a 3D model of the sound speed in such a room on a regular cubic grid—hot colors (red, yellow) indicating high velocity and cool colors (blue, green) low ones. In much the same way, 3D velocity models of the earth are constructed for use in modeling and imaging of seismic data. But an earth velocity model might be a cube several kilometers on a side, (see Fig. 1.3). As physical properties in the earth change the speed of seismic waves also vary. First order effects are increase of temperature and pressure with depth, and other factors include pore fluids, porosity, mineralogy, and fractures.

If wavespeed depends on direction of travel, the medium is said to be anisotropic. Our acoustic example is no help here, since sound waves are isotropic, traveling at the same speed in any direction. This is also very nearly true for seismic waves in thick, uniform rock units. But anisotropy affects both P- and S-waves and is common in sedimentary rocks due to fine layering, fracturing, regional stress fields, and other causes. In layered rocks such as shale, it is not unusual for a horizontally traveling P-wave to go 20% faster than a vertical one. In practice P-wave anisotropy is often overshadowed by other velocity effects and therefore difficult to discern on its own.

In a basin under lateral tectonic stress, wavespeed will vary with compass direction of travel, and fractured rocks give rise to the interesting phenomena of shear wave splitting. This is the seismic equivalent of the splitting of light by the mineral calcite to create a double image of objects viewed through it. Shear wave anisotropic effects, while interesting, are not as important in petroleum seismology as P-wave effects simply because the great majority of work is done with P-wave data.

If the wavespeed depends on wavelength, the medium is said to be dispersive. This is commonly the case for seismic surface waves (Rayleigh wave) when the near surface is layered. Seismic body waves, both *P* and *S*, are only slightly dispersive until pressed into higher frequency bands associated with sonic logging or laboratory rock physics measurements. For dispersive waves, it is necessary to distinguish the speed at which a recognizable feature propagates (phase velocity) from the speed of energy transport (group velocity). By necessity, dispersion is associated with absorption, which is the conversion of wave energy into heat. Absorption is one of an entire class of mechanisms that cause seismic amplitude to decay with distance traveled, and it is challenging to isolate absorption from the others.

Finally, if the wavespeed depends on the amplitude of the wave, the medium, it is said to be nonlinear. This is not a consideration for seismic waves in the 10–100 Hz band typical of petroleum seismic data. However, as with dispersion, there are measurable effects at sonic frequencies and beyond [46].

In summary, for the vast majority of seismic work the P-wave velocity can be considered a function of position $v(\underline{x}) = v(x,y,z)$ intertwined with anisotropy effects of up to 20%. Figure 1.3 shows a 2D velocity model with gray levels representing isotropic P-wavespeed. This particular model was created to study imaging algorithms, but models of similar complexity are derived from real data.

picture 01-03 scaled 650

Fig. 1.3 This 2D velocity model represents P-Wavespeed in the Earth. Such models are used in seismic imaging and modeling. Light shades represent low velocity and dark shades high velocity. (Modified from [58]).

1.6 Display of wavefield data

When a receiver records data, a time series is created that consists of some number of consecutive measured values separated in time by a constant time sample rate. Whatever the actual measured quantity (particle velocity, pressure, etc.), it is usually referred to as amplitude data. A time series recorded by any kind of seismic receiver is called a trace and will generally consist of a few thousand samples with a time sample rate 1, 2, or 4 ms (1 millisecond = 0.001 sec.). Many such traces taken together form a gather or section, which represents the measured wavefield or image created from it. The display of seismic wavefield data is generally in one of two forms, an image or a wiggle trace display. In the image display, gray levels or color values are associated with amplitudes.

A data panel consisting of 1000 traces with 500 samples per trace is shown in Figure 1.4. Even with 500,000 (500K) data samples, we clearly see seismic reflection events associated with the seafloor, faults, bed terminations, unconformity surfaces, and other geologic features. A wiggle trace display of this data at this scale would be unusable due to the data density (about 20K samples per square inch).

picture 01-04 scaled 850

Fig. 1.4 Grayscale image of a data panel consisting of 1000 traces and 500 time samples per trace. Trace spacing is 15 m and the time sample rate is 4 ms. A sedimentary interval is highlighted to show thickening. This is offshore data from Southeast Asia.

Figure 1.5 shows cascading levels of zoom for this same data panel. Note the distance bar on each plot for scale. Working down from the top, each data panel contains 500K, 100K, 50K, 25K, 5K, and 500 data samples. Even though all panels are physically the same size, the data density is different in each. When working with a lot of data (very high sample density), image displays are essential. For close work, the wiggle trace display is preferred as it allows the interpreter to analyze individual waveforms.

picture 01-05 scaled 850

Fig. 1.5 The top image contains the same data as Figure 1.4 and each lower panel is a zoom view of the one up and to the right of it. When the zoom level gets in close enough a wiggle plot is useful to show waveform details. Note horizontal scale bar on each plot.

1.7 Waveform

Think of a sound wave moving from left to right and a magic camera that can take a picture of the pressure field. If we take a snapshot, the wave is frozen at a particular instant in time. The pressure is now a function of *x* and will vary between high and low as indicated by shades of gray in Figure 1.6.

picture 01-06 scaled 900

Fig. 1.6 A sound wave moving to the right at fixed time.

The variation may also be plotted as the line graph shown in Figure 1.7. Since the horizonal axis in this plot is distance, the peak-to-peak interval is the wavelength. The deflection of the wave from zero is the amplitude representing an observed physical quantity, in this case pressure. The distance between any two recurring features (peak, trough, downslope zero crossing[1], etc.) is the wavelength.

picture 01-07 scaled 850

Fig. 1.7 Line graph of the sound wave in Figure 1.6.

Think now of a receiver (i.e., a microphone) fixed in space and measuring the passage of the same sound wave as a function of time. This is the typical case in acoustics and seismology. Figure 1.8 is the time series or trace for this case. Since the horizontal axis is now a time

[1]This distinction is necessary because the distance between successive zero crossings is actually the half wavelength.

coordinate the peak-to-peak interval is the period of the wave, *T*, given in seconds. The frequency associated with this period is

$$f = 1 / T \ , \qquad (2)$$

where the units of frequency are Hertz whose dimensions are 1/sec.

picture 01-08 scaled 850

Fig. 1.8 Time trace measured by a fixed receiver.

These two views of the sound wave (fixed time and fixed space) are connected by the general relationship

$$v = f\_, \ (3)$$

where *v* is wavespeed, *f* is frequency, and _ is wavelength. Clearly, we can use this equation to find $f = v/\_$ or $\_ = v/f$ as the situation requires.

The waves shown in Figures 1.7 and 1.8 are termed monochromatic, because they contain only one frequency or wavelength. In seismic data, we deal with a wavelet, or wave pulse, which contains many frequencies (Fig. 1.9) and is time-limited, meaning that it exists only for a finite time instead of oscillating forever like a sine wave. Since each frequency has its own period, we speak of the entire waveform as having a dominant period. The dominant frequency and dominant wavelength are then given by $f_{dom}=1/T_{dom}$ and $\_{dom}=v/f_{dom}$, respectively. The particular waveform shown in Figure 1.9 is a sinc wavelet whose mathematical expression is

$$sinc(t,f)=\frac{sin(2pft)}{2pft} \ , \quad (4)$$

where *f* is the dominant frequency.

picture 01-09 scaled 850

Fig. 1.9 A sinc wavelet contains many frequencies and therefore many periods. It is characterized by its dominant period.

If f_1 and f_2 are the minimum and maximum signal frequencies in a general wavelet, the dominant frequency is the average given by

$$f_{dom}=(f_1+f_2)/2 \ . \qquad (5)$$

Another common definition of dominant frequency is the peak of the frequency spectrum of a window in the seismic data. See appendix A for a discussion of the Fourier transform and associated topics such as the amplitude and phase spectrum. The bandwidth of the wavelet can be described in Hz or octaves

$$Hz \quad : \quad f_2\text{-}f_1$$

$$octaves \quad : \quad \ln_2(f_2/f_1) = \ln(f_2/f_1) \,/\, \ln 2 \ . \quad (6)$$

The octave range is a measure of how dominant the central peak is relative to sidelobes. In order to achieve a strongly peaked waveform, the bandwidth should be at least 2.5 octaves. An easy way to estimate bandwidth in octaves comes from the fact that an octave is the doubling of frequency. For example, consider a 10–90 Hz wavelet. To estimate the octave bandwidth, begin at the lowest frequency and start doubling until the highest frequency is met or exceeded—10, 20, 40, 80, 160. This wavelet has a bandwidth of a little over three octaves as detailed calculation would confirm.

It is instructive to put this in the context of human hearing. The speed of sound in air is about 335 m/s and the range of human hearing is 20–20,000 Hz. This is the range for a healthy, young adult. Children can hear up to 25,000 Hz and most of the energy in speech is in the 300–3000 Hz range.

The octave bandwidth of normal adult hearing is

$$\frac{\ln(20\,000/20)}{\ln 2}=9.97 \ , \quad (7)$$

which is much better than the 2 or 3 we get in typical seismic data. This impressive range allows us to distinguish chirps arriving almost simultaneously and determine the source direction.

The calculations can be done by the reader, but the approximate range of sound wavelengths we can hear range from 17 m down to 2 cm. So what is it about the human body as a sound receiver that has a length scale of 17 m? These low frequency (long wavelength) sounds feel more like a whole-body vibration or buzzing than a true sound. Examples are boom cars (ultrabase stereo), power transformers (f=120 Hz; _=3 m), and the buzz of a bee (f=20 Hz; _ =17 m). These sounds are received by the entire body, which on average is 2 m tall. So the lowest frequency that registers with you is one whose wavelength is about eight times your greatest dimension.

Now think about an elephant or a whale and the very low frequencies they must be able to hear. In fact, recent research in South Africa has shown that elephants communicate across great distances using sound that is far below the range of human hearing. Apparently that is the way they all know when to show up at the watering hole.

And what about the short wavelengths? What is there about our hearing apparatus which has a length scale of two centimeters? This is about the size of the small bones of the human inner ear. Small animals can hear much higher frequencies than we do. To cite one example, the big brown bat [168] responds to frequencies up to 100,000 Hz.

1.8 Impulse response

The simplest kind of source that can initiate wave propagation acts at one point in space and one point in time. This source is called a spike or delta function, _( ), and is discussed in appendix A. Such an impulsive point source located in a 3D medium at a point in space $\underline{x}_0=(x_0,y_0,z_0)$ and acting at time t_0, symbolically _($\underline{x}$-$\underline{x}_0$) _(t-t_0), generates an outgoing wavefield that is called the impulse response of the medium. The impulse response may be analytic, if the medium is simple, and the equations of motion are mathematically solved for the wavefield, or numerical, if the wavefield is simulated by computer.

A numerical impulse response for 2D acoustic waves is shown in Figure 1.10. The velocity model appears in each panel as layers in varying shades of gray, the darker layer colors being higher velocity. The source acted at time zero (not shown) at the precise center of the model $(x,z)=(1250\ m, 1250\ m)$. This source initiates a wavefield that is shown as a transparent overlay at progressively later times in each panel. So, the wavefield represents the time evolution of the impulse response. To be precise, the source was not actually a pure spike in this case but a waveform much like the one in Figure 1.9. However, considering the scale of the model and short duration of the source pulse, we can consider it a good approximation to the impulse response.

The impulse response is a useful concept for modeling, processing, and imaging. In each case, it allows us to isolate the effect of a very simple stimulus before dealing with complicated, real world issues.

picture 01-10 scaled 700

Fig. 1.10 A 2D impulse response. A point source at the center of the model gives rise to a wavefield that propagates outward. The wavefield is shown at four successively later times and the layered velocity model is seen as a transparent overlay with darker shades indicating higher velocity.

1.9 Reciprocity

As the wavefield in Figure 1.10 moves across any point in the model, we can imagine a receiver located there, which records its passage. We have a point source generating waves that are recorded by a point receiver. A powerful idea related to this situation is the reciprocity principle that in this case states:

For any velocity model and source wavelet, the seismic trace recorded at point (x_1,y_1) *due to a point source acting at* (x_2,y_2) *is identical to the trace recorded at point* (x_2,y_2) *due to a point source acting at* (x_1,y_1).

In other words, if we swap the location of the source and receiver, the recorded trace does not change.

Reciprocity is not just a property of 2D acoustic waves, but extends to much more complicated situations [200]. The usefulness of reciprocity is that it forms the basis of many seismic imaging and inversion methods [5], including surface related marine multiple suppression algorithms [71]. In difficult situations, such as extended sources shooting into multicomponent receivers through anisotropic rocks [59], it is also the source of recurring controversy.

1.10 Source and receiver directivity

The point source discussed above is an idealization. If such a source did exist in a fluid, then waves generated by it would spread out with uniform strength in all directions. However, any real source has a finite size and does not radiate equally in all directions. This agrees with our common experience concerning sound. When speaking to someone, you face them, partially out of habit and courtesy, but mainly because it is easier for that person to hear what you are saying. Someone behind you might be closer, yet not hear what is being said. From a physics perspective, your voice exhibits a radiation pattern; it is emitting more power in some directions than in others.

Figure 1.11 illustrates this point with 2D waves. The fact is that any physical source used in petroleum seismology, from a marine air gun to a land vibrator, will exhibit 3D source directivity. You can imagine that a source radiation pattern will affect all subsequent evolution of the wavefield. If possible, this directivity will be processed out of the data to avoid interference with other effects that carry more useful information.

picture 01-11 scaled 800

Fig. 1.11 Snapshots of two wavefields. The source in each case was at the center of the grid. (A) Waves from a symmetric point source are emitted identically in all directions.

(B) Actual sources behave more like this simulation. The power and emitted waveform vary with direction from the source forming a radiation pattern.

Directivity is not just an issue with the source but also the receiver. Our analogy with sound still works here, since the ear is a very directional receiver. What we measure generally depends on the direction the wave is traveling relative to the directivity pattern of the receiver. This effect needs to be understood and removed to preserve accurate amplitude information.

It should be noted that source and receiver directivity are fundamentally different than anisotropy. The effects we have described exist in isotropic media, and more complicated versions would exist in anisotropic media.

1.11 Wavefront and rays

Just as we see the wavefront moving away from a stone tossed into quiet water, we see the wavefront when looking at a wavefield snapshot such as Figure 1.10. Another viewpoint is that the wavefield is characterized by rays, which connect the source point with the wavefront. These are equivalent ways of looking at the situation, and which one we choose depends on the problem being solved.

In the simplest case of a point source in a constant velocity medium, the wavefront is circular (2D) or spherical (3D) and all rays are straight. The rays radiate from the source to land perpendicular on the wavefront as shown in Figure 1.12A.

picture 01-12 scaled 700

Fig. 1.12 Rays (arrowhead lines) connect the source point (center circle) with the wavefront (heavy line). (A) In constant velocity media the rays are straight and the wavefront is a circle (2D) or sphere (3D). (B) When velocity varies in the medium, rays bend and wavefronts deform.

As the wavefield evolves with time, the rays lengthen and the wavefront moves progressively farther from the source as seen in Figure 1.10. But you will notice that the outermost wavefront in the later snapshots are clearly not circular. When velocity varies from one point to the next, it has the effect of bending the rays and deforming the wavefront. Specifically, the wavefront is pulled into high velocity zones and appears to accelerate away from the source in these regions, while rays bend away from high velocity zones and toward low velocity regions.

Figure 1.12B is a cartoon of this situation. Computing the precise raypaths and wavefront geometry for complex velocity models is a significant problem—one either uses ray theory to find the rays or wave equation methods to find the wavefronts. Figure 1.13 is the rather complicated wavefront of Figure 1.10 (Time 4). It is marked up to show velocity model layers (L1, L2,...), interfaces (I1, I2,...), and some wavefield events, including I1 downgoing reflection, wave transmitted through I1, I2 reflection, I4 reflection, direct wave in L2, and I3 reflection.

While Figure1.13 includes wavefield events, such as reflections, for really complicated velocity models just computing direct arrival raypaths is a difficult problem. Figure 1.14 shows first arrival rays [139] for one shot location in the Marmousi model, [192] which has been used extensively to test and develop seismic imaging algorithms. The rays in this case are severely bent, cross each other, and leave holes where they have twisted around high velocity regions. Time along each ray increases away from the source and reflection events are not included. The wavefronts could be constructed by connecting the rays at some fixed time. You can imagine that the wavefronts would become extremely distorted far from the source. Unless the medium is anisotropic, rays will always be perpendicular to wavefronts.

picture 01-13 scaled 800

Fig. 1.13 Detail of the Time 4 snapshot in Figure 1.10. Velocity model layers and interfaces are labeled as well as some wavefield events.

picture 01-14 scaled 800

Fig. 1.14 (A) The Marmousi Velocity Model. (B) Direct arrival ray fan for a source located at 6000 m. Note intense folding of the ray field in response to strong velocity variations in the model. (Modified from [139]).

1.12 Huygens' principle

An important concept in wave propagation of any kind was given by Christiann Huygens in 1678 and published [86] in 1690. It states [21]

> *...every point of a wavefront may be considered as a center of a secondary disturbance which gives rise to spherical waves and the wavefront at any later instant may be regarded as the envelope of those wavelets.*

The precise meaning and mathematical details of this principle have occupied minds for centuries [9]. Born and Wolf [21] point out that, taken literally, the principle is incorrect because it would lead immediately to a secondary wavefront propagating toward the source, as well as the one we observe moving farther away as illustrated in Figure 1.15. The principle is saved, if the secondary sources have directivity patterns of such a nature that only the physical wave is generated by their summation [21]. This extended form is called the Huygens-Fresnel principle.

While this fundamental principle underlies much of seismic wave propagation, its direct use in solving problems is rare [159].

picture 01-15 scaled 1000

Fig. 1.15 Huygens' Principle. Each point on the upgoing wavefront W1 at time 1 is considered to be the center of a secondary spherical source (a circle here in 2D). After a small interval of time, each has a radius determined by the velocity at its center, and the envelope of these is the new wavefront W2 at time 2. The false wavefront A heading in the wrong direction is eliminated in an extended form of the principle (termed Huygens-Fresnel), where the amplitude emitted by each secondary source is directional.

1.13 Fermat's principle

We have seen that in variable velocity media wavefronts distort and rays bend. But why is this?

Think about two points in a complicated velocity model and imagine a source at one point shooting a ray into a receiver at the other point. Clearly there are an infinite number of ways to connect these points—a straight line, a parabolic curve, a general curve, etc. Along any path we care to choose, the traveltime can be calculated for each small step along that path by

$$dt=\frac{ds}{v(x,y,z)} \quad ,(8)$$

where dt is the traveltime over the short interval ds, and $v(x,y,z)$ is the velocity in that small interval. The total traveltime along the path from source point (p_1) to receiver point (p_2) will therefore be

$$t=\int_{p_1}^{p_2}\frac{ds}{v(x,y,z)} \quad . \qquad (9)$$

The problem is that for every path we choose to connect the fixed points (p_1,p_2), we get a different traveltime. This is where Fermat's principle comes in. In its original 1662 form, [198] it says, "The actual path between two points taken by a beam of light is the one which is traversed in the least time." This is also called the principle of least time. It leads directly to Snell's law and ray equations [18] for determining raypath geometry in complex media.

It is comforting to know that when the velocity is constant, and the ray is therefore straight (say along the x direction), the equation above gives $t=x/v$ as it should since time equals distance divided by speed.

1.14 Snell's law

In 1621, Willebrord Snell discovered the mathematical rule by which light bends as it passes from one material into another. Snell's law in some form is valid for all waves, but we will develop the idea here in the context of sound waves as this is the simplest case.

We consider two points, (p_1,p_2), in a model with two velocities separated by a horizontal interface, Figure 1.16. In drawing (A) one path is shown connecting the points, but an infinite number of such paths could be drawn. A convenient way to distinguish any given path is by the horizontal distance, x_1, from p_1 to the point where it strikes the interface. For any path, the traveltime between (p_1,p_2) will be the sum of the distances traveled in each layer divided by the velocity of that layer,

$$t(x_1)=\frac{L_1(x_1)}{v_1}+\frac{L_2(x_1)}{v_2} \quad . \qquad (10)$$

Armed with Fermat's principle, we know the actual path taken will be the one which takes the least time. As the distance x_1 takes on different values, the intercept point moves and the traveltime changes. Some of these times are longer and some shorter, but there is a unique time that is the smallest of all. This is a classic minimization problem in one variable and the solution is the value of x_1 satisfying

$$\frac{\partial t(x_1)}{\partial x_1}=0 \quad . \qquad (11)$$

The result of solving this problem is a relationship between the incidence angle, $\_{1}$, and the transmission angle, $\_{2}$,

$$\frac{sin\ q_1}{v_1}=\frac{sin\ q_2}{v_2} \quad , \qquad (12)$$

as illustrated in Figure 1.16B. The vertical dashed line is termed the normal to the interface (i.e., it is perpendicular to the interface).

It is interesting to note that Aristotle (384-322 BCE) understood this law to be

$$\frac{q_1}{v_1}=\frac{q_2}{v_2} \quad , \qquad (13)$$

which agrees with Snell's law in the small angle limit as $\_{1} \_ \_{0}$.

When $v_1 > v_2$ Snell's law says the transmission angle will be less than the incidence angle and the transmitted ray will bend toward the normal, (see Fig. 1.16D). From this case, we know that rays will bend into low velocity regions. If $v1 = v_2$, it follows that $\_{}_1 = \_{}_2$, and the ray does not bend, (see Fig. 1.16C). Not very exciting, but true. Finally, if the velocity increases across the interface, $v_1 < v_2$, the transmission angle is greater than the incidence angle, and the transmitted wave bends away from the normal, (see Fig. 1.16B). It follows that rays bend away from high velocity regions.

picture 01-16 scaled 550

Fig. 1.16 Snell's Law. (A) Geometry of a raypath connecting two points across an interface separating regions of different wavespeed. (B) When $v_1 < v_2$, the transmitted wave is bent away from the normal. (C) When $v_1 = v_2$, the ray does not bend. (D) For $v_1 > v_2$, the ray bends toward the normal.

If the interface is not horizontal, Snell's law still works because it is all relative to the normal to the interface. In fact, the model could be flipped upside down so that the incident wave is coming from below. In any case, we understand medium 1 in the formula refers to the incident side of the interface, and medium 2 is the transmission side. Also, Snell's law clearly extends to the 3D case, where the normal is associated with a plane separating v_1 and v_2. If the interface is curved, the tangent plane is used. The geometry gets more complicated, but there is nothing fundamentally new. All of these facts conspire to give the complicated raypath behavior seen in large-scale models such as Figure 1.14.

Figure 1.17 shows both reflected and transmitted rays which, taken together, are termed scattered waves. Since both the incident and reflected rays travel with speed v_1, Snell's law confirms the angle of incidence equals the angle of reflection. In the case we have been considering, it is assumed that either only one wave type exists, or if others exist, the incident wave does not excite them. More general cases will be dealt with later.

picture 01-17 scaled 800

Fig. 1.17 Acoustic reflection and transmission geometry. The reflected and transmitted waves are collectively called scattered waves.

1.15 Critical angle

The example in Figure 1.16B shows the transmitted ray bending away from the normal when velocity increases across the interface. At some angle of incidence, this will cause the transmitted ray to run parallel to the interface, and this incidence angle is called the critical angle, $\_c$. If the transmission angle is 90 degrees, then Snell's law takes the form

$$\frac{sin\ q_c}{v_1} = \frac{1}{v_2} \quad , \qquad (14)$$

which can be solved for the critical angle

$$\_c = sin^{-1}\left(\frac{v_1}{v_2}\right) \quad ; \; v_1 < v_2 \quad . \qquad (15)$$

To cite one example, sound waves traveling through air (v_1 =335 m/s) and passing into water (v_2 =1500 m/s) have a critical angle of about 13 degrees. This means that only sound striking the air-water interface at angles less than 13 degrees will be transmitted into the water.

1.16 Dimensional effects

1.16.1 Waveform. The real world is 3D, but seismic modeling is often done in 2D to save time and effort. The wavefield snapshots in Figure 1.10, for example, are 2D results. But a 2D point source is actually a 3D line source. In other words, the physics of the situation does not see a single source sitting in a 2D world, but a line of point sources stretching away into the third dimension.

Consider what happens when a 3D impulsive point source explodes in a constant velocity world. It generates a wavefront that spreads out in a spherical shell, and behind the wavefront it quickly goes still, because the source is no longer emitting any waves. This is not the case for a 2D impulsive point source. Behind the wavefront, there is always energy arriving from somewhere along the line source [3]. In effect, we need to add up the contributions from the infinite number of point sources along the *y*-axis to find the 2D result.

As a consequence, 3D wave modeling will propagate the actual waveform as emitted by the source, while 2D modeling will propagate the y-integral of the source waveform. The effect of this on an impulse would be the appearance of the spike stretching back toward the source in what is called a square-root-t tail. Figure 1.18 illustrates the significant difference this can represent for a more realistic emitted waveform. When very detailed modeling is used for stratigraphic or rock property investigations, this issue can sneak in and lead to confusion. This effect is sometimes corrected under the name of 2.5D wave equation modeling [110, 112].

picture 01-18 scaled 900

Fig. 1.18 3D Seismic modeling propagates the actual source waveform while 2D modeling propagates the Y-integral of that waveform. This can be confusing when comparing waveforms between modeled data and field data.

1.16.2 Energy density and geometric spreading. A property of any wave is that it gets weaker with distance traveled. There are many reasons for this. Most are related to material properties, but geometric spreading is a general effect that is different for waves propagating in a 2D or a 3D world. Of course, we live and work in a 3D world, but waves associated with surfaces are often confined to those surfaces and thus spread out in 2D. So the 2D case is relevant to us.

Consider an impulsive point source of waves in an infinite 3D medium. The source injects energy into the material for a brief period of time (say a few ms) then stops. This generates an outgoing wavefront that, if velocity is constant, will be spherical. At some short time after the source acted, all of the energy is concentrated on the surface of a small sphere. For this ideal case, the energy distributed over the sphere at any time represents the total energy injected by the source and is therefore a constant. The total energy, E_T, divided by the surface area, S, of the spherical wavefront is the energy density, E,

$$E = \frac{E_T}{S} \quad . \qquad (16)$$

Since the surface area of a sphere is

$$S = \frac{4}{3} \pi r^2 \quad , \qquad (17)$$

the energy density as a function of distance, r, from the source is

$$E = \frac{3E_T}{4pr^2} \propto \frac{1}{r^2} \quad , \qquad (18)$$

where the last form hides all the constants to isolate the dependence on distance. This equation says that as the wavefield evolves the energy density decreases as $1/r^2$.

When we make measurements of sound or seismic waves, the observed quantity is not energy density, but the amplitude of pressure, particle displacement, or some other physical characteristic of the wave. The energy density [177] is proportional to the square of the measured amplitude

$$E(r) \propto A^2(r) \quad , \qquad (19)$$

and it follows that

$$A(r) \propto \sqrt{E(r)} \quad , \qquad (20)$$

Now consider the wavefront at two distances from the source, (r_0, r), where r_0 is some small nonzero reference distance. The amplitude ratio between measurements at r_0 and r is

$$\frac{A_0}{A} = \sqrt{\frac{E_0}{E}} = \sqrt{\frac{r^2}{r_0{}^2}} = \frac{r}{r_0} \quad , \qquad (21)$$

and solving this for the amplitude at *r*,

$$A = \frac{A_0}{r_0}\frac{1}{r} \propto \frac{1}{r} \quad (22)$$

This effect is termed either geometrical spreading, spherical spreading, or spherical divergence. The equation tells us that for 3D spherical waves the amplitude decay due solely to geometry is inversely proportional to distance traveled.

As an analogy, consider a bright red balloon. The reason it has such intense color is because the amount of pigment per unit surface area is high, something we can call the pigment density. Now inflate the balloon and notice the color. As the balloon expands, it changes from bright red, to red, to pink, to light pink. This occurs because the amount of red pigment in the balloon is constant, but the surface area is increasing, and therefore the pigment density is decreasing. In the same way, sound waves spreading in 3D become weaker as time progresses, because a fixed amount of energy is distributed over a spherical wavefront of increasing radius.

The wavefront radius depends on the velocity and time as

$$r = vt \quad , (23)$$

which allows us to conclude

$$A \propto \frac{1}{t} \quad . \qquad (24)$$

The 2D case follows the same line of argument with one essential difference; the energy for a 2D wave is not spread over the surface of a sphere but along the perimeter of a circle. This is because 2D wavefronts are expanding circles like surface waves in water. The analysis for the 2D case leads to geometric amplitude decay of

$$A_{2d} \propto = \frac{1}{\sqrt{t}} \quad . \ (25)$$

The difference between $1/r$ and $1/\sqrt{r}$ amplitude decay is significant and grows more significant the farther you are from the source. Figure 1.19 illustrates this difference. In Figure

1.19A, a 2D wavefield is shown at five different times. Velocity is constant and each time corresponds to a circular wavefront progressively farther from the source and the peak amplitudes are seen slowly decaying. The 3D case, Figure 1.19B, is a horizontal slice through the spherical wavefield at the same time/distance values as Figure 1.19A. Note the more rapid loss of amplitude as expected from 1/*r* decay.

picture 01-19 scaled 800

Fig. 1.19 Geometric spreading. (A) The amplitude of 2D waves decays with distance as circular wavefronts spread out. (B) In 3D, the wavefronts are spherical and decay is more rapid, as seen in this 2D slice through a series of wavefronts.

To show a numerical example, we need the concept of dB (decibels). This is a logarithmic unit for expressing ratios [165]. If D is the decimal ratio, then the dB measure is

$$dB = 10\ log_{10}(D) \quad , \qquad (26)$$

which can be inverted to find the decimal ratio when the dB level is known,

$$D = 10^{dB/10} \quad . \qquad (27)$$

Extending a calculation in Telford et al. [177], we consider 3D and 2D wave amplitudes at various distances from a source relative to a reference location, r_0, of 200 m. The dB loss due to geometrical spreading in 3D, L_{3D}, is

$$L_{3D} = 10\ log_{10}\left(\frac{A_0}{A}\right) = 10\ log_{10}\left(\frac{r}{r_0}\right) = 10\ log_{10}\left(\frac{r}{200}\right) \quad , \qquad (28)$$

and for 2D the result is

$$L_{2D} = 10\ log_{10}\left(\frac{A_0}{A}\right) = 10\ log_{10}\left(\frac{r}{r_0}\right)^{1/2} = 5\ log_{10}\left(\frac{r}{200}\right) \quad . \qquad (29)$$

Figure 1.20 plots this case in two ways. On the left are curves for 2D and 3D spreading losses as a function of distance from the source. For example, at 1500 m the 3D wave is 20 dB down from the measurement at 200 m. In decimal form, this result says the 3D wave at 1500 m has 1/100th of the amplitude measured 200 m from the source. At the same distance, the 2D wave is only 10 dB down and thus is 1/10th of the 200 m amplitude. In other words, the peak amplitude of the 2D wave at this distance is 10 times that of the 3D wave.

picture 01-20 scaled 650

Fig. 1.20 Geometric spreading. (A) Another way of expressing different decay rates in 2D and 3D is to use decibels (dB). This example shows amplitude spreading losses relative to a reference measurement 200 m from the source. (B) The difference between 3D and 2D geometric spreading loss expressed in dB as a function of distance.

The right graph captures this information in another way by plotting the difference of 3D and 2D spreading loss. At 1500 m this difference is -10 dB showing again that the 2D wave has 10 times the amplitude of the 3D wave. Note that throughout this discussion, it has been assumed the 2D and 3D waves had equal peak amplitude 200 m from the source, an unlikely situation. However, if the peak amplitudes are unequal at the reference distance, they can be normalized, and the results would still hold.

The consequence of our geometric spreading analysis is to realize that 2D waves will decay much more slowly than 3D waves, and this fact is evident in seismic data. Seismic surface waves, which we consider noise, must be actively suppressed, or they will easily overpower weak reflection events which exhibit 3D decay.