Acoustics for Ultrasound Imaging

Ben Cox

January 2013

Contents

Aims

These notes give an introduction to the acoustics behind ultrasound imaging, i.e. the physics and mathematical description of ultrasound in biological tissue. They don't describe the details of the different types of ultrasound machines and how they work, and how imaging algorithms generate the images we see. (Those lectures will follow these.) The description of acoustics given here starts off quite general, and is therefore closely related to the acoustics used to describe submarine sonar, seismic waves, noise from aircraft, etc. Towards the end of this set of lectures more details specific to acoustics in biological tissue will be introduced.

As well as being able to describe ultrasound waves mathematically, by the end of this part of the course, you should

- be aware of the fundamental assumptions underlying models of acoustics,
- understand the physical processes underlying acoustic phenomena,
- have an intuitive grasp of how acoustic waves propagate (eg. how they spread out and are absorbed).

You may even be able to answer a few questions such as:

- Can ultrasound be used to treat cancer or battle-field wounds?
- Diagnostic ultrasound imaging is harmless, but physiotherapists use ultrasound to encourage healing. Can both be true?
- Can ultrasound encourage bone growth and healing?
- Bubble cavitation is the main cause of damage to ships propellers. Why might we want to introduce bubbles into our bodies, and cavitate them?

1 Introduction

1.1 What is a wave?

We all know intuitively what we mean by waves and the term is used in everyday speech to describe a diverse family of phenomena - crime waves, waving goodbye, waves of nausea, combustion waves, etc. - which share common features, eg.

- with all waves some quantity changes (with time or distance or both),
- some waves have a periodic motion or pattern,
- some waves propagate,
- some waves pass through a medium leaving it unchanged,
- some waves have all these characteristics.

For *transverse* waves the wave motion is perpendicular (at right-angles) to the wave direction:

- Mexican (stadium, audience) waves,
- shear waves in solids (seismic S-waves),
- electromagnetic waves (light, radio, microwave, bluetooth, etc.),
- rope waves,
- Love waves (a shear surface wave).

For *longitudinal* waves the wave motion is aligned with the wave direction:

- stop-go traffic waves on busy motorways caused by braking,
- compressional waves in solids (seismic P-waves),
- (ultra)sound waves.

Some waves combine both transverse and longitudinal motions:

- surface seismic waves that cause earthquake damage (Rayleigh waves),
- surface acoustic waves (SAW devices are used as filters in mobile phones, etc.),
- waves in the sea (sometimes called gravity waves, because the restoring force is gravity, and not to be confused with gravitational waves predicted by General Relativity which are yet to be directly detected),
- ripples on a pond (surface-tension, 'capillary' waves).

Propagating waves are *a means by which energy can be transferred from one point to another without transfer of matter.*

Ultrasound waves are longitudinal, compressional waves, that can be periodic or pulsed, propagate at roughly 1500 m/s in water or biological tissue, can leave the medium unchanged (diagnostic ultrasound), but at higher intensities can also change it (therapeutic ultrasound).

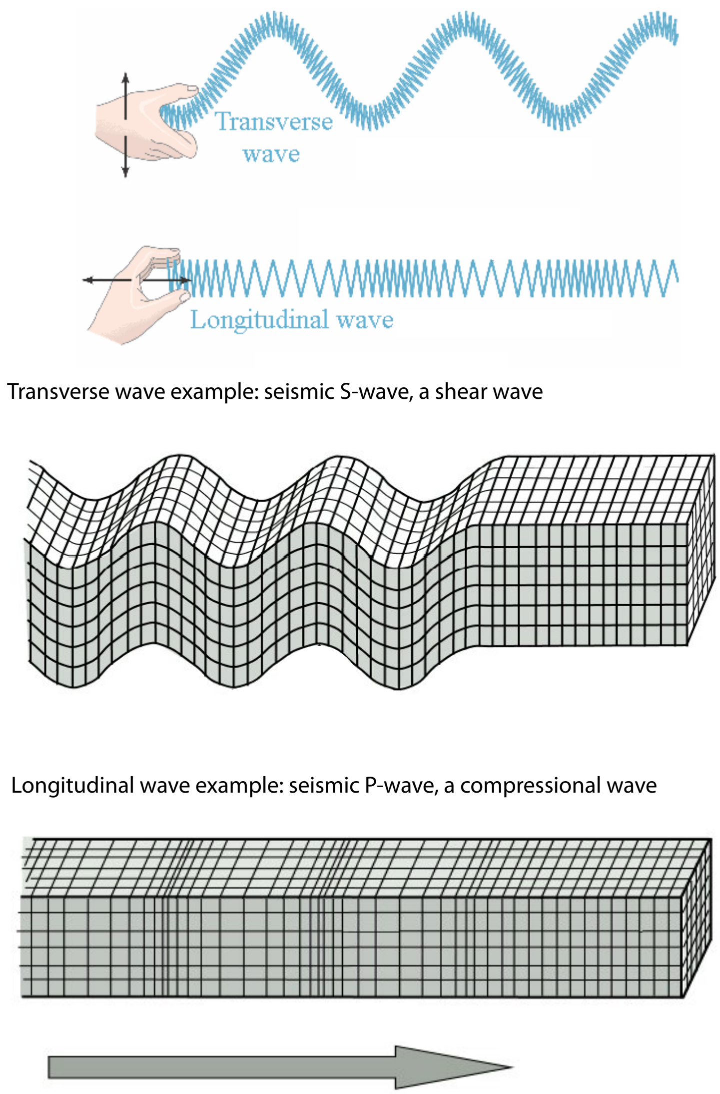


Figure 1: Examples of longitudinal and transverse waves with a ‘slinky spring’ and an elastic solid. P-waves (compressional waves) and S-waves (shear waves) in seismology are longitudinal and transverse waves respectively. (Sound waves are compressional waves.)

1.2 What is sound?

What makes a wave a *sound* wave? In his *Dictionary of Acoustics* Chris Morfey gives two definitions for 'sound':

> ***Sound*** *(1) a disturbance in pressure that propagates through a compressible medium. More generally, sound can refer to any type of mechanical wave motion, in a solid or fluid medium, that propagates via the action of elastic stresses and that involves local compression and expansion of the medium.*
>
> ***Sound*** *(2) the auditory sensation produced by transient or oscillatory pressures acting on the ear, or by mechanical vibration of the cranial bones at audio frequencies.*

The first definition is what we mean by sound in these notes. The second definition is what is meant by sound in everyday speech. Adult humans can hear sounds at frequencies between about 20 Hz and 15 kHz, which is roughly what definition (*2*) refers to as 'audio frequencies'. The terms *infrasound* and *ultrasound* are used for frequencies below and above this audible range.

- infrasonic $<$20 Hz
- audio frequency 20 Hz - 20 kHz
- ultrasonic $>$20 kHz, and sometimes $\gg$20 kHz

'Ultrasonic' therefore covers a huge range of frequencies from dogs hearing just above 20 kHz through a few MHz as used in ultrasound imaging to GHz used in acoustic microscopy (not covered here).

Children can hear higher frequencies than adults, which is the idea behind the child deterrent that works by emitting high frequency sounds at intensities that are very annoying to those who can hear it, but are too high for adults to hear. Exploiting the same phenomenon, 'mosquito' ringtones with frequencies above 15 kHz are available for the child who wants to avoid the teacher hearing his phone ringing in class (and who presumably hasn't managed to work out how to set it to vibrate).

The ranges of frequencies that some other animals can hear, or that they use for echolocation, are shown in Fig. 2 along with the frequency ranges of sonar (the lower frequencies tend to be used for military applications and the higher ones for fish/mine/shipwreck hunting), diagnostic ultrasound and HIFU (High Intensity Focussed Ultrasound, used for therapeutic purposes). Note that the scale on the frequency axis is a log scale which covers a huge range from 1 Hz to 100 MHz (8 orders of magnitude).

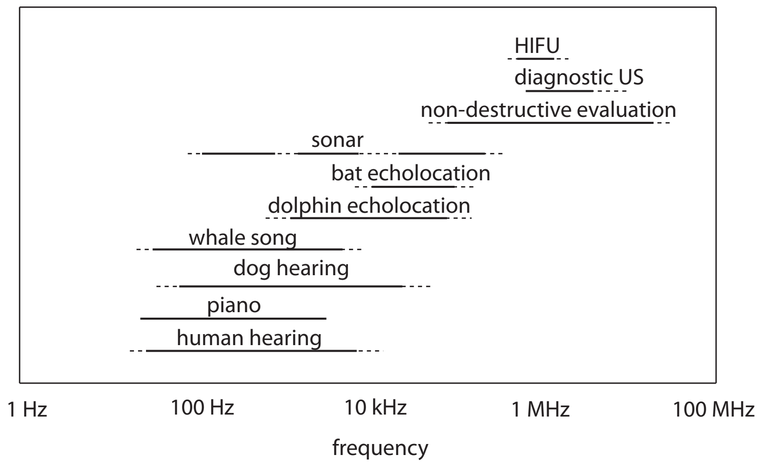


Figure 2: Ranges of sound frequencies used by certain animals and technologies.

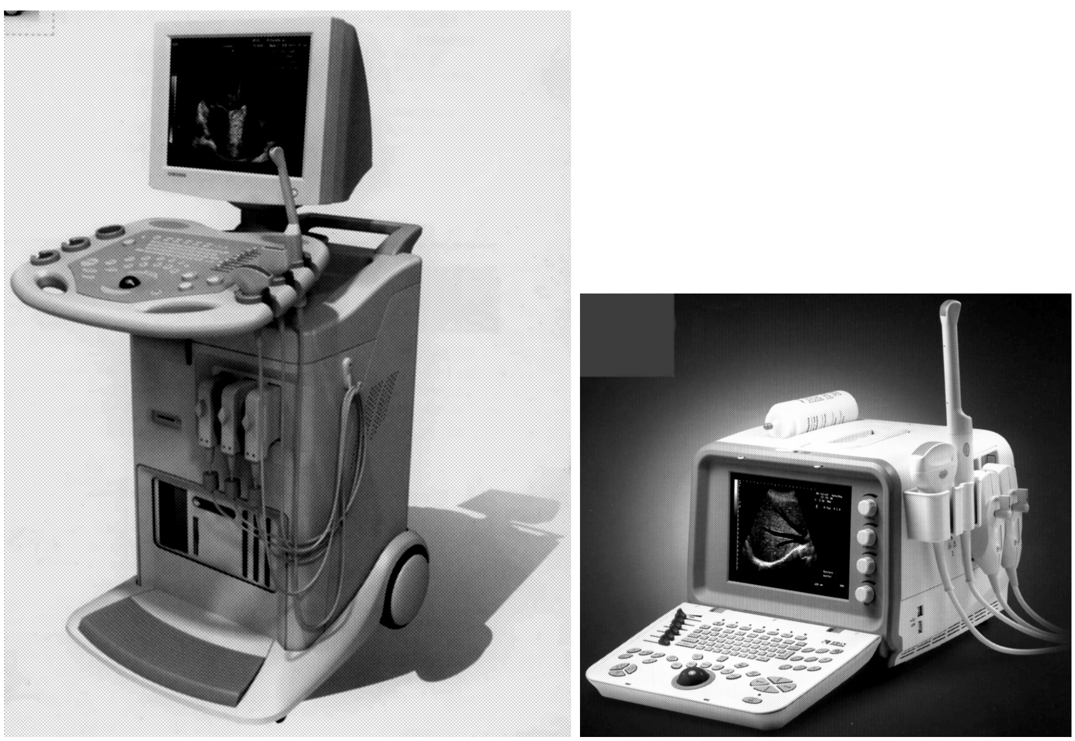
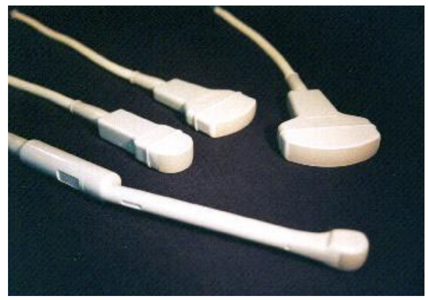
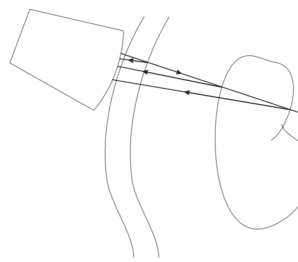


Figure 3: Typical diagnostic ultrasound imaging machines.

1.3 Very briefly, how does ultrasound imaging work?

Figure 3, on the previous page, shows two modern ultrasound imaging machines as might be found in a hospital. The details of ultrasound imaging machines and the image reconstruction algorithms they use are not covered in detail in these notes. Nevertheless, this page will give a quick overview of how ultrasound imaging works, as background to the maths and physics of ultrasound to follow. (Often ultrasound is abbreviated to US.)





• Ultrasound IN
An ultrasound *probe* is held against the skin, with some *coupling gel* in between to help the ultrasound get into the body efficiently. A beam of ultrasound pulses, or μs tone bursts at MHz frequencies, are sent into the tissue. The wavelength of the sound in the tissue is typically 0.1-1 mm.

•Ultrasound propagation and reflection
The ultrasound waves propagate through the tissue and partially reflect, or scatter, from acoustic heterogeneities (differences in acoustic impedance), eg. between different tissue types, fat, muscle, cysts, blood vessels, tumours, air pockets, etc. The reflected waves propagate back to the surface of the tissue.

•Ultrasound OUT
The reflected waves are detected at the surface, usually by the same transducer that sent them in, and recorded.

•Display
The time of arrival of a pulse back at the detector indicates the depth from which the reflection came. Using this information, the recorded waves are processed to form an image. At this stage, the image may also be processed to bring out certain features or to correct for the effect of acoustic absorption, for example.

In clinical ultrasound imaging the images update in real-time and so single-frame snapshots give a poor impression of the amount of information in the images. The correct interpretation of US images depends on the skill and experience of the user - who is therefore often a specially trained sonographer, radiographer or radiologist.

1.4 How and when did ultrasound imaging start?

Ultrasound imaging has its roots in **sonar** (SOund Navigation And Ranging), developed in the first half of the 20th century.

- *1912* Titanic sinks after hitting an iceberg, and Lewis Richardson patents an ultrasonic underwater detection device.
- *1914* Reginald Fessenden detects an iceberg from 3 km away by echo-ranging with a moving-coil transducer.
- *1918* Constantin Chilowsky and Paul Langevin detect a submarine at 1.5 km using a piezoelectric transducer.
- *1930s* Many ocean-going liners are fitted with echo-ranging devices.
- *1940s* Sonar is widely used for submarine detection in World War II.

Non-destructive testing, ultrasonic flaw detection, was developed soon after.

- *1928* Sergei Sokolov suggests ultrasonic flaw (defect) detection.
- *1941* Floyd Firestone invents "The Supersonic Reflectoscope" which used 5 MHz, 1 μs, pulses to locate flaws in materials.
- *late 1940s* Commercial ultrasonic flaw detectors are produced in UK (Kelvin Hughes), Germany (Siemens), Austria (KretzTechnik AG), France (Ultrasonique), Japan (Mitsubishi).

The application of ultrasound to the human body using flaw detection equipment soon followed, and was developed throughout the second half of the 20th century.

- *1940* H. Gohr and T. Wedekind suggest ultrasound for medical diagnosis.
- *1942* Karl Dussik attempts ultrasonic transcranial brain imaging using 1.2 MHz pulses of 100 ms duration.
- *1949* D. Howry converts old gun turret into scanner to measure cross-sections through the body. George Ludwig publishes report on the use of ultrasound for diagnosis, which includes a mean value of tissue sound speed of 1540 m/s and recommends an optimal frequency of 1 and 2.5 MHz.
- *1953* I. Edler and C. Hertz detect heart motions, first example of *echocardiology*.
- *1954* John Wild and John Reid build the first hand-held B-mode scanner, which worked at 15 MHz, and used it to detect breast cancers.
- *1958* Ian Donald diagnoses operable ovarian cancer in a woman thought to have inoperable stomach cancer.
- *1962* Shigeo Satomura suggests blood-flow measurement using Doppler-ultrasound.
- *1965* First real-time scanner built by Siemens.
- *1970s* Rapidly-growing interest in diagnostic ultrasound.
- *1980s* Ultrasound screening for pregnant women introduced.
- *1990s* Tissue harmonic imaging improves image resolution.
- *2000s* Real-time 3D ultrasound imaging.

1.5 Medical applications of ultrasound imaging

Ultrasound imaging, or ultrasonography, is very widely used in medicine for diagnosis and guiding interventional procedures, and the following list represents just some of the more common applications in use in hospitals and clinics today.

Obstetrics

- routine foetal development
- foetal location (intrauterine vs ectopic)
- multiple pregnancy
- check for physical abnormalities/viability
- determine sex

Cardiology (transthoracic or transoesophageal echocardiography)

- cardiac anatomy (hole in heart)
- assessment of cardiac function following heart attack
- measurement of ejection fraction
- heart valve assessment

Vascular ultrasound

- arterial and venous haemodynamics
- deep vein thrombosis (DVT)
- carotid doppler for stroke-risk assessment

Ophthalmology

- measure the cornea-retinal distance
- image the parts of the eye when there is no direct visual access (eg. swollen)

Abdominal ultrasound

- liver cirrhosis
- renal: kidney size, cysts, obstruction
- gall stones in gall bladder or bile duct
- appendicitis

Uro-genital

- uterus, bladder, testicular cancers
- ovarian cysts

Transrectal ultrasound

- Prostate and bowel cancer staging (not diagnosis)

2 Acoustics Basics

This section covers some basic notation and terminology used in acoustics, and some of the fundamental physical principles. We start with an apparently simple question:

2.1 Is soft tissue solid or liquid?

Should we treat biological tissue as liquid or solid, at least, as far as ultrasound is concerned? Some tissues are obviously solid - bones, for instance - but what about soft tissues such as skin or muscle? On the one hand they are very malleable and consist largely of water, on the other hand, we don't flow out of our beds and can stand up without our skin pooling around our ankles.

One of the differences between a solid and a liquid is that a solid has rigidity and can support a *shear* force. Imagine gluing the palm of your hand to a table and then trying to push your hand along the table top. You can move your hand a bit as the skin deforms but it will soon reach a point where you can't push it any further (without tearing the skin). Your skin will be supporting a shear force - it seems to behave like a solid.

Remove your hand from the table and your skin will return to the same shape it was originally - it is an *elastic* solid. This suggests we should treat soft tissue as an *elastic solid*. Nevertheless, soft tissue is usually modelled as a fluid as far as ultrasound is concerned.

In ultrasound imaging, soft tissue is usually modelled as a fluid.

Why do we treat soft tissue as a fluid when it is actually an elastic solid? The pragmatic reason is that this approximation has proven to be reasonably accurate and useful over half a century of ultrasound studies. Another motivation is that wave propagation in fluids is much simpler to visualise and model mathematically than wave propagation in solids. The reason why we get away with it, though, is because treating tissue as a fluid is equivalent to ignoring shear waves, and there are good reasons why shear waves can usually be neglected in ultrasound imaging:

- shear waves are not generated efficiently by ultrasound transducers which send short *compressive* pulses into the tissue,

- shear waves are strongly absorbed by soft tissue so don't travel far through it,

- shear waves travel much more slowly than compressional waves, so even if a shear wave travels far enough for a reflection to arrive back at the detector, it will arrive long after a compressional wave that set off at the same time.

2.2 Wavelength, frequency and wave speed

The terminology introduced in this section will be useful for describing waves and wave propagation. The terms in this section are quite general, and are not specific to ultrasound waves, so they could be used to describe any type of wave.

• **Wavelength**, λ. For waves with a periodic pattern in space, such as sand ripples on a beach, it is possible to define a characteristic length, the wavelength, as the distance between two adjacent peaks or troughs. Wavelength is often written using the Greek letter lambda λ, and has units of length, typically on a 0.1-1 mm scale for medical ultrasound.

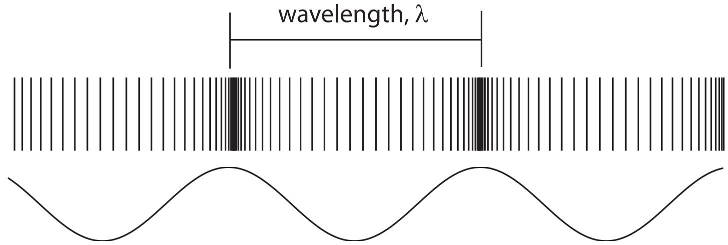


Figure 4: Wavelength: distance between two points of equal phase, eg. two peaks or troughs.

• **Frequency**, f. If a periodic pattern propagates (think ripples on a pond spreading out) then the rate at which the peaks pass a fixed point - the number of peaks passing per second - is called the frequency of the wave, with units of cycles per second or Hertz (Hz). Medical ultrasound typically uses frequencies from 1-15 MHz. The reciprocal of the frequency is called the **period** of the wave. The term *circular* or *angular* frequency, ω, which is the number of *radians* per second, is also used. It is related to f by

$$\omega = 2\pi f \tag{1}$$

(Whenever ω is used in these notes it refers to the angular frequency in rads/s even when it is just called the 'frequency'.)

• **Wave speed**, c. Propagating waves often travel at characteristic speeds in a given medium. It is usually denoted c or v. Sound travels at about 340 m/s in air and 1500 m/s in water. Mexican waves travel at typically 12 m/s apparently (in air presumably). The wavelength, frequency and wave speed are related by

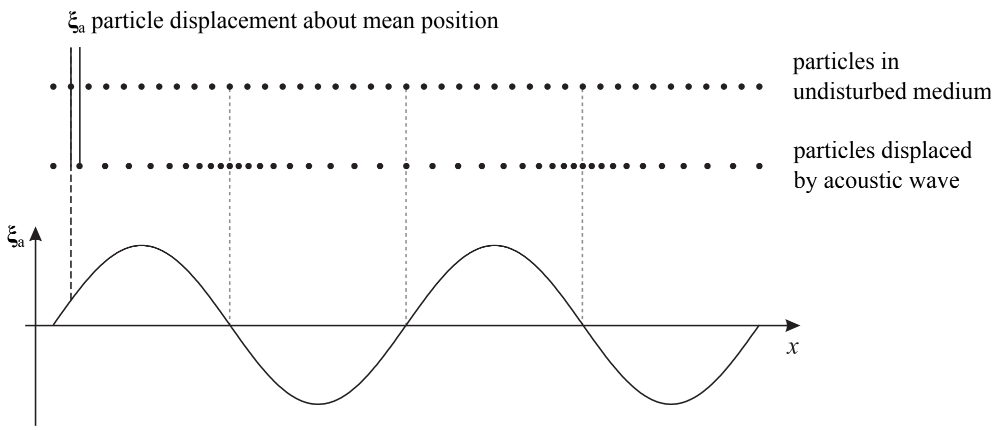
$$\boxed{c = f\lambda} \tag{2}$$

• **Wavenumber**, k is defined as $k = 2\pi/\lambda$. It is the number of radians per metre, rad/m, but the units are sometimes written as just m^{-1} as radians are dimensionless. e.g. if $\lambda = 1/3$ m so that 3 wavelengths fit into 1 m, then there are $3 \times 2\pi$ radians in 1 m so the wavenumber $k = 6\pi$ rads/m. The wavenumber is sometimes called the *spatial* frequency of the wave to distinguish it from the *temporal* frequency, f. (Confusingly, sometimes in *optics* the wavenumber is defined as $1/\lambda$ not $2\pi/\lambda$ as here.)

2.3 Acoustic variables

We now need to introduce some terms so that we can talk more specifically about acoustic (or ultrasound) waves, which as we saw above we will treat as compressional waves in a fluid medium.

• **Particle displacement** [nm, μm] Fluid 'elements' or particles of a medium move - oscillate about their mean position - as a sound wave passes through. (What do we mean by a fluid 'element'? See Appendix A.1.) For a point in space $\mathbf{x}$, the *acoustic particle displacement* will be written using the symbol, $\boldsymbol{\xi}_a(\mathbf{x})$ (pronounced *z-eye* or *k-see* with the subscript '$_a$' to indicate that the displacement is due to an acoustic wave and not something else, such as a flow). The particle displacement is a vector quantity, which means that is consists of both a magnitude and a direction. In practice, vectors are often defined with respect to a fixed coordinate system, and the vector is given in terms of its components along the coordinate axes. eg. for the Cartesian coordinates (x, y, z), the particle displacement could be written as $\boldsymbol{\xi}_a(\mathbf{x}) = (\xi_{a,x}(\mathbf{x}), \xi_{a,y}(\mathbf{x}), \xi_{a,z}(\mathbf{x}))$. For more on vectors see Appendix A.3.



• **Particle velocity** [m/s] In general, the fluid velocity vector will include a non-oscillatory net flow velocity $\mathbf{u}_0$ (which will often be assumed to be zero, $\mathbf{u}_0 = 0$) as well as the oscillatory acoustic velocity vector, which for each fluid particle is the time derivative of the acoustic particle displacement $\mathbf{u}_a = \partial\boldsymbol{\xi}_a/\partial t$. The total particle velocity $\mathbf{u}$ is therefore

$$\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_a. \tag{3}$$

• **Acoustic pressure** [kPa, MPa], **density** [kg/m^3], and **temperature** fluctuations [K] As a sound wave passes through a point in space and the particles there are displaced, the pressure, density and temperature there will fluctuate too. If the ambient (non-fluctuating) values of each quantity are denoted with subscript $_0$, and the fluctuations due to the acoustic wave denote with subscript $_a$, then the total pressure p, density ρ (rho), and temperature T can be written as

$$p = p_0 + p_a, \qquad \rho = \rho_0 + \rho_a, \qquad T = T_0 + T_a. \tag{4}$$

Note that not all books use exactly this notation. Some use p or p_e ('excess pressure') for p_a.

2.4 Energy, power and intensity

One property of propagating waves is that they transfer energy from one point to another without the transfer of matter. In acoustics, this flow of energy is called the *acoustic intensity*. Another term is *acoustic energy flux vector*.

Think of an imaginary square (with unit area) in a fluid through which a sound wave is travelling. The energy that crosses that surface per unit time is the intensity in the direction normal to the square. Think of another, similar, square at a different angle. The amount of energy that flows across that surface will be different (e.g. if the square is aligned with the direction of the wave then no energy will cross it), so the acoustic intensity is a *vector* - it has direction asociated with it as well as magnitude.

• **Instantaneous acoustic intensity**, $\mathbf{I}(\mathbf{x}, t)$, in a time-varying acoustic field, is a vector defined as

$$\mathbf{I}(\mathbf{x}, t) = p_a(\mathbf{x}, t)\mathbf{u}_a(\mathbf{x}, t) \qquad [\mathrm{J/s/m^2} = \mathrm{W/m^2}] \tag{5}$$

where p_a is the acoustic pressure and $\mathbf{u}_a$ the acoustic particle velocity. (Is this a plausible definition of intensity? Recall that, in general terms, pressure = force per unit area, velocity = distance per unit time and 'work done' = force × distance. Acoustic intensity is a measure of power per unit area = work done per unit time per unit area = pressure × velocity.)

• **Acoustic intensity** is also used sometimes to refer to the *time-averaged acoustic intensity*

$$\mathbf{I}_{\mathrm{av}} = \frac{1}{T}\int_0^T \mathbf{I}(\mathbf{x}, t)dt = \frac{1}{T}\int_0^T p_a\mathbf{u}_a dt = < p_a\mathbf{u}_a > \qquad [\mathrm{W/m^2}] \tag{6}$$

where $<>$ is used to indicate a time-average. The length of the time-average, T, is either a whole number of cycles ($2\pi n$ radians, $n \in \mathbb{N}$) or a very long time (infinite number of cycles). It should be clear from the context whether or not the time-average is implied.

• **Sound power**, P in Watts, is the rate at which acoustic energy is flowing across a surface S

$$\mathcal{P} = \iint_S \mathbf{I} \cdot \mathbf{n} dS \tag{7}$$

where $\mathbf{n}$ is the unit normal to the surface S. If S is a closed surface, then it is only non-zero when the surface surrounds a *source*.

• **Acoustic energy density**, w in $\mathrm{J/m^3}$, is a scalar (field) and is the sum of the time-averaged kinetic and potential energy densities. Kinetic energy is related to motion (in this case particle motion) and potential energy refers to 'stored' energy (here stored in the compressibility of the medium, like elastic potential energy in a spring). The two terms of the acoustic energy density are therefore given by

$$w(\mathbf{x}) = w_{\mathrm{kin}}(\mathbf{x}) + w_{\mathrm{pot}}(\mathbf{x}) = \frac{1}{2}\rho_0 < u_a^2 > + \frac{1}{2}\frac{< p_a^2 >}{B_s} \tag{8}$$

The acoustic energy density is related to the *acoustic radiation pressure* which will be described in Section 6.

2.5 Decibels and sound levels

Be careful with decibels (dB) as it is not a unit in the usual fundamental sense (like Pascals for pressure or metres for length). However, it is a widely used unit to describe the 'loudness' of sounds, among other things, so it is worth understanding.

- **Decibel: ratio of powers** Wherever decibels occur, they are usually the log of a ratio of two quantities with units of *power* (Watts):

$$10\log_{10}\left(\frac{\text{power (W)}}{\text{reference power (W)}}\right) \qquad [\text{dB}] \tag{9}$$

  (The 'deci' refers to the 10 in front of the definition; without this the unit is called the Bel, and is very rarely used.)

  Measures of sound are often given in decibels (dB) because the range of acoustic pressure amplitudes involved in sound waves of interest to humans is huge so a log scale helps to compress it to a scale with a range of 10-100 or so.

- The **sound power** $\mathcal{P}$ can be written in decibels as

$$\text{sound power level} = 10\log_{10}\left(\frac{\mathcal{P}}{\mathcal{P}_{\text{ref}}}\right) \tag{10}$$

  where the reference power is defined as

$$\mathcal{P}_{\text{ref}} = 10^{-12}\text{W} = 1 \text{ picowatt.} \tag{11}$$

- **Sound pressure level** Acoustic pressure does not have units of power, but is often written in dB too. In plane or spherical sound waves, the acoustic pressure and intensity are related by $|\mathbf{I}| = p_a^2/\rho_0 c_0$, so the sound pressure level is defined as

$$\text{Sound pressure level, } L_p = 10\log_{10}\left(\frac{p_a^2}{p_{\text{ref}}^2}\right) = 20\log_{10}\left(\frac{p_a}{p_{\text{ref}}}\right) \tag{12}$$

  where the reference pressures are given by convention as

$$p_{\text{ref}} = \begin{cases} 1\,\mu\text{Pa, in water, tissue, etc} \\ 20\,\mu\text{Pa, in air or other gases} \end{cases} \tag{13}$$

sound source	sound pressure level SPL (dB re 20 μPa)	acoustic pressure amplitude (Pa)
jet engine (10m away)	150	630
pain threshold	130	63
pneumatic drill	100	2
hearing damage (long term exposure)	85	0.6
major road	80	200×10^{-3}
conversation	60	20×10^{-3}
countryside at night	40	2×10^{-3}
normal breathing	10	63×10^{-6}
normal threshold of hearing	0	20×10^{-6}

Figure 5: Typical sound levels for common sources of sound in air.

sound source	SPL (dB re 1 μPa)	acoustic pressure (MPa)	frequency range	acoustic displacement
LFA sonar (22 km - near)	180-240	0.001-1	100-500 Hz	0.2 μm-1.1 mm
HIFU	260	10	1-2 MHz	0.5-1 μm
dolphin sonar at 1 m	180-220	0.001-0.1	50-150 kHz	0.7 nm-0.2 μm
diagnostic ultrasound	230-250	0.5-5	3-15 MHz	3 nm - 0.1 μm
physiotherapy ultrasound	230	0.5	1-3 MHz	10-40 nm

Figure 6: Typical sound levels for common sources of sound in water. Note that because of the different frequency ranges involved, although the SPL values are roughly similar, the displacement amplitudes vary by a factor of 10^5.

3 Acoustic Wave Equation

Acoustic waves can be described using the 'wave equation', one of the classic partial differential equations of physics. In one dimension it is written (for the acoustic pressure p_a in this case) as

$$\frac{1}{c^2}\frac{\partial^2 p_a}{\partial t^2} = \frac{\partial^2 p_a}{\partial x^2} \tag{14}$$

and it appears almost anywhere that propagating waves appear: sound, vibration, electromagnetic waves, water waves, etc. Extending it to three dimensions gives

$$\frac{1}{c^2}\frac{\partial^2 p_a}{\partial t^2} = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) p_a \quad \text{also written} \quad \left(\frac{1}{c_0^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) p_a = 0 \tag{15}$$

(See Appendix A.3 to brush up on the gradient operator ∇.)

Acoustic waves can propagate in any medium that has the properties of compressibility and inertia (the tendency of something with mass to resist acceleration or deceleration). In this Section, this linear acoustic wave equation will be derived in this section from three fundamental principles.

- The first principle is that the medium (tissue, water, etc) is compressible, ie. the density can change, and this is captured in the *conservation of mass* or *continuity equation*.
- The second is that the medium has inertia and this is expressed in an equation that describes the *conservation of momentum* - essentially a version of Newton's second law $F = ma$, and sometimes known as *Euler's equation*
- The third is a relationship between various thermodynamic quantities, and is expressed as an *equation of state* or *pressure-density relation*.

3.1 Conservation of mass - continuity equation

The basic idea here is simple: we equate the rate of increase of mass in a fixed (imaginary) volume V, Fig. 7, with the net amount of mass that flows into V, per unit time. In other words, if mass goes into V then, because mass is conserved, the amount of mass inside V must be going up. The mass density is written as $\rho(\mathbf{x}, t)$ kg/m^3 so the total mass in V is the volume integral $\iiint_V \rho dV$. eg. if V is a cube of uniform density then its mass is ρL^3, or $\rho((4/3)\pi r^3)$ for a sphere. (For a reminder about volume integrals see Appendix A.5.) The rate of change of mass in V can therefore be written as

$$\text{rate of increase of mass in } V = \frac{\partial}{\partial t}\iiint_V \rho dV. \tag{16}$$

The surface of V (bounding V) is S. The net flow of mass out through a small part of the surface, dS at point $\mathbf{x}_s$, is the dot product $\rho\mathbf{u}\cdot\mathbf{n}dS$ (units kg/s) where $\mathbf{n}$ is the *outward* unit normal to dS. The net mass flowing *into* V per unit time is therefore

$$\text{net mass flowing } \textit{into } V \text{ per unit time } = -\iint_S \rho\mathbf{u}\cdot\mathbf{n}dS. \tag{17}$$

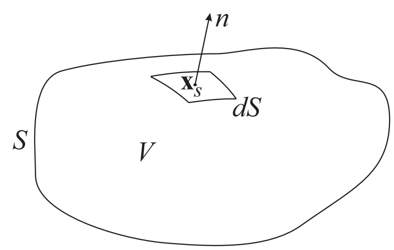
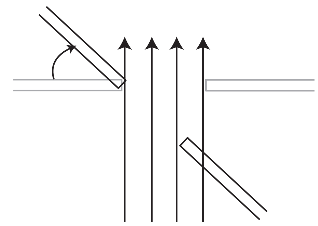


Figure 7: Control volume for derivation of mass conservation equation.



To understand better where the $\mathbf{u} \cdot \mathbf{n}$ comes from consider a beam of fixed width passing through an aperture of the same width, as illustrated in Fig. 3.1. When the angle between the beam and the normal to the aperture (call it α) is zero, all the beam gets through, but as α increases, less and less gets through until, when they are at right-angles none gets through. In general, the amount getting through is $\mathbf{u} \cdot \mathbf{n} = |\mathbf{u}| \cos \alpha$.

The *divergence theorem*, sometimes called *Gauss' divergence theorem*, equates the outward flux of a vector field over a closed surface with the volume integral of the divergence of the vector field over the region enclosed by the surface:

$$\iint_S \mathbf{A} \cdot \mathbf{n} dS = \iiint_V \nabla \cdot \mathbf{A} dV, \tag{18}$$

where $\mathbf{A}$ is a vector field. This can be used to convert Eq. (17) to

$$\text{net mass flowing into } V \text{ per unit time } = - \iiint_V \nabla \cdot (\rho \mathbf{u}) dV. \tag{19}$$

(To understand more intuitively where the $\nabla \cdot (\rho \mathbf{u})$ in Eq. (19) comes from, see Fig. 8.) Equating Eqs. (16) and (19), and rearranging, gives

$$\iiint_V \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) \right) dV = 0 \tag{20}$$

As this equation is zero for any arbitrary volume V, its integrand must also be zero, so

$$\boxed{\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{u})} \tag{21}$$

A different, and perhaps more intuitive, derivation of this equation of mass conservation is given in Fig. 8.

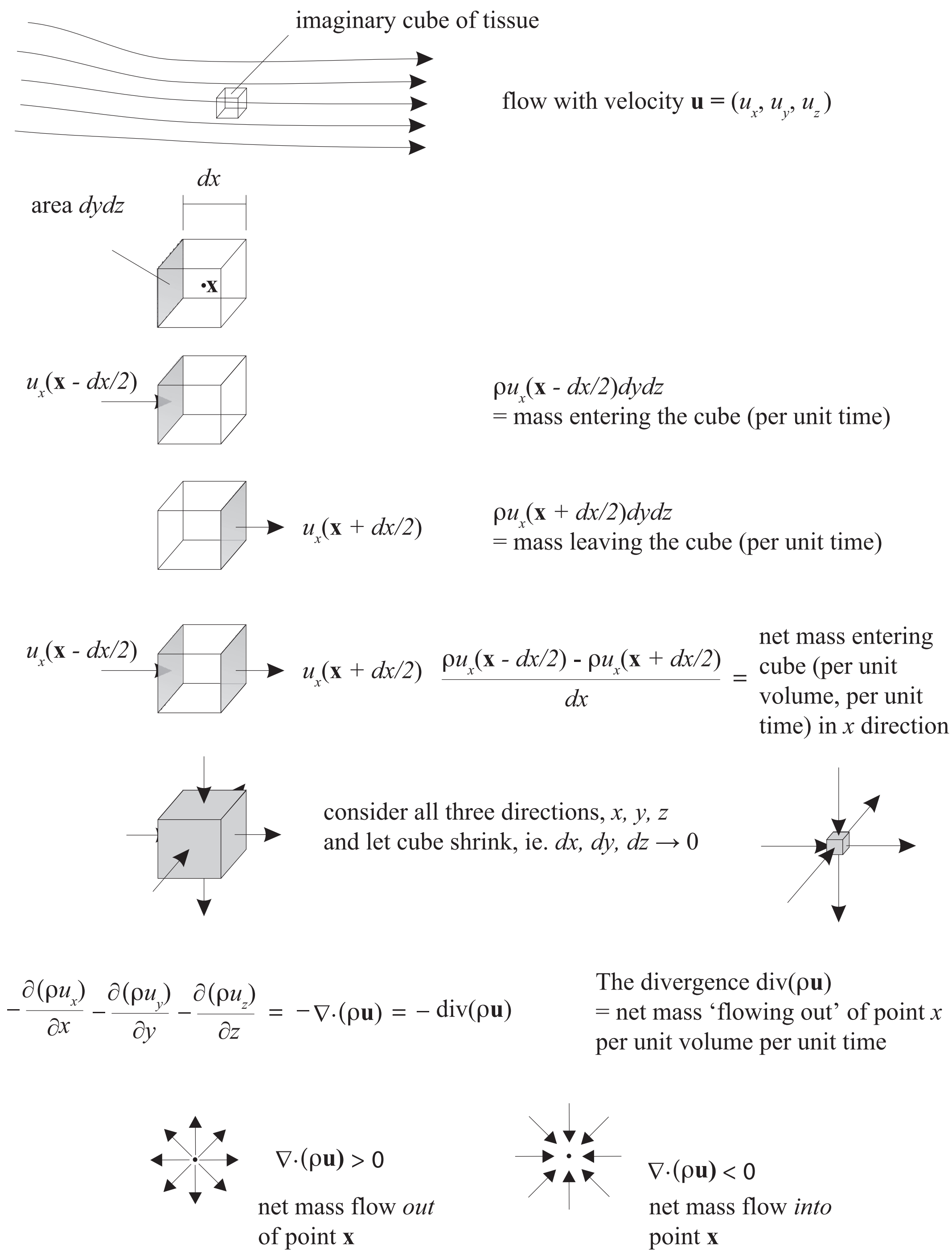


Figure 8: Convervation of mass

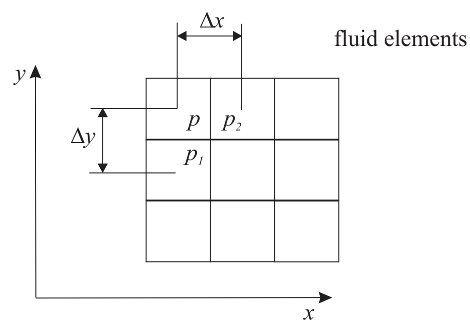
3.2 Conservation of momentum - Euler's equation of motion

The equation derived in this section expresses the conservation of momentum in a similar way to that in which the previous section described the conservation of mass. It is called *Euler's equation of motion*, and is essentially Newton's second law applied to a fluid element. Newton's second law equates force to the rate of change of momentum. For a single fluid element, so small that its density is constant throughout its volume, the momentum may be written as

$$\text{rate of change of momentum per unit volume of a fluid element} = \rho \frac{D\mathbf{u}}{Dt} \tag{22}$$

where D/Dt is the material derivative operator (see Appendix A.4).

The forces on the fluid element (ignoring forces due to external fields such as gravity) are due to pressure differences with neighbouring fluid elements:



$$\text{force on the fluid element in } x \text{ direction} = \frac{p - p_2}{\Delta x} \to -\frac{\partial p}{\partial x} \tag{23}$$

$$\text{force on the fluid element in } y \text{ direction} = \frac{p - p_1}{\Delta y} \to -\frac{\partial p}{\partial y} \tag{24}$$

(The minus sign arises because the gradient in direction x, say, is defined as positive for an increase in the x-direction, and an increasing pressure in the x-direction would give a force in the negative x-direction.)

Combining the forces in all three directions gives the net force on the fluid element as the vector

$$\text{net force per unit volume on the fluid element} = -\nabla p \tag{25}$$

Equating (22) and (25) gives

$$\boxed{\rho \frac{D\mathbf{u}}{Dt} = -\nabla p} \tag{26}$$

3.3 Equation of state

A relationship between thermodynamic variables in a compressible fluid is called an *equation of state* or, in acoustics, the rather more prosaic *pressure-density relation.* The general form may be written as

$$p = p(\rho, s) \tag{27}$$

where s is the specific entropy (see Appendix A.2).

• **No heat flow** Laplace's hypothesis is that there is no heat flow in the fluid during the passage of a sound wave. At an acoustic pressure maximum in a sound wave, there will be a temperature rise, and at a pressure minimum (rarefaction) a temperature fall. If the distance between the temperature maxima and minima (half a wavelength) is much wider than the distance the heat can diffuse in one period (1/frequency), then there will be very little heat flow and the temperature will not equilibriate. As the distance heat can diffuse in time t is proportional to $\sqrt{t}$, but the wavelength is proportional to the time period, so at lower frequencies (longer wavelengths) the adiabatic approximation is better, and at very high frequencies it becomes more isothermal.

Is it true that the heat cannot flow far enough in one time period? The speed of sound in tissue is about $c = 1540$ m/s, so an ultrasound wave with frequency $f = 3$ MHz has a wavelength $\lambda \approx 0.5$ mm and so the max and min temperatures are separated by $\lambda/2 = 250$ μm. The thermal penetration depth is defined as $z = \sqrt{4Dt}$, where the thermal diffusivity of water $D = 114 \times 10^{-9}$ m^{2}s^{-1} and t is a characteristic time, here the time period $t_p = 1/f = 0.33$ μs. The thermal penetratoin depth is therefore about 0.4 μm, which is much less than 250 μm, so there is no way the heat can cover the distance in the time.

The no-heat-flow condition can be stated by saying that the specific entropy remains constant for any given fluid particle,

$$\frac{Ds}{Dt} = 0. \tag{28}$$

(The 'material derivative' operator D/Dt is explained in Appendix A.4.) If the entropy is originally the same everywhere then this says that the entropy will not change anywhere, so it need not be explicitly considered and the equation of state can be written as a 'pressure-density relation':

$$p = p(\rho). \tag{29}$$

• **Pressure-density relation** Fig. 9 shows the isentropic pressure-density relationship for water at 20°C, starting at the undisturbed, ambient, values p_0, ρ_0. It is clearly nonlinear, although for small amplitude waves it can be assumed to be linear. The acoustic pressure $p_a = p - p_0$, can be written in terms of the density change $\rho_a = \rho - \rho_0$ as a **Taylor series expansion** about the point (ρ_0, p_0):

$$p_a = \rho_a \left(\frac{\partial p}{\partial \rho}\right)_{s,0} + \frac{\rho_a^2}{2!}\left(\frac{\partial^2 p}{\partial \rho^2}\right)_{s,0} + \frac{\rho_a^3}{3!}\left(\frac{\partial^3 p}{\partial \rho^3}\right)_{s,0} + \dots \tag{30}$$

$$= A\left(\frac{\rho_a}{\rho_0}\right) + \frac{B}{2!}\left(\frac{\rho_a}{\rho_0}\right)^2 + \frac{C}{3!}\left(\frac{\rho_a}{\rho_0}\right)^3 + \dots \tag{31}$$

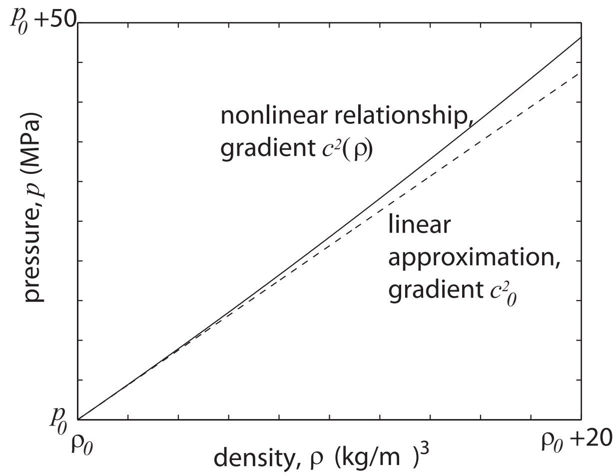


Figure 9: Pressure-density relation for water at 20°C, and the linear approximation.

The subscript $_0$ indicates that the derivatives are evaluated at the ambient values (p_0, ρ_0), and we have introduced the constants

$$A \equiv \rho_0 \left(\frac{\partial p}{\partial \rho}\right)_{s,0}, \quad B \equiv \rho_0^2 \left(\frac{\partial^2 p}{\partial \rho^2}\right)_{s,0}, \quad C \equiv \rho_0^3 \left(\frac{\partial^3 p}{\partial \rho^3}\right)_{s,0}. \tag{32}$$

• **Sound speed** As well as being the intuitively understandable speed at which a sound wave travels, the sound speed is also a thermodynamic property of a material, defined as

$$c^2 \equiv \left.\frac{\partial p}{\partial \rho}\right|_s \tag{33}$$

which is the gradient of the solid line in Fig. 9. The sound speed at the ambient density ρ_0 is the gradient of the dashed line. It is defined as

$$c_0^2 \equiv \left.\frac{\partial p}{\partial \rho}\right|_{s,0} = \frac{A}{\rho_0} \tag{34}$$

• **Linearised pressure-density relation** Substituting the equation for c_0, Eq. (33), into Eq. (31), and neglecting the higher order terms, gives a linear acoustic pressure-density relation

$$\boxed{p_a = c_0^2 \rho_a} \tag{35}$$

As can be seen from Fig. 9, this linear relationship only holds true for small amplitude waves. It is not always true for the amplitudes of waves encountered in diagnostic and therapeutic ultrasound, but it is a good approximation most of the time.

• **Bulk modulus** The definition of A given above is just the definition of the isentropic (constant entropy) bulk modulus, B_s, so $A = B_s$. (Do not confuse the B which appears in Eq. (31) with the isentropic bulk modulus B_s, they are different.)

3.4 Linear acoustic wave equations

To derive the linear wave equation, we need these three equations from the last few sections:

- linearised pressure-density relation, $p_a = c_0^2 \rho_a$ (36)
- mass conservation equation, $\partial \rho / \partial t = -\nabla \cdot (\rho \mathbf{u})$ (37)
- momentum conservation equation, $D\mathbf{u}/Dt = -\nabla p / \rho$ (38)

We will also make the approximation that

$$\rho^{-1} \approx \rho_0^{-1}. \tag{39}$$

• **Linearised mass conservation equation** Expanding the right-hand side of Eq. (37) using the chain rule, and using approximation in Eq. (39), Eq. (37) may be written as

$$\frac{1}{\rho_0}\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{u} - \frac{\mathbf{u} \cdot \nabla \rho}{\rho_0} \tag{40}$$

Recall that $\rho = \rho_0 + \rho_a$ and $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_a$. If the ambient density does not change with time, $\partial \rho_0 / dt = 0$, there is no net flow $\mathbf{u}_0 = 0$, and we make the *linearising assumption* that

$$|\nabla \cdot \mathbf{u}| \gg \left| \frac{\mathbf{u} \cdot \nabla \rho}{\rho_0} \right|, \tag{41}$$

Eq. (40) becomes the linearised mass conservation (or continuity) equation for the acoustic pressure

$$\boxed{\frac{\partial \rho_a}{\partial t} = -\rho_0 \nabla \cdot \mathbf{u}_a}. \tag{42}$$

• **Linearised momentum conservation equation** Expanding the material derivative in Eq. (38) (see Appendix A.4) and again using the approximation in Eq. (39) gives

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{\nabla p}{\rho_0} \tag{43}$$

If the ambient pressure is the same everywhere, $\nabla p_0 = 0$, there is no net flow $\mathbf{u}_0 = 0$, and we make the *linearising assumption* that

$$\left| \frac{\partial \mathbf{u}}{\partial t} \right| \gg |(\mathbf{u} \cdot \nabla)\mathbf{u}| \tag{44}$$

Eq. (43) becomes the linearised momentum conservation (or force or Euler's) equation:

$$\boxed{\frac{\partial \mathbf{u}_a}{\partial t} = -\frac{\nabla p_a}{\rho_0}} \tag{45}$$

• **Linearising assumption** Both of the linearising assumptions, Eqs. (41) and (44) hold when the total particle speed, $|\mathbf{u}|$, is much less than the sound speed, ie. the Mach number is much less than 1, $M = |\mathbf{u}|/c_0 \ll 1$.

• **Wave equation for heterogeneous media** Differentiating Eq. (36) with respect to time gives:

$$\frac{\partial p_a}{\partial t} = c_0^2 \frac{\partial \rho_a}{\partial t} \tag{46}$$

Substituting this into Eq. (42) gives

$$\frac{1}{\rho_0 c_0^2} \frac{\partial p_a}{\partial t} = -\nabla \cdot \mathbf{u}_a. \tag{47}$$

Differentiating this with respect to time, and differentiating Eq. (45) with respect to space, gives the two equations

$$\frac{1}{\rho_0 c_0^2} \frac{\partial^2 p_a}{\partial t^2} = -\frac{\partial}{\partial t} \left(\nabla \cdot \mathbf{u}_a\right), \qquad -\nabla \cdot \left(\frac{\partial \mathbf{u}_a}{\partial t}\right) = \nabla \cdot \left(\frac{\nabla p_a}{\rho_0}\right) \tag{48}$$

As $\partial/\partial t(\nabla \cdot \mathbf{u}_a) = \nabla \cdot (\partial \mathbf{u}_a/\partial t)$, Eqs. (48) can be combined to give a linear wave equation for heterogeneous media:

$$\boxed{\frac{1}{c_0^2} \frac{\partial^2 p_a}{\partial t^2} - \rho_0 \nabla \cdot \left(\frac{\nabla p_a}{\rho_0}\right) = 0} \tag{49}$$

• **Wave equation for homogeneous media** When the ambient density ρ_0 is the same everywhere, Eq. (49) becomes the wave equation for homogeneous media:

$$\boxed{\frac{1}{c_0^2} \frac{\partial^2 p_a}{\partial t^2} - \nabla^2 p_a = 0} \tag{50}$$

3.5 How does the wave equation describe propagating waves?

Greater insight into solutions of the wave equation can be gained by considering the two first-order equations that were used to derive it, than by looking at the wave equation itself. In one-dimension, Euler's equation, Eq. (45), and the equation combining the conservation of mass and the acoustic equation of state, Eq. (47), can be written as

$$\frac{\partial u_a}{\partial t} = -\frac{1}{\rho_0}\frac{\partial p_a}{\partial x}, \tag{51}$$

$$\frac{\partial p_a}{\partial t} = -\rho_0 c_0^2 \frac{\partial u_a}{\partial x}. \tag{52}$$

Eq. 51 says that the rate of change of the particle velocity at a point is proportional to the negative of the gradient of the acoustic pressure there. Similarly, and somewhat symmetrically, Eq. 52 says that the rate of change of the acoustic pressure at a point is proportional to the negative of the gradient of the particle velocity (the divergence in higher dimensions). These two equations can be combined to give the wave equation, as described in the previous section.

Graphs may better indicate why these two 'rate equations' lead to wave propagation. In Fig. 10 the wave is propagating to the right, in the positive x direction, towards $x = +\infty$. The acoustic pressure p_a and particle velocity u_a are in phase. Consider the position $x = a$. The gradient of the pressure is positive, and so the rate of change of the particle velocity is negative, according to Eq. (51), and the particle velocity there will decrease. The gradient of the particle velocity at position $x = a$ is also positive, and so the rate of change of the ppressure is negative, according to Eq. (52), and the pressure there will decrease. Consider the position $x = b$. The gradients of the pressure and particle velocity there are both negative, and the pressure and particle velocity there will therefore increase, as time moves forward. In this way, we can see that the pulse moves to the right. The opposite case is shown in Fig. 11, where the pressure and particle velocity are out of phase and the wave travels to the left.

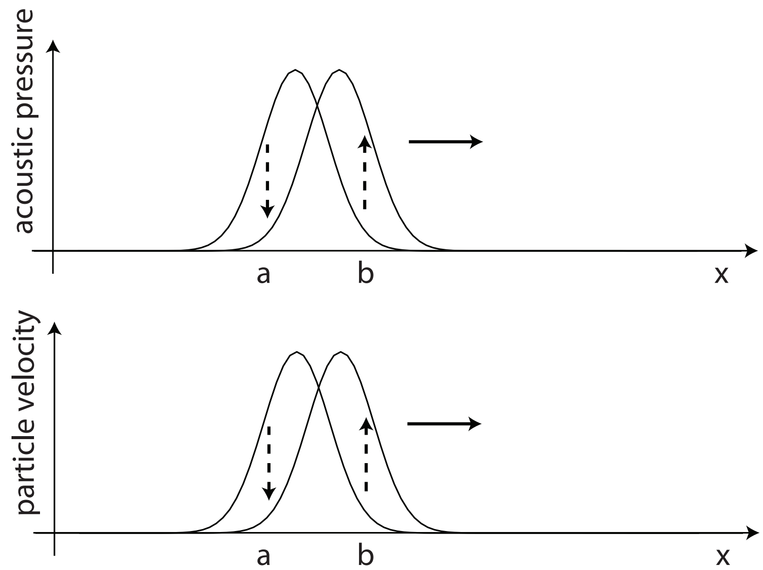


Figure 10: Propagation in +x direction

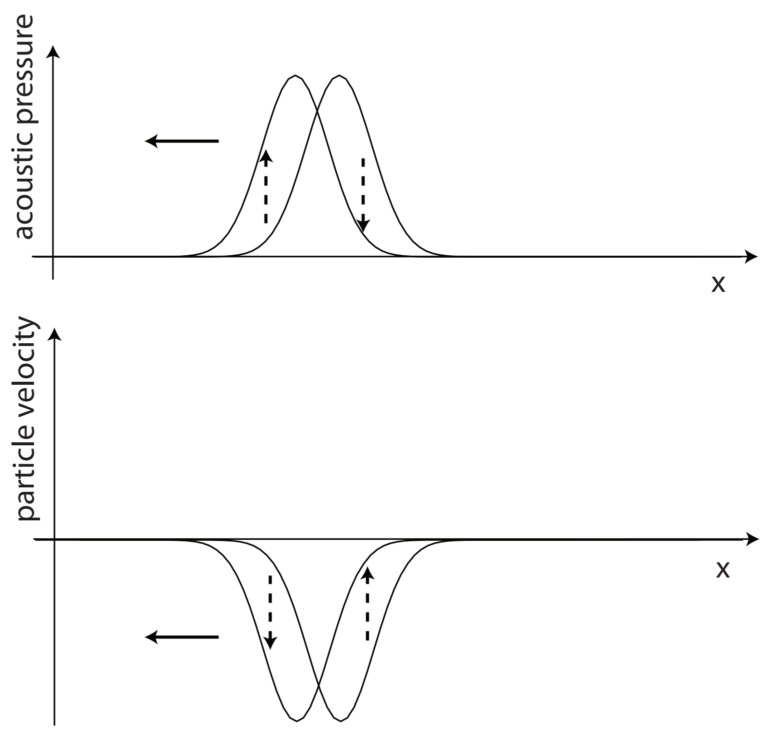


Figure 11: Propagation in -x direction

3.6 Single-frequency plane waves

Think of a complicated acoustic wave field, eg. at a concert. That wave field, and the waves that constitute it, are all solutions to the wave equation, although they may be too difficult to write down. Numerical approaches such as finite-difference models and ray tracing can be used to calculate complicated wave-fields when necessary. In the next few sections, three simple but important types of wave will be described: plane, cylindrical and spherical waves. These canonical solutions can be very helpful in several ways: they can help to give us an intuitive understanding of how acoustic waves behave in certain situations, they can be used as building blocks to construct more complicated solutions (this is the theory of Green's functions which we won't go into in detail), and they can approximate many real life cases.

• **Propagating waves** Two solutions to the one-dimensional wave equation, Eq. (14) can be written in the forms

$$f(x - ct) \text{ or } f(x/c - t) \text{ for a wave that propagates towards } x = +\infty \tag{53}$$

$$f(x + ct) \text{ or } f(x/c + t) \text{ for a wave that propagates towards } x = -\infty \tag{54}$$

where f is an arbitrary function of any shape (and nothing to do with frequency, also sometimes written f). To check these are solutions, substitute $f(x \pm ct)$ into Eq. (14) and check the equation still holds true. As $\partial^2 p/\partial t^2 = c^2 f''(x - ct)$ and $\partial^2 p/\partial x^2 = f''(x - ct)$ in both cases, they are solutions.

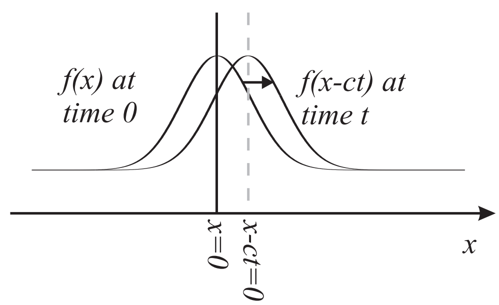


Figure 12: In this example, the peak always occurs where the argument of f is zero, so it can be seen that $f(x - ct)$ or $f(x/c - t)$ must be moving in the positive x-direction. After time t, the wave has moved a distance of $+ct$.

• **Plane waves** A plane wave is a wave in which surfaces of equal phase are flat planes. Plane waves are the most important type of acoustic wave to understand. As we will see later, wave-fields of any shape or pattern can be formed by adding together plane waves with different frequencies and directions, so understanding them will be a significant step towards understanding all types of wave-fields. Also plane, or very nearly plane, waves occur in real life, eg. waves propagating along ducts or pipes, or radiating from large vibrating surfaces.

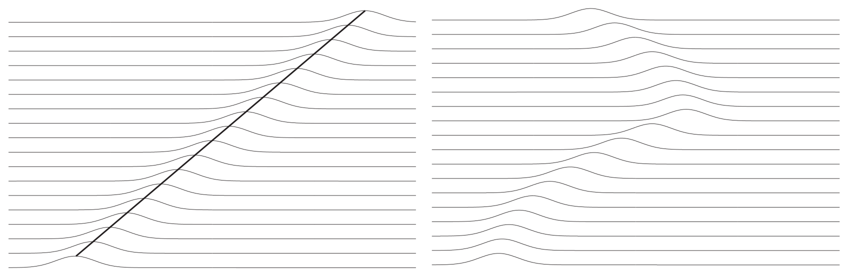


Figure 13: A plane wave with one of the lines of equal phase marked, and a non-plane wave.

• A plane wave that **varies sinusoidally in space** can be written in any of the forms

$$a_0 \sin(kx), \qquad a_0 \cos(kx), \qquad \mathrm{Re}(a_0 e^{ikx}) \tag{55}$$

where a_0 is the *amplitude* of the wave and $k = 2\pi/\lambda$ is the wavenumber or spatial frequency [rad/m]. ($\mathrm{Im}(a_0 e^{ikx})$ would do too, but it is not commonly used.)

• A **single-frequency wave** is any wave that varies sinusoidally in time, and can be written in the forms

$$f(x)\sin(\omega t), \qquad f(x)\cos(\omega t), \qquad \mathrm{Re}(f(x)e^{i\omega t}) \tag{56}$$

where ω is the temporal (circular) frequency [rad/s].

• A **single-frequency harmonic plane wave** is a propagating plane wave that varies sinusoidally both as a function of position and as a function of time, and can be written in the forms

$$a_0 \sin(kx \pm \omega t), \qquad a_0 \cos(kx \pm \omega t), \qquad \mathrm{Re}(a_0 e^{i(kx \pm \omega t)}) \tag{57}$$

where $-$ corresponds to propagation in the positive x-direction, and $+$ to propagation in the negative x-direction.

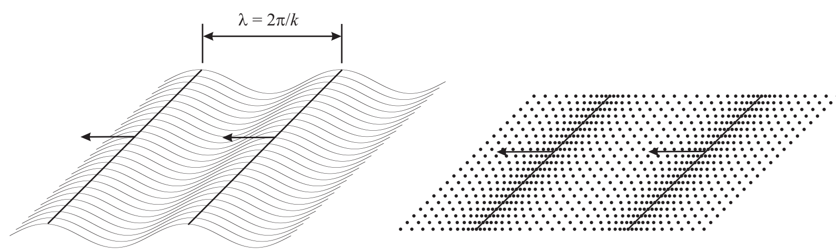


Figure 14: Single-frequency, harmonic, plane waves, with two lines of equal phase shown.

• **Dispersion relation** A relationship between spatial and temporal frequency is, in general, called a dispersion relation. This is because ‘dispersive’ is the term used to describe a medium in which the sound speed depends on frequency, $c = c(\omega)$. In all the applications of interest here, the sound speed is non-dispersive, and the temporal and spatial frequency are related by $\omega = c_0 k$.

• **Simple relationships** between acoustic variables occur for single-frequency harmonic plane waves. These equations are useful when trying to obtain rough estimates of acoustic quantities for more complex wavefields. As such they are widely used, but it is important to remember that they are only strictly true for single-frequency harmonic plane waves. Consider a wave travelling in the positive x-direction. The acoustic quantities describing it can be related as follows:

- if the particle displacement is $\xi_a = \xi_0 \cos(kx - \omega t) = \mathrm{Re}(\xi_0 \exp(i(kx - \omega t)))$,
- particle velocity, $u_a = \partial \xi_a/\partial t = \omega \xi_0 \sin(kx - \omega t) = \mathrm{Re}(-i\omega\xi_0 \exp(i(kx - \omega t)))$,
- acoustic pressure, $p_a = \rho_0 c_0 u_a = \rho_0 c_0 \omega \xi_0 \sin(kx - \omega t)$,
- acoustic intensity, $I = p_a u_a = p_a^2/(\rho_0 c_0)$,
- time-averaged acoustic intensity, $I_{\mathrm{av}} = p_a^2/(2\rho_0 c_0)$.

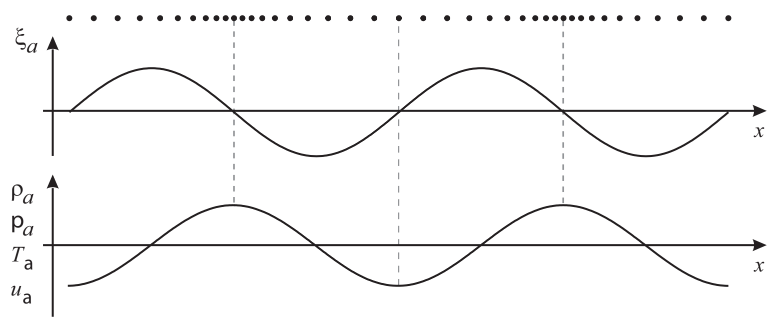


Figure 15: In a plane wave propagating in the positive x-direction, the acoustic pressure is out-of-phase with the particle displacement, but in-phase with the acoustic particle velocity, temperature and density.

• **Direction matters!** Note that the particle velocity u_a is a vector, so when the wave is propagating in the opposite direction, towards $x = -\infty$, some minus signs appear, e.g. the wave is written as $\exp(i(kx + \omega t))$ rather than $\exp(i(kx - \omega t))$ so the relationship between the pressure and particle velocity becomes $p_a = -\rho_0 c_0 u_a$ (see Section 3.5).

• **Phase** In the wave described above, propagating in the positive x-direction, the acoustic pressure and particle velocity are $\pi/2$ (90°, $\lambda/4$) out-of-phase with the particle displacement, but in-phase with each other. In other words, the pressure maxima and minima occur where the particle displacement is zero. The fluctuations in the temperature and density are in phase with the pressure too. (The density gets out of phase with the pressure when there is acoustic absorption - see Section 5.)

• **Characteristic acoustic impedance,** Z is defined as $\rho_0 c_0$, and is the constant relating the pressure and particle velocity for a plane wave. It will reappear in the discussion of reflection coefficients.

3.7 Plane wave decomposition

In this section, we see that any wavefield can be written as a sum (integral) of harmonic plane waves with different frequencies and directions.

• **Plane waves in three dimensions** One-dimensional, single-frequency harmonic plane waves can be written as $a_0 \exp(i(kx \pm \omega t))$ as described above. (We will drop the Re denoting the real part from now on, but acoustic pressure is a real quantity, so it should be clear where it is implied.) In a sense, all plane waves are one-dimensional in that they are only travelling in one direction. However, if that direction is not along one of the coordinate axes but at an angle to them, then the wave must be written as a combination of components in the directions of the coordinate axes.

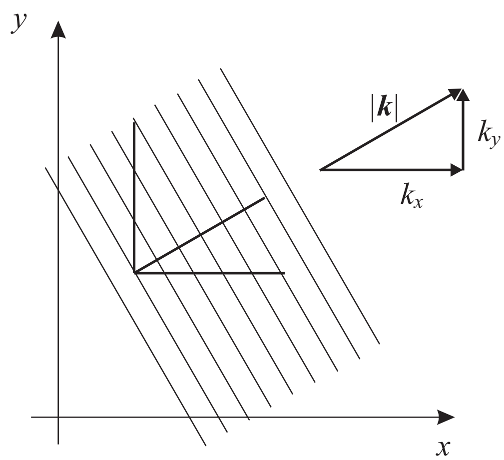


Fig. 3.7 shows a plane wave travelling in direction $\mathbf{k} = (k_x, k_y)$ where k_x and k_y are the magnitude of the wavenumber vector $\mathbf{k}$ in the directions x and y. The magnitude of the wavenumber in the direction of propagation is given by $k = |\mathbf{k}| = \sqrt{k_x^2 + k_y^2}$. This can be seen by comparing how many crests are covered by the unit lengths drawn parallel to the direction of propagation, and the x and y axes. (Recall the wavenumber is measured in radians/m and crest-crest is 2π radians. Here $k_x \approx 5$, $k_y \approx 3$, and $k = \sqrt{k_x^2 + k_y^2} \approx \sqrt{5^2 + 3^2} \approx 6$.)

In general,

- A plane wave has direction given by its wavevector $\mathbf{k} = (k_x, k_y, k_z)$
- The magnitude of the wavevector is called the wavenumber and is given by

$$k = |\mathbf{k}| = \sqrt{k_x^2 + k_y^2 + k_z^2} \tag{58}$$

- A general plane wave with unit amplitude can therefore be written as

$$e^{ik_x x} e^{ik_y y} e^{ik_z z} e^{\pm i\omega t} = e^{i(k_x x + k_y y + k_z z \pm \omega t)} = e^{i(\mathbf{k}\cdot\mathbf{x} \pm \omega t)} \tag{59}$$

• **Harmonic plane waves as a solution to the wave equation** To check the single frequency, harmonic plane wave $p_a(\mathbf{x}, t) = a_0 e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)}$ is a solution to the wave equation we simply substitute this into Eq. (50) and both sides of the equation are still equal. Noting that

$$\frac{\partial p_a}{\partial t} = -i\omega a_0 e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)} = -i\omega p_a, \qquad \frac{\partial^2 p_a}{\partial t^2} = -\omega^2 p_a \tag{60}$$

and

$$\nabla p_a = i\mathbf{k}a_0 e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} = i\mathbf{k}p_a, \qquad \nabla^2 p_a = -k^2 p_a \quad \text{where } k = |\mathbf{k}| \tag{61}$$

gives

$$-\left(\frac{\omega^2}{c_0^2}\right) p_a + k^2 p_a = 0 \tag{62}$$

which is true when

$$\boxed{\omega = c_0 k} \tag{63}$$

which is the *dispersion relation* stated earlier.

• **Spatial Fourier transform** In Section A.7 in the appendix, it is explained how any temporal signal $f(t)$ can be described in terms of frequency components $F(\omega)$ via a Fourier transform. The same can be done for functions of space. For instance, if $p(x)$ describes a one-dimensional shape, then it can be split into *spatial frequency* components exactly as if it were a time signal:

$$p(x) = \int_{-\infty}^{\infty} P(k_x)\exp(ik_x x)dk_x \tag{64}$$

The spatial frequency is k_x, the wavenumber in the x-direction.

This idea can be straightforwardly extended to higher dimensions, so a two dimensional pattern $p(x, y)$ can be written as a sum of spatial frequencies in the two directions x and y

$$p(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(k_x, k_y)\exp(i(k_x x + k_y y))dk_x dk_y. \tag{65}$$

This can be straightforwardly extended to 3D, and written more succinctly as

$$p(\mathbf{x}) = \iiint P(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{x})d\mathbf{k}. \tag{66}$$

The integrand is recognisable from the discussion of a plane wave above, and it is a short step to write a complete, time-varying sound field $p(\mathbf{x}, t)$ as a sum of propagating plane waves with amplitudes $P(\mathbf{k})$:

$$\boxed{p(\mathbf{x}, t) = \iiint P(\mathbf{k})\exp(i(\mathbf{k}\cdot\mathbf{x} - \omega t))d\mathbf{k}} \tag{67}$$

This **plane wave decomposition** describes any arbitrary, time-varying, acoustic pressure field $p(\mathbf{x}, t)$ as a sum of plane waves with directions $\mathbf{k}$ and amplitudes $P(\mathbf{k})$. Note that $P(\mathbf{k})$ is not written explicitly as a function of temporal frequency ω because $\omega = c\sqrt{k_x^2 + k_y^2 + k_z^2} = ck$ according to the dispersion relation, Eq. (63).

This is a very important result, because plane waves are simpler to deal with than arbitrarily shaped waves or wave fields, so the ability to break them down into plane waves, which can be dealt with separately and then recombined, is very powerful.

3.8 Spherical and cylindrical waves

• **Spherical waves** Image a point source of sound, such as a tiny, spherical, pulsating bubble. The waves that travel away from the source will be spherical. These are a useful type of wave to study because many sources can be approximated as emitters of spherical waves, and - more fundamentally - any distribution of sources can be represented as a collection of point sources, each emitting spherical waves.

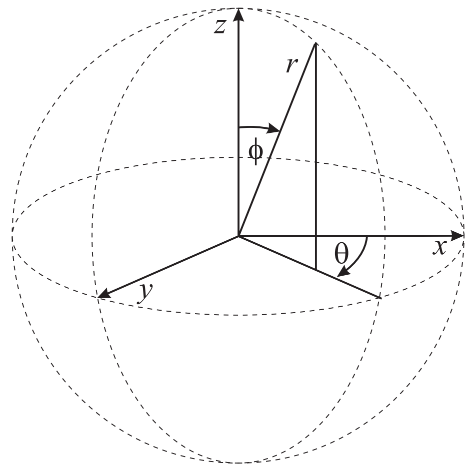


Figure 16: The relationship between spherical coordinates (r, θ, ϕ) and rectangular 'Cartesian' coordinates (x, y, z).

To find a spherically-symmetric solution to the wave equation, we first move from cartesian (x, y, z) coordinates to spherical (r, θ, ϕ) coordinates, and then look for a solution that does not depend on the angles θ and ϕ, but just on r (and time). The solution must be of the form $f(r, t)$, where $r = \sqrt{x^2 + y^2 + z^2}$.

We know that the Laplacian in Cartesian coordinates, ie. (x, y, z), is written as $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$. It can be shown that when this is written in spherical coordinates, and when there is no dependence on θ and ϕ, it takes the form:

$$\nabla^2 p(r, t) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(rp) \tag{68}$$

Substituting this into Eq. (50) gives the wave equation in spherical coordinates:

$$\boxed{\frac{1}{c_0^2}\frac{\partial^2(rp)}{\partial t^2} - \frac{\partial^2(rp)}{\partial r^2} = 0} \tag{69}$$

as $r\partial^2 p/\partial t^2 = \partial^2(rp)/\partial t^2$. This is just the one-dimensional wave equation, Eq. (14) but in (rp) instead of just p. We know two solutions to Eq. (14), Eq. (53) and (54), so the solution to (69) is therefore

$$rp = g(r \pm ct) \qquad \text{or} \qquad p(r, t) = \frac{g(r \pm ct)}{r} \tag{70}$$

where r is the distance from the centre of the wave, and g is any function.

Unlike plane waves whose amplitude is constant, the amplitude of a spherical wave decreases with distance r from the centre, so decreases with time for an outgoing wave $f(r - ct)/r$. This is known as *spherical spreading*.

• **Cylindrical waves** Plane and spherical waves were described above. Plane waves travel in just one direction, and spherical waves spread out in three dimensions (x, y, z). In between these two there is a third case worth considering: cylindrical waves. Imagine a flat layer of water, perhaps a metre deep and infinitely wide, like a huge flat lake. If a string made of explosives were hung vertically in the water and the explosives detonated, a cylindrical wave would travel out through the lake.

Cylindrical waves spread out in two dimensions and undergo *cylindrical spreading*, which is less severe than spherical spreading as the pressure falls with distance as only $1/\sqrt{r}$.

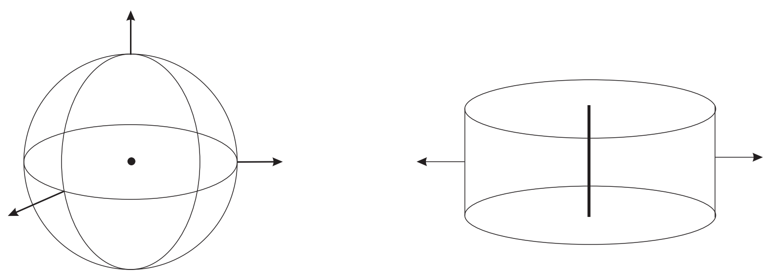


Figure 17: Spherical and cylindrical waves from point and lines sources of sound respectively.

• **Geometric spreading** refers to the $1/r$ and $1/\sqrt{r}$ decrease in the pressure amplitude of outgoing spherical and cylindrical waves respectively. The reason for this decrease can be understood by considering the energy per unit area.

If a spherical wave is emitted with total energy E, then its intensity (energy per unit surface area) will be $E/(4\pi r^2)$, where r is the radius of the spherical wave. As the acoustic intensity is proportional to the acoustic pressure squared, the acoustic pressure amplitude of the spherical wave will fall off as $1/r$.

Similary with a cylindrical sound wave. The energy would be spread over a cylindrical surface, so the intensity would be $E/(2\pi rh)$ where h is the height of the cylinder and this r is a cylindrical coordinate (rather than the spherical coordinate r above). The pressure amplitude in a cylindrical wave therefore falls off as $1/\sqrt{r}$.

4 Reflection, Refraction and Scattering

So far, we have looked waves propagating in acoustically homogeneous materials ie. materials in which the density, ρ_0, and sound speed c_0 are constant. If this were the case in soft tissue then ultrasound imaging would not work. There need to changes in the sound speed or the density in order for the ultrasound waves to be reflected. More precisely, the *characteristic acoustic impedance* of the material, $\rho_0 c_0$, must vary between different tissue types.

4.1 Acoustic impedance

• **Specific acoustic impedance**, z_n, is defined as the ratio of the acoustic pressure p_a and particle velocity in an acoustic wave in a specified direction $\mathbf{u} \cdot \mathbf{n}$, where $\mathbf{n}$ is a unit vector giving the direction of interest and $\mathbf{u}$ is the vector particle velocity of the wave. For any given direction $\mathbf{n}$, the specific acoustic impedance z_n is the scalar quantity

$$\text{specific acoustic impedance, } z_n = p_a/(\mathbf{u} \cdot \mathbf{n}) \tag{71}$$

• **Characteristic acoustic impedance** was encountered in Section 3.6 as the relationship between the acoustic pressure and particle velocity *in a plane wave*, $p_a = \rho_0 c_0 u_a$. The 'characteristic acoustic impedance', $\rho_0 c_0$, is purely a property of the medium through which the wave is travelling. The units are Rayls = kgm^{-2}s^{-1}.

$$\text{characteristic acoustic impedance, } Z = \rho_0 c_0 \tag{72}$$

tissue type	sound speed c_0 (m/s)	density ρ_0 (kg/m^3)	$Z =$ $\rho_0 c_0$ (Rayls)	B/A
air at STP	330	1.2	400	0.4
water at 20°C	1480	1000	1.48×10^6	5.0
water at 37°C	1530	990	1.51×10^6	5.4
blood	1585	1060	1.68×10^6	6
brain	1560	1035	1.62×10^6	6.5
breast	1510	1020	1.62×10^6	9.6
fat	1430	930	1.33×10^6	10.3
heart	1555	1060	1.65×10^6	5.8
kidney	1560	1050	1.64×10^6	9.0
liver	1580	1050	1.66×10^6	6.7
muscle	1580	1040	1.65×10^6	7.4
spleen	1565	1055	1.65×10^6	7.8
bone	3200	1990	6.36×10^6	-

Figure 18: Typical values of the sound speed c_0, density ρ_0, characteristic acoustic impedance $\rho_0 c_0$, and nonlinearity parameter B/A (see Section 6) for air, water and various tissues.

4.2 Reflection and transmission coefficients

• **Boundary conditions** When a wave reaches a boundary, part will be reflected and part transmitted. These three parts (incident wave, reflected wave, transmitted wave) must obey two boundary conditions:

(1) *Continuity of pressure* The acoustic pressure must be the same on both sides of the boundary. There must be no net force.

(2) *Continuity of normal particle velocity* The particle velocities normal to the boundaries must be equal. The fluid must stay in contact.

• **Normal incidence pressure reflection and transmission coefficients**

When an acoustic pressure wave with amplitude p_i is normally incident on an interface (a change in characteristic acoustic impedance), a wave with amplitude p_r will be reflected and another wave, with amplitude p_t will be transmitted. These wave amplitudes define the pressure reflection and transmission coefficients, R and T:

$$R = \frac{p_r}{p_i}, \qquad T = \frac{p_t}{p_i} \tag{73}$$

Boundary conditions (1) and (2) imply that at the boundary

$$p_i + p_r = p_t, \qquad u_i + u_r = u_t \tag{74}$$

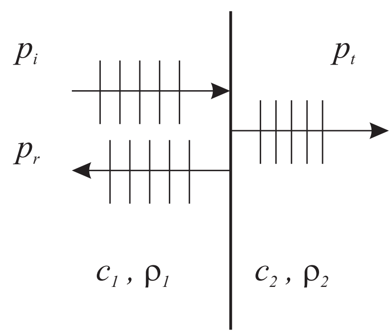


Figure 19: Reflection of a normal incident plane wave

where u is the particle velocity. Dividing these equations gives

$$\frac{p_i + p_r}{u_i + u_r} = \frac{p_t}{u_t} \tag{75}$$

Recalling that, in a plane wave, p/u is the characteristic acoustic impedance, we can make the substitutions

$$\frac{p_i}{u_i} = \rho_1 c_1, \qquad \frac{p_r}{u_r} = -\rho_1 c_1, \qquad \frac{p_t}{u_t} = \rho_2 c_2 \tag{76}$$

where the minus sign arises because u_r is in the opposite direction to u_i. Substituting these into Eq. (75) and rearranging gives

$$\boxed{\text{normal incidence pressure reflection coefficient, } R = \frac{\rho_2 c_2 - \rho_1 c_1}{\rho_2 c_2 + \rho_1 c_1} = \frac{Z_2 - Z_1}{Z_2 + Z_1}} \tag{77}$$

As the pressure-continuity boundary condition, Eq. (74), implies $1 + R = T$, the transmission coefficient can be found to be

$$\boxed{\text{normal incident pressure transmission coefficient, } T = \frac{2\rho_2 c_2}{\rho_2 c_2 + \rho_1 c_1} = \frac{2Z_2}{Z_2 + Z_1}} \tag{78}$$

• **Oblique incidence pressure reflection and transmission coefficients**
Generalising this to a plane wave that is incident at angle is straightforward. Simply replace the particle velocity with its component normal to the surface, eg. $u \cos\theta$, where θ is the incidence angle. The boundary conditions are therefore

$$p_i + p_r = p_t, \qquad u_i \cos(\theta_i) + u_r \cos(\theta_r) = u_t \cos(\theta_t) \qquad \text{at the boundary} \tag{79}$$

The angle of incidence must equal the angle of reflection $\theta_i = \theta_r$, and **Snell's law** requires the incidence and transmittance angles to be related by

$$\frac{\sin\theta_i}{c_1} = \frac{\sin\theta_t}{c_2} \tag{80}$$

Following a similar procedure as for the normally incident wave, the reflection and transmission coefficients can be written as a function of incident angle as

$$\boxed{\text{pressure reflection coefficient, } R(\theta_i) = \frac{\rho_2 c_2 \cos(\theta_i) - \rho_1 c_1 \cos(\theta_t)}{\rho_2 c_2 \cos(\theta_i) + \rho_1 c_1 \cos(\theta_t)}} \tag{81}$$

$$\boxed{\text{pressure transmission coefficient, } T(\theta_i) = \frac{2\rho_2 c_2 \cos(\theta_i)}{\rho_2 c_2 \cos(\theta_i) + \rho_1 c_1 \cos(\theta_t)}} \tag{82}$$

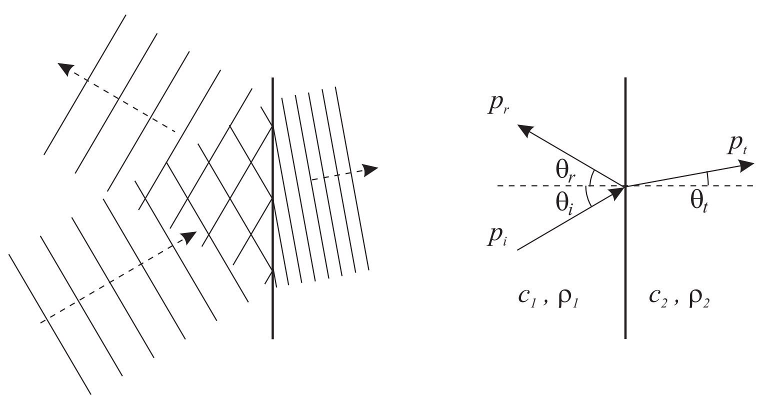


Figure 20: Reflection of a plane wave incident at an oblique angle.

• **Intensity reflection and transmission coefficients** Note that the reflection coefficients described above are *pressure* reflection coefficients and they are related (slightly counter-intuitively but in agreement with the boundary conditions in (74)) as

$$1 + R = T \tag{83}$$

When considering energy or intensity, the law of conservation of energy must apply, which says that the amount of energy transmitted must equal the amount of energy incident plus the amount reflected:

$$1 = R_e + T_e \tag{84}$$

where the subscript $_e$ indicates that these are energy reflection and transmission coefficients. In a plane wave, the intensity is proportional to p_a^2, so

$$R_e = |R|^2 \tag{85}$$

and therefore the energy transmission coefficient is related to the pressure reflection coefficient by

$$T_e = 1 - |R|^2 \tag{86}$$

4.3 Why ultrasound imaging works

With any medical imaging modality the input signal (here an ultrasound pulse, in CT scans an X-ray beam) must be affected by the medium - the body - enough to be able to carry information about it, but not so much that it is impossible to get a signal out, or the signal is so distorted it is useless. eg. whole body medical imaging using neutrinos would be no use as they don't interact with tissue, and using UV light would be no good as it is completely absorbed within a few microns. In ultrasound imaging, there is a balance too:

- There must be sufficient difference between the acoustic impedances of the different tissue types that there is some reflection (but not total reflection or nothing beneath that layer would be imaged).
- The sound speeds must not vary too much between the tissue, or it would be impossible to calculate how deep the reflections are from their arrival times.

As luck would have it, this is pretty much the case in the human body, as Fig. 21 shows.

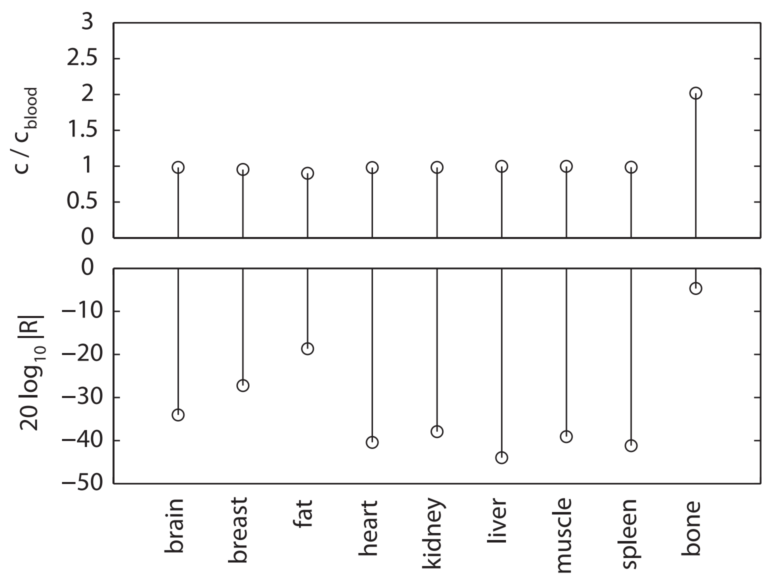


Figure 21: (Top) Sound speeds of various tissue types normalised by the sound speed in blood. (Bottom) Reflection coefficients for waves travelling from blood to various tissue types, expressed in dB (total reflection is 0 dB, no reflection is $-\infty$ dB).

4.4 Refraction

Refraction is the reason a stick looks bent in water; the light changes direction as it crosses the water-air boundary because of the difference in the speed of light in the two media. The same phenomenon occurs with sound. (Fig. 20 shows a sound wave changing direction as it crosses a boundary).

• **Huygens' principle** Given the position of a wavefront and the sound speeds, Huygens' principle (pronounced 'hoy-jens') allows us to construct the position of the wavefront a moment later. *Every point on the wavefront is considered to be a tiny source of spherical waves.* The spherical waves move at the local sound speed, so in time Δt the radii of the spherical waves are $c(x)\Delta t$, where $c(x)$ indicates that the sound speed depends on position. The 'envelope' of the spherical wavefronts gives the position of the next wavefront. (The small wavefronts used construct the new wavefront are sometimes referred to as 'wavelets'. These should not be confused with the wavelets used in linear algebra and to compress jpeg images.)

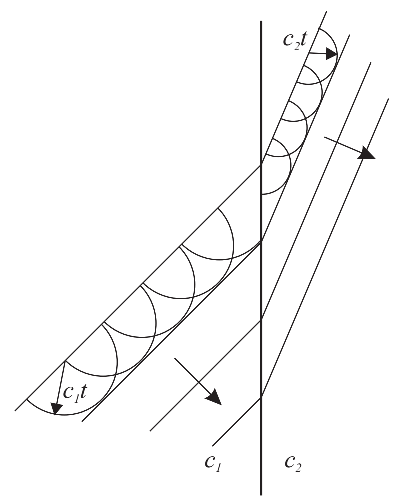


Figure 22: Explanation of refraction using Huygens' principle.

Note that the pressure across the boundary must be continuous (ie. the pressure on one side must be the same as on the other, *boundary condition (1)* in Section 4.2) so when drawing the wavefronts at their different angles they must meet at the boundary.

As a wave travels from one medium to another with a different sound speed it will change direction (and wavelength). Huygens' principle can be used to show in which direction the wave will travel in the new medium. Fig. 22 shows that when a wave travels into a medium with a *slower* sound speed the direction of the wave will move *closer to the normal*, ie. the boundary normal (also see Fig. 20). The opposite is true too: a wave moving into a faster medium will be bent away from the normal.

• **Snell's law** A consequence of Snell's law is that a wave entering a medium with a different sound speed will change direction. This can be understood using Huygens' principle.

- If $c_2 < c_1$ the direction of propagation moves *towards* the normal to the boundary.
- If $c_2 > c_1$ the direction of propagation moves *away from* the normal to the boundary.

• **Critical angle: total internal reflection** When the incidence angle is such that $\sin\theta_i = c_1/c_2$ then the transmitted wave will be travelling at 90° to the boundary. This is called the *critical angle*, and if the incidence angle becomes any larger, then $|R| = 1$ and the incident wave will be totally reflected.

4.5 Scattering and diffraction

In one way, scattering is the most important aspect of ultrasound imaging, because without it there would be no image at all. Scattering is a *nonlinear* phenomenon, in the sense that a wave scattered from one scatterer could then scatter from another scatterer, and again from the first scatterer and so on. Sometimes it is useful (because it makes the maths easier) to separate the total pressure field into a sum of the undisturbed incident field (the incident field that would exist if there were no scatterers), a component consisting of waves that have been scattered once, a component consisting of waves that have been scattered twice, ...

Scattering and diffraction are related and overlap in meaning, but roughly the terms are used as follows:

- **Scattering** refers to the reflection of sound from surfaces or heterogeneities in a medium. It is quite a general term and includes reflection and diffraction.

- **Diffraction** is usually used to refer to the 'leakage' of sound into 'shadow zones'. Diffraction is the reason you can hear someone talking in the next room even though you can't see them; the sound waves 'bend' around the corners more than light waves do as they have a much longer wavelength. (*Dif* fraction is quite different from *re*fraction and the two should not be confused.)

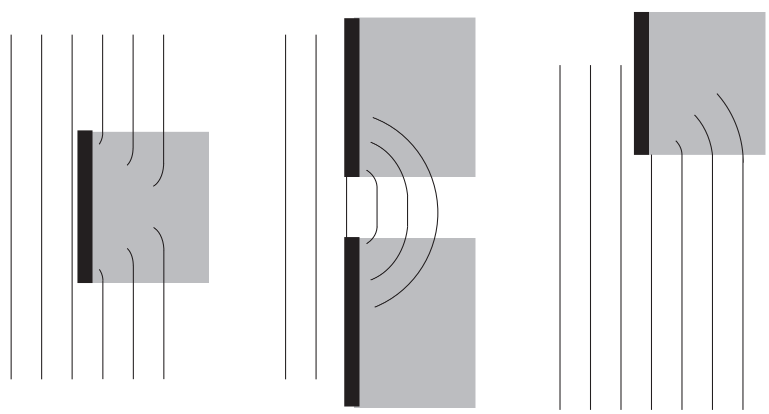


Figure 23: Three examples of diffraction, past an obstacle, through an aperture and at a corner. The *shadow zones*, into which the sound waves would not reach if they followed purely geometric, 'ray', trajectories are shown in grey. The sound waves moving into these shadow zones is one of the main characteristics of diffraction.

Scattering classification Scattering occurs on a wide range of scales, typically from $0.03\lambda - 300\lambda$. At different scales different effects dominate the scattering. A categorisation that divdes types of ultrasonic scattering roughly into five classes has been proposed. The middle three classes, 1-3 are the most important.

- **Class 0: molecular** Effects on a molecular scale; molecules bumping into each other, for instance. Some types of acoustic absorption (not usually considered as scattering) would fall into this category.

- **Class 1: diffusive** Small scatterer (or long-wavelength regime) in which $\lambda \gg a$ for a scatterer with characteristic size a, (one reference suggests $ka < 0.35$ where $k = 2\pi/\lambda$, giving $a \approx 20\ \mu$m). This class includes scatterers such as small clusters of living cells, blood cells or extra-cellular tissue matrix. This 'Rayleigh scattering' is weak, and the ratio of scattered to incident intensity $I_s/I_i \propto k^4 \propto \lambda^{-4}$. (For acoustic pressure the dependence is $p_s/p_i \propto k^2$.)

  Microbubble contrast agents are typically this sort of size and enhance the image by increasing the diffusive scattering in the regions in which they are located, eg. in a blood vessel.

- **Class 2: diffractive** $\lambda \approx a$, so $a \approx 0.1 - 1$ mm. This class includes small cysts or microcalcifications. Diffractive effects are important, so the scatter is anisotopic and there is variable frequency dependence.

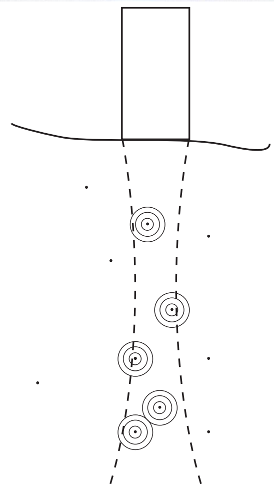
- **Class 3: specular** $\lambda \ll a$. Straightforward reflections as discussed in the previous sections. Large cysts, vessel or organ boundaries. There is no wavelength dependence.

- **Class 4: moving** This class is reserved for moving scatterers, so red blood cells could also fall into this category.

4.6 Image artefacts

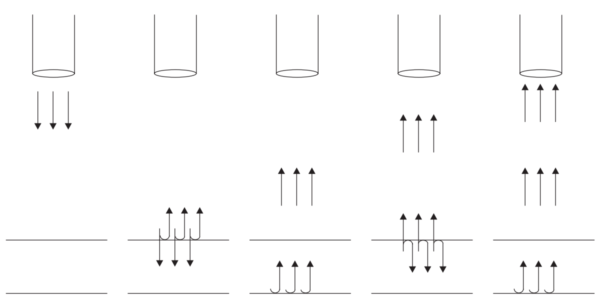
Some image artefacts are due to reflection and scattering within the tissue, so they are briefly mentioned here. On the whole, they are not desirable in that they degrade the image, but sometimes, if the cause of the artefact is understood it may provide some useful information about the tissue.

Speckle The noisy, textured, background in ultrasound images is called speckle and, while it is due to the micro-scatterers in the tissue, it does not carry any useful information about them.



- Speckle is due to constructive and destructive interference of the scattered sound waves from the tissue microstructure (and re-reflection of the scattered field from nearby surfaces).
- However, the actual image-texture of the speckle depends more on transducer response than the actual texture of the tissue, so it is not always representative of tissue microstructure and is not necessarily of diagnostic value.
- Speckle is not random but deterministic (move the transducer away and back again and the same speckle pattern returns), so it can be useful for tracking tissue motion, eg. distortion or strain.

Reverberation This refers to multiple reflections, such as when an ultrasound wave reflects back and forth between two layers, each time before returning to the transducer.



Acoustic shadows Despite diffraction, for large enough objects, the ultrasound waves are unable to travel into the shadow zone, so there will be nothing on the image at this point.

5 Acoustic Attenuation and Absorption

Up to now we have assumed (implicitly) that a plane wave with amplitude A at $t = 0$ would still have an amplitude A at some later time $t = \tau$, even though it would have moved a distance $c\tau$. Actually, ultrasound waves are partially absorbed as they travel along, ie. some of the energy is converted into heat, so their amplitude decreases.

- **Absorption** refers to the dissipation (conversion) of acoustic energy into heat.
- **Attenuation** is sometimes used to mean absorption, but it is also used to refer to any reduction in the amplitude of a sound wave, whether due to reflection, scattering, geometric spreading or dissipation as heat.

5.1 Mechanisms of absorption

In all of the modelling so far, we have considered the medium, eg. water or tissue, to be continuous (see Appendix A.1). While that remains true when absorption is taken into account, in the sense that bulk properties such as pressure may still be used to describe the sound field, we can no longer ignore the fact that the medium consists of molecules, which will bump into each other and which can store energy in several different ways. When molecular processes are included in a continuum model they are sometimes referred to as *internal* processes. Acoustic absorption can be explained in terms of internal processes.

There are several possible mechanisms by which the energy in ultrasonic waves can be dissipated as heat: viscous losses due to shear and compressive stresses, conversion of the wave energy into internal molecular energy states, or thermal conduction between regions of the wave at different temperatures (which is usually negligible at frequencies we are interest in).

• **Viscous absorption** Viscosity leads to shear forces within the tissue as the sound wave passes through, and some energy is converted, via friction to heat. More strictly, *shear viscosity* occurs due to diffusion of momentum, ie. molecules colliding, between regions of differing particle velocity. (*Bulk* or *volume* viscosity refers to losses during compression, when some of the energy is not stored as elastic potential energy but is converted into heat.) Viscosity contributes to absorption over a wide range of frequencies, but depends on frequency as ω^2, so is much more significant at higher frequencies:

$$\alpha_{\text{visc}} \approx \frac{\omega^2}{2\rho_0 c_0^3}\left(\frac{4}{3}\eta + \eta_B\right) \tag{87}$$

where η is the coefficient of shear viscosity and η_B the coefficient of bulk viscosity and ω is the frequency in radians/s.

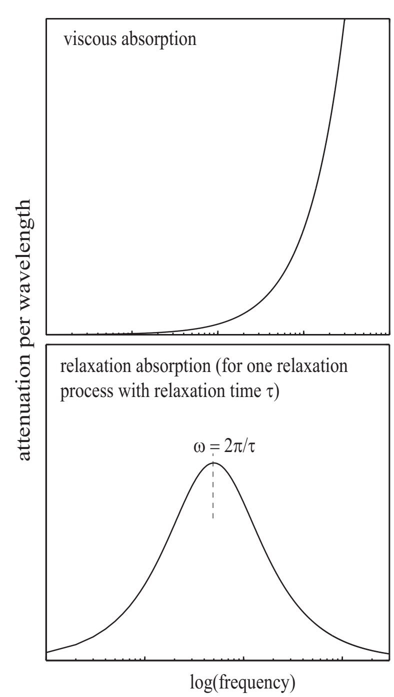


Figure 24: Characteristic curves for viscous and relaxation absorption.

• **Molecular relaxation absorption** The term *relaxation* (think 'languid', 'lagging behind' rather than 'calm' or 'tranquil') refers to any molecular process in a fluid that has some time delay associated with it, such as ionisation, evaporation, chemical reaction and transfer of energy between molecular modes. We consider only the last of these mechanisms here.

Imagine a small cube containing fluid - fluid made up of millions of individual molecules - that is compressed suddenly by moving the walls of the cube inwards. The density of the fluid increases instantly, as there are still the same number of molecules but they are taking up less space. The translational kinetic energy of the molecules also increases as they are knocked by the container walls as they move in, like table tennis balls off a bat. As the molecules will now hit the walls faster, the pressure increases too.

However, the molecules can store energy in a number of ways: translational kinetic energy, rotational kinetic energy (the molecules themselves spinning, or parts of them spinning), and vibrational energy (think of two parts of a molecule 'tied together' by nuclear forces which act as elastic thread allowing them to oscillate against each other), and the *equipartition theorem* states that at thermal equilibrium, the energy of the molecules will be shared equally among these various modes.

As the walls of the cube were moved abruptly, the fluid is currently not at thermal equilibrium, so some of the energy in the translational mode will be redistrbuted in the rotational and vibrational modes. As it is only the translational mode that affects the pressure, there will be a slight drop in pressure. The time that the energy takes to 'leak' from the translational to a rotational or vibrational mode is called the 'relaxation' time of that mode. Every possibly route (pathway) of energy redistribution will have its own relaxation time.

For a given pathway, and corresponding relaxation time, if the frequency of the acoustic wave is too high, then there will be no time for the relaxation to take place. Relaxation absorption is only significant - for any particular relaxation pathway - near frequencies for which the period of the acoustic wave ($1/f$) is close to the relaxation time. The *absorption per wavelength*, $\alpha_{\rm rel}\lambda$, for a pathway with a relaxation time of τ, is proportional to

$$\alpha_{\rm rel}\lambda \propto \frac{\tau f}{1+(\tau f)^2} \tag{88}$$

In practice, there is almost always more than one relaxation mechanism involved - and sometimes very many - so the total relaxation absorption will consist of a sum of several different relaxation curves.

5.2 Plane wave absorption

A harmonic plane wave, in the absence of absorption, can be written as $p_a(x,t) = A\exp(i(kx-\omega t))$. If the medium has an non-zero absorption coefficient, then the plane wave amplitude will be attenuated as it travels, so an additional exponential-decay factor is introduced

$$p_a(x,t) = Ae^{i(kx-\omega t)}e^{-\alpha x} \tag{89}$$

where α is called the **absorption coefficient** of the medium. After travelling a distance $x = 1/\alpha$, the amplitude of the plane wave will have fallen to $1/e \approx 0.37$ of its original value.

- **Units** Beware! Absorption coefficients are quoted in different units.
  - The absorption coefficient α, as described above, has units of $1/\text{cm} = \text{cm}^{-1}$. Sometimes α is written in Nepers/cm, Np/cm. Nepers are a (fairly pointless) dimensionless unit, as $1\ \text{Npcm}^{-1} = 1\ \text{cm}^{-1}$.
  - Often, the absorption coefficient is written as a decibel reduction (the minus sign is usually omitted):

$$\alpha_{\text{dB}} = 20\log_{10}\left(\frac{Ae^{-\alpha}}{A}\right) = (-)8.7\alpha\ \text{dB/cm} \tag{90}$$

- **Absorption and intensity** Intensity, in a plane wave, is proportional to p_a^2, and so the attenuation of the intensity has a factor of 2 in the exponent

$$\frac{I(\mathbf{x})}{I_{\text{initial}}} = \frac{(Ae^{-\alpha x})^2}{A^2} = \exp(-2\alpha x) \tag{91}$$

- **Complex wavenumber** The absorption coefficient can be lumped together with the wavenumber, to give a complex wavenumber $\hat{k}$. For example, eq. (89) can be rewritten as

$$p_a(x,t) = Ae^{i(kx-\omega t)}e^{-\alpha x} = Ae^{i(k+i\alpha)x-i\omega t} = Ae^{i\hat{k}x-i\omega t} \tag{92}$$

where the complex wavenumber is defined as

$$\hat{k} = k + i\alpha \tag{93}$$

- **Complex wavespeed** The dispersion relation $\omega = ck$ links the wavenumber and wave speed. Another way of including absorption is therefore using a complex wave speed:

$$\hat{c} = \frac{\omega}{k + i\alpha} \tag{94}$$

5.3 Pressure-density relations with absorption

• **Stokes' equation of state** Recall that in linear acoustics the sound speed relates the pressure and density changes via the linearised equation of state, $p_a = c^2 \rho_a$. Absorption can be incorporated into the equations of linear acoustics by adapting this equation, and many additional terms have been proposed over the years. For instance, Stokes proposed adding the following equation of state to describe absorption due to viscosity:

$$p_a = c_0^2 \rho_a + \frac{\zeta}{\rho_0} \frac{\partial \rho_a}{\partial t} \tag{95}$$

where ζ is the fluid viscosity. The effect of this term is to introduce a time delay between the pressure and the density, a lag. Why does this lag lead to, or signify, absorption?

• **Pressure-density phase delay** Before answering that question directly, it is informative to note that writing the sound speed as complex, as described above, also introduces a time delay (a difference in phase) between the density and the pressure. If c is complex then c^2 is also complex, and - writing c^2 in magnitude and phase form - the (linear) pressure-density relation becomes:

$$p_a = |c^2| e^{i\phi} \rho_a \tag{96}$$

where ϕ is a *phase* term. The physical significance of ϕ can be seen by writing the density ρ_a as a harmonic plane wave, $\rho_a = Ae^{i(kx-\omega t)}$

$$p_a = |c^2| A e^{i\phi} e^{i(kx-\omega t)} = |c^2| A e^{i(kx-\omega t+\phi)} \tag{97}$$

The acoustic pressure, p_a, and density fluctuations are *no longer in phase*. There is a time delay between them. On the pressure-density graph this phase delay makes the straight-line relationship between p_a and ρ_a shown in Fig. 9 become elliptical (like a Lissajous figure on an oscilloscope):

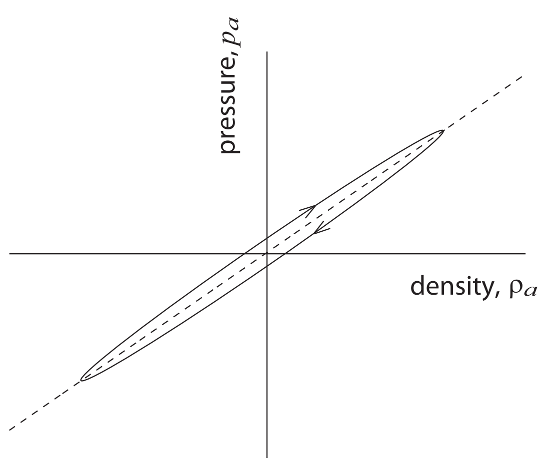


Figure 25: Pressure-density relation showing the effect of absorption.

• **Pressure leads density?** This difference in phase between the acoustic pressure and density is often described as the changes in density lagging behind - following after some delay - the changes in acoustic pressure. How does this fit with our example above of a cube of fluid undergoing compression, so the density changes instantaneously, and the equilibrium pressure is reached only after a lag. It sounds - at first glance - as though the density changes leads the pressure changes, not the other way around.
To see why the pressure is said to lead the density consider a step change in density (as in the cube example above). In this case, a pressure increase occurs at the same time as the density increase, and then - after a delay - the pressure reduces by a little. Now consider a series of step changes in density making up a rough approximation to a sine wave, the thinner solid curve in Fig. 26. At each step, the acoustic pressure, the thicker curve, will (a) increase in proportion to the change in density, but also (b) decrease a little because of the relaxation loss from the previous step. The effect of this is shown in Fig. 26 and it is clear that the pressure curve is *leading* the density curve, i.e. the pressure curve reaches its peaks earlier in time than the density curve.

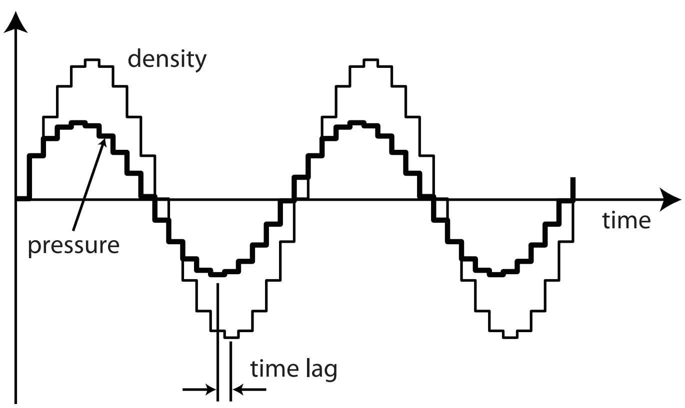


Figure 26: When absorption is taken into account, acoustic pressure changes lead the density changes, for a sinusoidal excitation.

5.4 Absorption in tissue

The frequency dependence of the absorption coefficients of many substances empirically follow power laws:

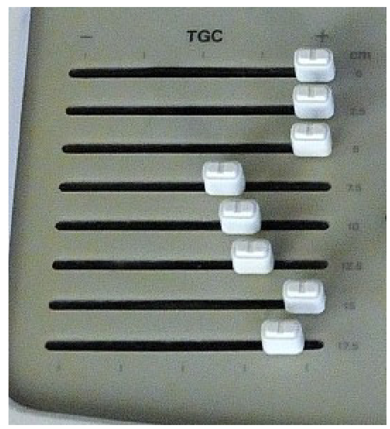
$$\alpha \approx \alpha_0 f^y \text{ Np/cm} \tag{98}$$

In decibels this becomes

$$\alpha_{\mathrm{dB}} \approx 20 \log_{10}(\exp(-\alpha_0 f^y)) = 8.7\alpha_0 f^y \text{ dB/cm} \tag{99}$$

and is often quoted as $(\alpha_{\mathrm{dB}}/f^y)$ in units of dB/cm/MHz^y.

Time-gain compensation Because the amount of the signal that has been absorbed is greater the longer the signal has travelled, ultrasound pulses reflecting back from deep within the tissue are much more strongly attenuated than pulses reflected from structures close to the surface. The image, if displayed without any adjustment for this, the contrast in the image between the surface and several cms deep would be great and the image would be unrepresentative of the tissue. *Time-gain compensation* is an approximate way to correct for this attenuation. It is a set of gain controls (see right), with each slider controlling gain of a band of the image around a certain depth eg. 0-5mm, 5-10mm, etc. It is a crude but practical way to account for absorption in real-time imaging.



tissue type	absorption coefficient (dB/cm/MHz^y)	frequency dependence y
water at 20°C	2.17×10^{-3}	2
blood	0.14	1.21
brain	0.58	1.3
breast	0.75	1.5
fat	0.6	1
heart	0.52	1
kidney	10	2
liver	0.45	1.05
muscle	0.57	1
spleen	0.4	1.3
bone	3.54	0.9

Figure 27: Absorption coefficients for various tissue types, including frequency dependence.

6 Nonlinear Acoustics

So far, we have been discussing acoustics in the *linear* regime, which has the following characteristics:

- The shape of a sound wave remains the same as it propagates (in the absence of absorption).
- Two sounds waves can pass through each other without interacting.
- The total acoustic pressure due to two sound waves is the sum of the individual acoustic pressures (superposition).
- The sound speed is a property of the medium and does not depend on the sound itself.

For large amplitude waves, none of these useful properties are necessarily true, and the sound propagation is described as *nonlinear*.

Nonlinear effects that cannot occur in linear acoustics include

- Steepening of waves and shock formation during propagation
- Harmonic generation in propagating waves
- Radiation pressure
- Streaming
- Cavitation

Complex notation Another difference between the linear and nonlinear regimes is that complex notation, e.g. $\exp(i\omega t)$, cannot be used in the nonlinear case. The reason is simply that raising the real part of a complex exponential to some power is different from taking the real part of a complex exponential raised to some power. For example:

$$\left(\mathrm{Re}\left(e^{i\omega t}\right)\right)^3 = \frac{1}{4}\left(3\cos(\omega t) + \cos(3\omega t)\right) \tag{100}$$

$$\neq \mathrm{Re}\left(\left(e^{i\omega t}\right)^3\right) = \mathrm{Re}\left(e^{3i\omega t}\right) = \cos(3\omega t) \tag{101}$$

Complex notation is one of several mathematical tools that are only useful when discussing linear propagation, including superposition, Green's functions (not covered in this course), plane wave decomposition, etc.

6.1 Sound speed increases with pressure

The speed at which the waves propagate, the sound speed c_0, has so far been assumed to be a property of the medium, and related to the ambient properties of that medium, p_0 and ρ_0. The story so far has been that different tissue types have different sound speeds, so the ultrasound waves travel at different speeds in different tissues, but *for a given medium* all ultrasound waves, whatever their frequency, shape or amplitude, will travel at the same speed.

In fact, this is not quite true for large amplitude waves, such as those used in diagnostic and, especially, therapeutic ultrasound, as the speed of the wave depends on the total pressure $p = p_0 + p_a$ and not just on its ambient value, so when the acoustic pressure p_a is large compared to p_0 the wave speed can be different at different parts of the wave. There are two main mechanisms that lead to this phenomenon, which were removed in Section 3 by linearising. The linearisations were

1. **Assumption of low Mach number** Consider a plane wave, with particle velocity u_a, moving in the positive x-direction. If $|u_a|$ is a significant proportion of c_0, then - because the pressure and particle velocity are often in phase (see Fig. 15) - the parts of the wave at higher pressure move with speed $c_0 + |u_a|$ and those with lower pressure at $c_0 - |u_a|$. The sound speed will therefore be faster at higher pressures. This is sometimes called a 'convective nonlinearity'.

$$\textbf{Convective nonlinearity}\text{: wave propagation velocity, } c = c_0 + u_a \tag{102}$$

2. **Linearisation of the equation of state** The use of the constant sound speed c_0 is equivalent to assuming a linearised equation of state (see Fig. 9). In practice the pressure-density relation is nonlinear, ie. when the fluid is compressed sufficiently its stiffness (bulk modulus) will increase. The sound speed will therefore be faster at higher pressures (as above, but by a different mechanism). This is sometimes called a 'material nonlinearity'.

   Nonlinear pressure-density relation In Section 3.3 a Taylor series expansion of the equation of state was used to obtain a linear pressure-density relation, $p_a = c_0^2 \rho_a$. By keeping the next term in the Taylor series expansion, a 'nonlinear pressure-density relation', or 'nonlinear equation of state', can be obtained:

$$p_a = A\left(\frac{\rho_a}{\rho_0}\right) + \frac{B}{2!}\left(\frac{\rho_a}{\rho_0}\right)^2 + \dots \tag{103}$$

   where we recall the definitions of A and B as

$$A \equiv \rho_0 \left(\frac{\partial p}{\partial \rho}\right)_{s,0} = \rho_0 c_0^2, \qquad B \equiv \rho_0^2 \left(\frac{\partial^2 p}{\partial \rho^2}\right)_{s,0}. \tag{104}$$

   The ratio B/A is called the **nonlinearity parameter** and is a fundamental measure of nonlinearity in acoustics. (Slightly unusually, there is no symbol for this ratio other than 'B/A' and so it is called simply "the nonlinearity parameter, B over A".)

Substituting the nonlinear pressure-density relation, Eq. (103), into the definition of the sound speed, Eq. (33), gives:

$$c = \sqrt{\left.\frac{\partial p}{\partial \rho}\right|_s} = \sqrt{\frac{A}{\rho_0} + \frac{B}{\rho_0}\left(\frac{\rho_a}{\rho_0}\right) + \ldots} \tag{105}$$

so, because $c_0^2 = A/\rho_0$,

$$\frac{c}{c_0} = \sqrt{1 + \frac{B}{A}\left(\frac{\rho_a}{\rho_0}\right) + \ldots} \tag{106}$$

$$\approx 1 + \frac{B}{2A}\left(\frac{\rho_a}{\rho_0}\right) + \ldots \tag{107}$$

where the last step used a binomial expansion.[1] Linear acoustics says that for a plane wave $\rho_a/\rho_0 \approx u_a/c_0$, so the material nonlinearity can be written as

$$\textbf{Material nonlinearity}: \text{ wave propagation velocity, } c \approx c_0 + \left(\frac{B}{2A}\right) u_a \tag{108}$$

- **Coefficient of nonlinearity** Combining the *convective* and *material* nonlinearities gives:

$$c \approx c_0 + u_a + \left(\frac{B}{2A}\right) u_a \tag{109}$$

$$\approx c_0 + \beta u_a \tag{110}$$

where the *coefficient of nonlinearity* is defined as

$$\beta \equiv 1 + \frac{B}{2A} \tag{111}$$

[1]Binomial expansion: $(1+x)^n = 1 + nx + n(n-1)x^2/2! + \ldots$ for any n when $|x| < 1$.

6.2 Nonlinear propagation and wave steepening

The *convective* and *material* nonlinearities both act to increase the propagation speed of the higher pressure part of the wave, and slow the speed of the lower pressure part. Each peak propagating nonlinearly will therefore gradually catch up with the trough in front of it, eventually developing into a shock. The sequence of events, pictured in Fig. 28, is as follows:

1. A sinusoidal wave starts to distort as peaks catch troughs,
2. Wave develops into a shock wave,
3. This steepening has moved some of the energy in the wave to higher frequencies,
4. The higher frequencies are more strongly absorbed,
5. The shock is dissipated and the wave and ends up as a low amplitude sine wave ('old age').

• **Shock parameter** A number which is useful when trying to calculate whether or not a significant amount of distortion has occurred or not is the *shock parameter*, σ (unhelpfully, this is sometimes also called the 'nonlinearity parameter').

$$\sigma = \frac{\beta \omega z p_{\mathrm{source}}}{\rho_0 c_0^3} \tag{112}$$

where z is the distance the wave has travelled and p_{source} is the acoustic pressure at the source. In Fig. 28 the value of σ has reached 1 by the third box down, and will go higher as the wave steepens further before decaying into 'old age' as shown.

• **Harmonic generation** The frequency spectrum, $P_a(f)$, of a wave, $p_a(t)$, is given by the Fourier transform $P_a(f) = \{p_a(t)\}$. If a wave changes shape, so $p_a(t)$ changes, then its Fourier transform will change. If the signal is originally a sine wave with a frequency f_0, then during nonlinear propagation, energy will be shifted into multiples of this fundamental frequency, nf_0, where n is a positive integer. The frequency content of the wave at each stage is shown in Fig. 28.

Why is the energy shifted to multiples of the fundamental frequency? The appearance of harmonics, integer multiples of the fundamental frequency, is a classic effect of nonlinearity that appears in many types of systems, not just acoustic. For a very simple example, consider a sine wave, $\sin(\omega t)$ that is passed through a system whose output is the *square* of the input. The output will be $\sin^2(\omega t)$, which using trigonometric identities becomes $\sin^2(\omega t) = \frac{1}{2} - \frac{1}{2}\cos(2\omega t)$, in which a component at 2ω has appeared.

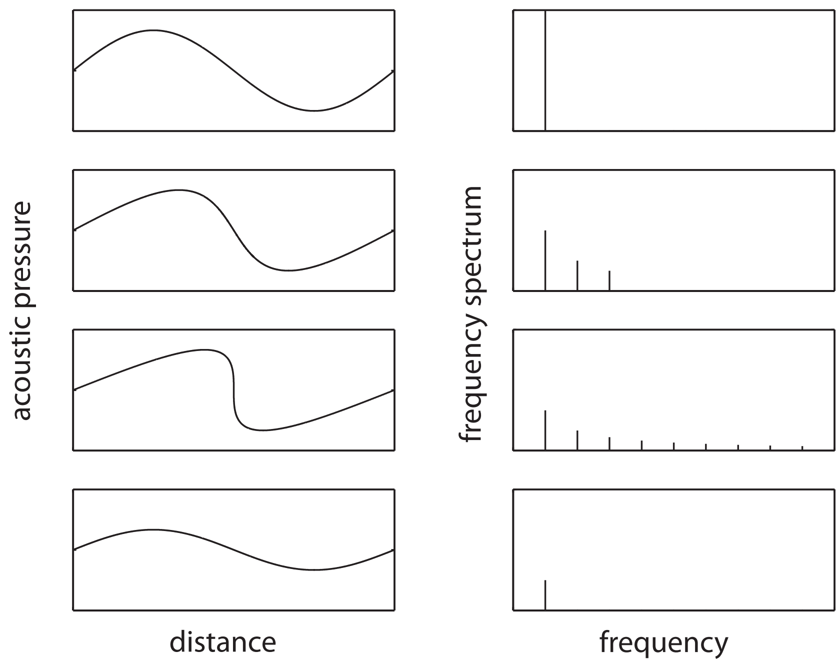


Figure 28: Wave steepening due to nonlinear propagation and consequent harmonic generation: (from the top) (1) initial large amplitude single frequecy wave, (2) steepening of the wave as the higher pressure parts of the wave travel faster, (3) shock formation, (4) sinusoidal 'old age' due to the strong absorption of the higher frequencies.

6.3 Tissue harmonic imaging

Many commerical ultrasound machines now allow users to do 'harmonic imaging', also called 'tissue harmonic imaging' and 'second harmonic imaging', to obtained better image resolution, i.e. finer detail is clear in the image. This imaging modality arose by accident, but to understand how we must briefly discuss ultrasound contrast agents.

Ultrasound contrast agents are gas filled microspheres (microbubbles, diameter 1-10 μm) which have a shell, made of lipid (fat) or a polymer, to prevent the gas dissolving in the tissue. They are typically injected into the blood stream and work by strongly reflecting the ultrasonic waves because of the large difference in characteristic acoustic impedance, $\rho_0 c_0$, between the gas in their interior and the tissue outside. As well as reflecting the ultrasound, the bubbles are also 'excited' and vibrate. This vibration is nonlinear, because the bubble is stiffer under compression than it is under expansion, and so harmonics - multiples of the driving frequency - are generated. When ultrasound contrast agents first became available manufacturers of the imaging machines allowed users to filter out the reflections at the frequency of the incident pulse (the usual ones used to form a normal ultrasound image) so that the nonlinear, higher frequency, signals from the microbubbles could be seen on their own.

Nonlinear propagation However, ultrasonographers found that they could get images of the tissue *even when they didn't use the microbubble contrast agents*, so something else was behaving nonlinearly and shifting energy from the driving frequency into harmonics. It turned out that it was the tissue itself that was behaving nonlinearly and nonlinear propagation of the incident signal was generating the harmonics (as described in the previous section).

Improved image resolution A 'tissue harmonic' image often has better resolution than a conventional ultrasound image for two reasons: (1) the harmonics are at higher frequencies, and so shorter wavelengths, (2) the harmonics are generated within the tissue, rather than being sent in by the source, so they have less tissue to travel through to get back to the transducer and will therefore be absorbed less.

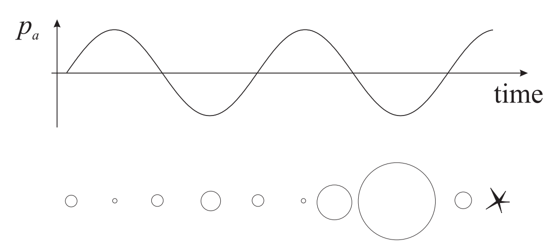
6.4 Inertial cavitation

The next most important effect that does not fit within the linear framework is cavitation. The word 'cavitation' has two slightly different, but closely related, meanings: the *formation and growth* of gas bubbles in a liquid in response to a pressure change, and their *collapse*.

Negative pressure The total pressure was described earlier as the sum of a ambient and an acoustic part, $p = p_0 + p_a$. If the pressure fluctuations in the acoustic wave are of very high amplitude, then it is plausible that $p_a > p_0$. In a gas this would be unphysical as a negative pressure makes no sense, but in a liquid it does. A (positive) pressure in a liquid comes from compressing it. The ambient pressure at a point comes from the liquid above that point it squashing it down, eg. the atmospheric pressure experienced by someone is due to the weight of the column of air above them. A liquid, unlike a gas, can also be stretched - it can support a tensile force - because of the intermolecular bonds formed by the closely packed liquid molecules. A tensile force is the opposite of compressive, and so is called a negative pressure.

Cavitation inception Once a certain negative pressure is reached, the bonds between the water molecules are not strong enough to keep the liquid together, and it 'tears apart', producing a void or cavity - a gas bubble - which is filled with liquid vapour and other dissolved gases that come out of solution. This is called 'cavitation' and can occur at narrowing in pipes where the pressure falls due to the flow speeding up, and at the tips of ships' propellers when the speed of the propeller in the water reduces the pressure low enough.

Stable cavitation A gas bubble in an ultrasound field will resonate - its radius will continually oscillate from small to large and back. This can be likened, to some extent, to a mass on a spring, expect here the 'spring' is provided by the gas and the mass is provided by the fluid surrounding the bubble. During the compressive cycle of the oscillation, as the radius becomes very small, the temperature of the gas in the bubble can reach very high temperatures (1000s K, sun surface ≈6000 K) and emit picosecond pulses of light (sonoluminescence). 'Stable' cavitation is when such oscillations continue indefinitely, emitting a pulse of light on each compressive pulse.



Inertial cavitation It is this type of cavitation that has consequences for ultrasound in medicine. In inertial caviation a gas bubble grows during each cycle as more gas diffuses into the bubble during the expansion phase than diffuses out during the compressive phase. This is known as 'rectified diffusion'. When the bubble reaches a certian size, it undergoes inertial collapse - rapid reduction in radius started by the pressure change in the liquid by accelerated by the mass of the liquid - the inertia - around the bubble. The bubble usually breaks up. Because of the inertia associated with this form of collapse, it is violent and can cause damage to surrounding tissue or boundaries. Cavitation close to a boundary often takes the form of 'jetting' where the bubble collapses asymmetrically, smashing a 'jet' of fluid into the boundary.

6.5 Radiation force and acoustic streaming

The nonlinear effects that are most important in ultrasound in medicine are the nonlinear propagation and inertial cavitation described above. There are, however, other effects of ultrasound which don't fit into the linear picture and yet which are worth knowing about, if only for a broader understanding of the sorts of effects ultrasound can have.

Radiation force If a plane wave, with acoustic intensity I in the direction of propagation, reaches a wall (perpendicular to the propagation direction) and is *completely absorbed by the wall*, the wall experiences a force per unit area, a 'radiation pressure', of

$$p_{\text{rad,abs}} = \frac{I_{\text{av}}}{c_0} = \frac{1}{2}\frac{p_a^2}{\rho_0 c_0^2} \quad \text{Pa} \tag{113}$$

where p_a is the maximum acoustic wave amplitude. Note that P_{rad} is a pressure, so is written in units of Pascals. This effect can be used to measure the power output of an ultrasound transducer using an 'ultrasonic power balance'. For an acoustic beam with power W, the force on a surface larger than the beam

$$F_{\text{rad}} = \frac{W}{c_0} \quad \text{N} \tag{114}$$

In water, the force is 0.69 μN/mW. If the wave is reflected from the wall, twice as much force is experienced by the wall:

$$p_{\text{rad,ref}} = \frac{2I_{\text{av}}}{c_0} = \frac{p_a^2}{\rho_0 c_0^2} \quad \text{Pa} \tag{115}$$

Acoustic streaming When a plane wave is attenuated, the acoustic pressure may take the form $p_a = A\exp(i(kx - \omega t))\exp(-\alpha x)$ where α is the absorpion coefficient. If a thin (imaginary) absorbing surface were placed in the flow at different values of x we can see that the acoustic pressure at the surface, and therefore the radiation pressure on the surface, would decrease as x got larger. In other words, in a decaying acoustic field there is a 'radiation pressure *gradient*'.

$$\text{Force per unit volume} = -\frac{dP_{\text{rad}}}{dx} \approx \frac{1}{c_0}\frac{dI_{\text{av}}}{dx} \approx \frac{-2\alpha I_{\text{av}}}{c_0} \tag{116}$$

A pressure gradient is a force, so there is a force on the fluid in the positive x-direction that causes it to flow. This acoustically-induced flow is called 'acoustic streaming'. Streaming therefore only occurs when the wave is being attenuated and the medium can flow, which is not always true of soft tissue. Acoustic streaming is sometimes called 'quartz wind', as the early ultrasound transducers were quartz crystals, or 'sonic wind'.

7 Bioeffects: Safety and Therapeutic Ultrasound

Ultrasonic bioeffects, the ways in which ultrasound can *change* tissue, fall into two categories: thermal (heating) and non-thermal (principally cavitation). As far as ultrasound imaging is concerned, any process that might alter or change the tissue in any way must be carefully examined to see whether or not it might cause damage - obviously an undesirable side-effect for an imaging modality. On the other hand, if a process involving ultrasound can affect the tisue *in a controllable way*, then it may be possible that it could be used for surgery or other procedures or therapies in which a controlled change in the tissue is desirable (such as stopping blood flow or killing tumour cells).

7.1 Mechanical and thermal indices

To give the user of an US imaging machine some idea as to the risk of damage from mechanical or thermal effects, two indices are required by law to be displayed on the US monitor: mechanical index, MI, and thermal index, TI.

• **Mechanical Index, MI** The mechanical index is designed to give a rough idea of the risk of mechanical effects during insonification. Because the main risk is due to inertial cavitation, the definition of MI is proportion to the peak *rarefactional* pressure, ie. peak negative pressure, P_r (in MPa):

$$\mathrm{MI} = \frac{P_r}{\sqrt{f}} \tag{117}$$

where f is the working frequency in MHz. (How to estimate P_r in practice is given in the *draft* British Standard *BS EN 62359 "Ultrasonics - Field Characterization - Test methods for the determination of thermal and mechanical indices related to medical diagnostic ultrasound fields"*.

The British Medical Ultrasound Society (www.bmus.org) give the following guidelines:

> **MI > 0.3** There is a possibility of minor damage to neonatal lung or intestine. If such exposure is necessary, try to reduce the exposure time as much as possible.
>
> **MI > 0.7** There is a risk of cavitation if an ultrasound contrast agent containing gas micro-spheres is being used. There is a theoretical risk of cavitation without the presence of ultrasound contrast agents. The risk increases with MI values above this threshold.

• **Thermal Index, TI** In diganostic ultrasound, the image quality is (obviously) important. Using higher intensities can improve the signal-to-noise ratio (SNR) and using higher frequencies, because they have shorter wavelengths, can improve the spatial resolution - the 'sharpness' of the image. Unfortunately, because of the increased ultrasound absorption at high frequencies, both higher frequencies and higher intensities will lead to increased heating in the tissue. Heat can damage tissue (see section below), so it is clearly important that thermal effects must be kept to a minimum for safety reasons. A *trade-off* is therefore necessary between the need for a better quality image and the risk of thermal damage. This is, ultimately, a decision made by the ultrasonography (the user of the US scanner), but to aid them, all US machines display the 'thermal index', TI, which gives an estimate of the risk to thermal damage. TI is defined as

$$\mathrm{TI} = \frac{P_p}{P_{\mathrm{deg}}} \tag{118}$$

where P_p, the 'power parameter' is related to the power input to the tissue from the ultrasound field, and P_{deg} is the power required to raise the tissue by 1 K. Because both of these quantities will depend on the particular situation and tissue type, there are different definitions of them, depending on whether the tissue is *soft tissue*, this index is called TIS, *bone at the surface*, TIC, or *bone below the surface*, TIB. There is also a difference depending on whether the US probe is scanned across the surface or held still (non-scanning mode). The specifics of the definition are described in the British Standard mentioned above. Among other guidelines, the BMUS give the following:

TI > 0.7 The overall exposure time (including pauses) of an embryo or fetus should be restricted in accordance with the following Table:

TI	0.7	1.0	1.5	2.0	2.5
fetal exposure (s)	60	30	15	4	1

TI > 1.0 Eye scanning is not recommended, other than as part of a fetal scan.

TI > 3.0 Scanning of an embryo or fetus is not recommended, however briefly.

7.2 Safety statement

The following statement was issued by the council of the British Medical Ultrasound Society (BMUS) in 2000 and reconfirmed in 2007.

Statement on the Safe Use and Potential Hazards of Diagnostic Ultrasound

"Ultrasound is now accepted as being of considerable diagnostic value. There is no evidence that diagnostic ultrasound has produced any harm to patients in the four decades that it has been in use. However, the acoustic output of modern equipment is generally much greater than that of the early equipment and, in view of the continuing progress in equipment design and applications, outputs may be expected to continue to be subject to change. Also, investigations into the possibility of subtle or transient effects are still at an early stage. Consequently diagnostic ultrasound can only be considered safe if used prudently.

Thermal hazard exists with some diagnostic ultrasound equipment, if used imprudently. A temperature elevation of less than 1.5°C is considered to present no hazard to human or animal tissue, including a human embryo or fetus, even if maintained indefinitely. Temperature elevations in excess of this may cause harm, depending on the time for which they are maintained. A temperature elevation of 4°C, maintained for 5 minutes or more, is considered to be potentially hazardous to a fetus or embryo. Some diagnostic ultrasound equipment, operating in spectral pulsed Doppler mode, can produce temperature rises in excess of 4°C in bone, with an associated risk of high temperatures being produced in adjacent soft tissues by conduction. With some machines colour Doppler imaging modes may also produce high temperature rises, particularly if a deep focus or a narrow colour box is selected. In other modes, temperature elevations in excess of 1°C are possible, but are unlikely to reach 1.5°C with equipment currently in clinical use, except where significant self-heating of the transducer occurs.

Non-thermal damage has been demonstrated in animal tissues containing gas pockets, such as lung and intestine, using diagnostic levels of ultrasound (mechanical index values of 0.3 or more). In view of this, it is recommended that care should be taken to avoid unnecessary exposure of neonatal lung , and to maintain MI as low as possible when this is not possible. In other tissues there is no evidence that diagnostic ultrasound produces non-thermal damage, in the absence of gas-filled contrast agents. However, in view of the difficulty of demonstrating small, localised, regions of damage in vivo, the possibility of this cannot be excluded. The Mechanical Index, if displayed, acts as a guide to the operator. The use of contrast agents in the form of stabilised gas bubbles increases the probability of cavitation. Single beam modes (A-mode, M-mode and spectral pulsed Doppler) have a greater potential for non-thermal hazard than scanned modes (B-mode, Colour Doppler), although the use of a narrow write-zoom box increases this potential for scanning modes."

7.3 Thermal effects

Ultrasonic heating can cause increases in temperature of several tens of degrees K, such as in the case of high-intensity focussed ultrasound (HIFU) therapy. On the other hand, for diagnostic applications, the amount of heating is usually very low. The key acoustic field quantity when considering heating effects is the acoustic power deposited in the tissue through absorption. In an absorbing medium, acoustic energy is removed from a wave and deposited as heat at a rate per unit volume Q, given by

$$Q = 2\alpha|\mathbf{I}| \quad \mathrm{W/cm^3} = \mathrm{J/s/cm}^3 \tag{119}$$

where $|\mathbf{I}|$, in W/cm^2, is the magnitude of the acoustic intensity vector and α (cm^{-1}) is the absorption coefficient. The rate of change of the temperature due to this rate of heating depends on the *specific heat capacity* of the tissue (J/kg/K):

$$\frac{\partial T}{\partial t} = \frac{Q}{\rho_0 C_p} = \frac{2\alpha|\mathbf{I}|}{\rho_0 C_p} \quad \mathrm{K/s} \tag{120}$$

(This follows from the definition of the specific heat capacity as the amount of energy in Joules required to raise 1 kg of the medium by 1 K.) The heat will then be conducted through the tissue, and so become redistributed according to the heat diffusion equation:

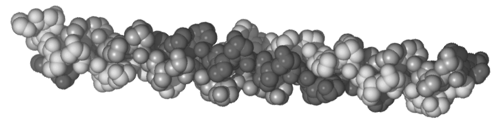
$$\frac{\partial T}{\partial t} - D\nabla^2 T = \frac{Q}{\rho_0 C_p} \tag{121}$$

where D, in m^2/s, is the thermal diffusivity of the tissue.

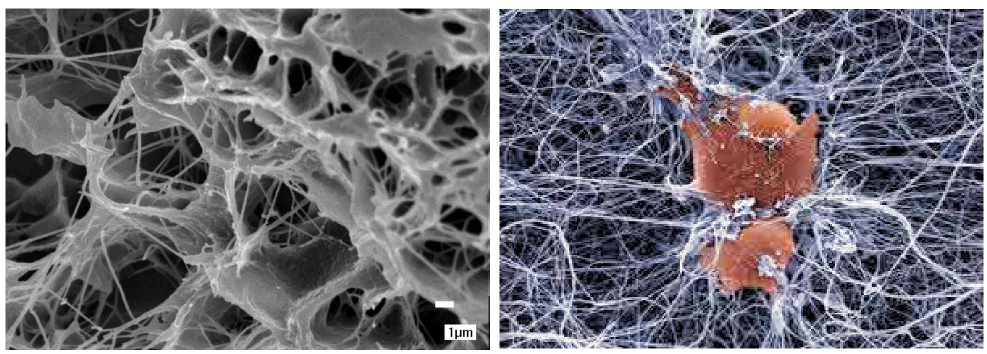
Pennes' bio-heat equation When the movement of the blood (perfusion) needs to be taken into account (some of the heat is carried the blood through vessels etc.), there is an additional term in the heat equation, and the equation is known as Pennes' bio-heat equation:

$$\frac{\partial T}{\partial t} - D\nabla^2 T = \hat{Q} - Q_p \tag{122}$$

where $Q/\rho C_p$ has been rewritten as $\hat{Q}$, and $Q_p = P\rho_b C_b(T - T_b)$ is the heat removed by perfusion, ρ_b and C_b are the density and specific heat capacities of blood, $(T - T_b)$ is the temperature difference between the tissue and the blood flowing through it, and P is the perfusion, the volume of blood flowing through unit volume of tissue per second. (Sometimes $P\rho_b$ is called the perfusion, which is the *mass* of blood per unit volume per second.)



Tissue composition The two main constituents of tissue, as far as ultrasound bioeffects are concerned, are *water* (mostly in cells) and *extra-cellular matrix*, ECM, the collagen 'scaffold' which the cells sit on or within (pictured below). The ratio of the amount of ECM and the amount of cells depends on the type of tissue: bone is high in ECM, liver has much less ECM holding it together. ECM is largely constructed from tri-helical collagen fibres, that begin to unravel (gelatinise, coagulate) well below 100°C.



Thermal tissue damage Irreversible tissue damage depends on the time at which the tissue is held at a given temperature. eg. 63% of epidermal tissue will die if it is held at 45°C for 9 hours, but the same amount of damage will occur in about 1 second at 60°C. In other words, the amount of damage, at a point in the tissue $\mathbf{x}$, depends both on the temperature history of the tissue $T(\mathbf{x}, t)$ and on the type of tissue. Hyperthermia is a general term used to describe tissue damage due to heating to temperatures below 100°C.

- $\sim$ 37°C is normal body temperature, and for the first 1.5°C or so of heating no irreversible changes occur.
- At $\sim$ 41°C, cell proteins - both membrane and cytoplasmic proteins - start to undergo conformational (shape) changes. They change conformation because the hydrogen bonds keeping them in their native state are broken by the increasingly violent vibrations of the molecule as the temperature increases. When a protein molecule changes shape it can often no longer fulfill its function within the cell. For instance, when enzymes, whose catalytic functions depend crucially on their shape, begin to deform, reaction rates within cells slow down. Even at small increases in temperature some cells will die because of these effects. The rate at which cells necrose or apoptose increases with the temperature.
- From $\sim$ 45°C the collagen fibres forming the ECM begin to shrink as the collagen's tri-helical structure breaks apart. The collagen softens and gelatinises. (Gelatin is just tangled, random coils of collagen.) Tissue starts to coagulate.

7.4 Therapeutic and Surgical Ultrasound

These notes are mainly concerned with ultrasound *imaging*, and the use of ultrasound for surgery or therapeutic applications doesn't strictly fit into that category. However, to give a more complete picture of the sorts of things ultrasound is used for, a brief summary of some non-imaging medical uses for ultrasound is given in this section.

There are numerous uses for ultrasound and ultrasonically-vibrating devices in medicine and related fields, and no attempt has been made to include them all. Many fall neatly into two groups: *low intensity* therapeutic applications, which include bone healing, and those that use *high intensity* ultrasound, such as lithotripsy. Other procedures that could have been discussed include ultrasonic dental cleaners and physiotherapy.

sound source	frequency range (MHz)	acoustic pressure (MPa)	acoustic intensity (W/cm^2)
lithotripsy	0.5	>10	low
HIFU	0.8-2	10	400-10000
Doppler ultrasound	1-10	0.5-5	0.1-10
physiotherapy ultrasound	1-3	<0.5	3
B-mode ultrasound	1-15	0.5-5	<0.1-1
fracture healing	1.5	< 2	0.03-0.5

Figure 29: Indicative values for various medical ultrasound modalities.

7.5 Low power therapeutic applications

The predominant mechanisms responsible for tissue effects at low ultrasonic intensities are more likely to be non-thermal than thermal. Non-thermal mechanisms could include increased movement of molecules due to cavitation, jetting, streaming or micro-streaming around microbubbles, or activation of cells' mechanoreceptors.

• **Fracture healing** There is no clear consensus on whether ultrasound can increase the rate of bone healing and repair, but there is some evidence to suggest that it might have a beneficial effect. One commercial device, Exogen®, intended for use in treating fracture non-unions (when a fracture has failed to heal after several months), is available. It emits low intensity ultrasound (lower than diagnostic levels), in 200 μs pulses at 1.5 MHz at a pulse repetition frequency of 1 kHz. One trial of 67 patients (although with no control group) suggests that it may speed up fracture repair. At this frequency and low intensity the mechanism is unlikely to be thermal, and it has been suggested that a cellular mechanism via mechano-receptors may be at work. Any cell membrane protein that can detect the mechanical environment around the cell (pressure, flow, shearing, etc) is called a 'mechano-receptor'. Integrins are surface proteins, one end of which secures the cells to the collagen extra-cellular matrix (ECM) while the other is attached to the cytoskelton inside the cell. They are mechano-receptors in the sense that when the ECM distorts - such as when an ultrasound wave passes through - they may change configuration and cause something to happen inside the cell. One possibility, then, is that this mechano-transduction pathway initiates an intracellular mechanism to speed up bone repair.

• **Sonophoresis** Many drugs are applied through the skin, typically in the form of a cream or patch. Drugs with large molecular weights penetrate the skin less easily than small molecules, and so can have a reduced uptake. By insonifying the site at which the drug is applied can increase this uptake, but the mechanisms by which this occurs are not fully clear. Local heating may be a reason when high frequencies are used, and cavitation temporarily increasing the permeability of the *stratum corneum* (the outer layer of the epidermis) is another possibility.

• **Sonoporation** In sonoporation, cell membranes are temporarily altered to allow large molecules, that would not normally be able to cross the membrane, to pass into the cell. The mechanism, as with many of these therapies, is poorly understood, but is more likely mechanical (including inertial cavitation, microstreaming and jetting) than thermal, although the latter cannot be ruled out. Microbubbles may enhance all of these processes. Sonoporation can be useful in delivering drugs, eg. to encourage porphyrin uptake in cancerous cells for photodynamic therapy, and microbubbles may even contain the drug to be delivered. Delivering drugs to the brain, in particular, is complicated by the existence of the 'blood-brain barrier', a layer of endothelial cells between the blood and central vervous systems that restricts the passage of many things, such as bacteria, large molecules, etc, from the bloodstream to the brain. Sonoporation can render this barrier temporarily permeable allowing drug treatments of the brain without an invasive procedure.

• **Sonothrombolysis** Thrombosis refers to the formation of a blood clot (thrombus) inside a blood vessel, obstructing the flow of blood putting the patient at risk of stroke and heart attack. Thrombolysis refers to the removal (dissolution) of the clot. A standard treatment for thrombosis is intravenous infusion of 'tissue plasminogen activator' (tPA) into the vessels. It has been shown that ultrasound can enhance the dissolution of the clot when used in conjunction with tPA and/or microbubbles. The mechanism of action is not fully understood, but it is thought that ultrasonic streaming might facilitate permeation of the clot by the drug. Inertial cavitation may also open up pores in the clot further enhancing drug uptake.

• **Cataract surgery** A cataract is the clouding of the crystalline lens of eye, and is usually treated by removing the clouded lens and replacing it with a synthetic lens. One way in which the lens can be removed is by *phacoemulsification*. In this procedure, a small incision is made in the front of the eye, and a focussed ultrasonic probe is used to break up the lens into small enough pieces which can be removed easily. The mechanism is presumably just the mechanical breaking apart of the crystalline lens, rather than a thermal effect. The frequency used is typically around 40 kHz at an intensity of about 1 W/cm^2.

7.6 High power therapeutic applications

• **Extra-corporeal shockwave lithotripsy (ESWL)** Surgery to remove kidney or gall stones that have become too large to pass naturally through the urethra can be dangerous. One alternative is extra-corporeal shockwave lithotripsy (ESWL), in which a large amplitude pulse is focussed onto the region where the kidney stones are in an attempt to break it into smaller pieces which can be passed naturally. Normally, a bipolar (positive and negative) pulse a few microseconds long with an amplitude of tens of MPa is focussed on the kidney stones using a parabolic reflector. Ideally, no damage is done to the healthy tissue outside the region of the focus of the lithotripter. ('Aiming' is usually done with the help of X-ray images.) Typically, a treatment may consist of several hundred pulses over half an hour, and sometimes repeat treatments are necessary. The mechanism by which lithotripsy breaks up kidney stones is not fully understood, but is almost certainly a non-thermal mechanism involving cavitation and jetting.

• **High intensity focussed ultrasound (HIFU)** HIFU has been designed primarily as a treatment for cancers, although non-invasive tissue ablation could be used for numerous applications. In the same way that in lithotripsy large amplitude pulses are focussed onto the treatment region, in HIFU a high-intensity ultrasound beam is focussed onto a chosen region with the intention of killing the tissue. The mechanism is thermal (heating due to absorption), and the aim is to deliver sufficient energy to raise the tissue to a temperature at which it will necrose and hold it there long enough for the tissue to be non-viable. In practice, tissue raised to 55°C for 1 s or longer will undergo coagulative necrosis and immediate cell death. The heating due to absorption can be enhanced by inertial cavitation when the intensity used is very high.

As HIFU transducers typically produce a cigar-shaped lesion about a mm diameter and a few mms long, one lesion is typically insufficient to ablate the required tissue region. In these cases, several lesion are positioned side-by-side to destroy a block of the diseased tissue. MRI or ultrasound imaging are used to guide HIFU delivery. MRI has the advantage that the region of temperature can be monitored, ensuring the focus is at the right place.

HIFU has been used successfully to treat kidney, liver, breast and bone cancers; it is used to treat prostate cancer by delivering the ultrasound with a transrectal probe, and has even been proposed to treat arterial fibrillation, essential tremor, and pain. HIFU can also be used for *haemostasis* (stopping bleeding) and *vascular occlusion* and so has been proposed as a treatment for battle-field wounds to reduce the morbidity due to loss of blood.

Bibliography

This bibliography contains a list of some of the sources used when putting these notes together, but there are numerous other books about waves, acoustics, ultrasound and the related mathematics in the library, web, or bookshops. For reminders about some of the mathematical basics, such as differentiation or complex numbers, for instance, there are resources available on the web, such as `mathworld.wolfram.com`, the MIT Opencourseware site `ocw.mit.edu` or even Wikipedia (if double-checked against a more reliable source).

Obvious though this might sound, one way to understanding something better is finding the source (webpage, book, person, lecture notes, ...) that explains it on the level you are ready for. This will often mean looking through numerous books or websites before finding the one that is written for someone with exactly your background at this moment.

Books

T.L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*, Elsevier, 2004

C.R. Hill, J.C. Bamber, G.R. ter Haar (Editors), *Physical Principles of Medical Ultrasonics*, 2nd Edition, Wiley, 2004

C.L. Morfey, *Dictionary of Acoustics*, Academic Press, 2000

A.D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*, Acoustical Society of America, 1989

L.E. Kinsler, A.R. Frey, A.B. Coppens and J.V. Sanders, *Fundamentals of Acoustics*, 4th Edition, Wiley, 2000

J. Lighthill, *Waves in Fluids*, Cambridge University Press, 1978

R. Beyer, *Nonlinear Acoustics*, Acoustical Society of America, 1997

Journal articles

Progress in Biophysics and Molecular Biology, **93**, 2007. Many of the articles in this issue review particular areas of medical ultrasound.

Special Issue on HIFU, *International Journal of Hyperthermia*, **23**, 2007

D. Dalecki, "Mechanical Bioeffects of Ultrasound", *Annual Review of Biomedical Engineering*, **6**, 229-248, 2004

A. Shaw and G. ter Haar, "Requirements for Measurement Standards in High Intensity Focussed Ultrasound (HIFU) Fields", *NPL Report*, DQL AC 015, 2006

Websites

`www.acoustics.salford.ac.uk/feschools/index.htm` Good reminder of the basic principles of waves and sound, with several good animations.

J. Woo, "A short History of the development of Ultrasound in Obstetrics and Gynecology" `www.ob-ultrasound.net/site_index.html` A useful history of many aspects of clinical ultrasound, despite the narrow focus suggested in the title.

A Some Fundamentals and Mathematical Reminders

These appendices are included to provide a description of some of the basic assumptions which are sometimes left unspoken, and the background knowledge that is required to understand the main text of the notes, such as some mathematical notation.

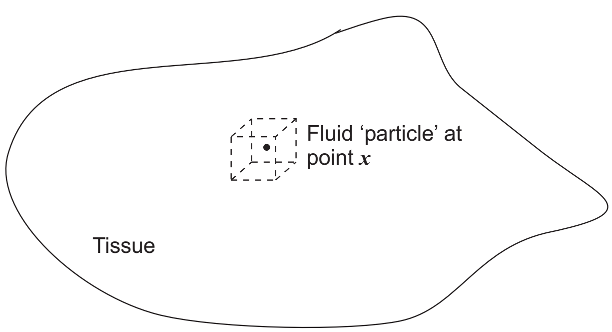
A.1 Continuum hypothesis

If we zoom in gradually on a piece of soft biological tissue, such as skin, we will first see things like hair follicles and cells (10-100 μm), then, on a smaller scale, different types of large molecular complexes and other structures inside and outside the cells (10-100 nanometres). If it were possible to zoom in further we would reach the nanometer scale and we would see individual small molecules. A fundamental assumption usually made when modelling ultrasound is that the description of the tissue at this level is too detailed, ie. the molecular picture of tissue is not relevant to describing ultrasound propagation (although the molecular constitution of the tissue is relevant to ultrasonic absorption, Section 5).

When modelling ultrasonic waves travelling through tissue, water, or some other medium, that medium is usually assumed to be continuous, ie. however small a piece of it you have, you can always divide it in half again - you never get down to a single molecule as you would in real life. This assumption, almost universal in fluid dynamics and acoustics, is called the continuum hypothesis.

> ***Continuum hypothesis****: the medium through which the ultrasound wave is travelling (tissue, water, etc) is continuous.*

Why is the continuum hypothesis widely accepted when it is not really true? For things on a scale much larger than the average spacing between the molecules, treating the medium as continuous is a good approximation. The spacings between the molecules in tissue is of the order of nanometres, so as long as we are only interested in distances longer than a few hundred nanometres, we can safely think of it as though it were continuous. The wavelength of the ultrasound waves we will be interested in is of the order of $\lambda \approx 0.1 - 10$ millimetres - long enough for the continuum hypothesis to be a very good approximation.



Fluid elements and particles An *element* or*particle* of a continuous fluid (or solid) is an imaginary, abstract, small amount of the fluid, that moves with the fluid. It can be a useful theoretical device when formulating equations and when thinking about some quantitites, such as stress and fluid acceleration. This sort of imaginary ‘particle’ is a purely theoretical construct and should not be confused with molecules or cells, which are real, finite-sized, constituents of the tissue, but are ignored for our purposes here. (It should also not be confused with a small fixed region of *space* through which the fluid flows, which is used when deriving the law of mass conservation, for example.)

A.2 Thermodynamics and stress

One advantage of treating tissue as a continuum is that we can talk of the mass density, temperature, etc. *at a point* in the tissue. As well as density and temperature, the pressure and specific entropy are fundamental quantities of interest here. These are all widely used properties of materials, but at the risk of repetition, and to help develop an intuitive grasp of what they are and how they behave, consider the following simple analogies.

• **Gas** Consider a box containing gas molecules which behave like miniscule bouncy balls. In this analogy, the *density* is defined as the total mass of balls in the box divided by the box volume, the *temperature* of the gas is related to how fast the balls are flying around - the faster they fly the hotter the gas, and the *pressure* on a wall of the box is proportional to how often and how fast the balls hit it. The *specific entropy* (entropy per unit mass) is a measure of the likelihood that the box of balls, taken as a whole, is in the thermodynamic state it is in given that the individual balls can move around randomly. eg. if all the balls, moving randomly, happen to end up on one side of the box, the probability of which is low, the entropy will be low. It is sometimes said that entropy is a measure of 'disorder'. These are the ideas underlying the *Kinetic Theory of Gases*.

• **Elastic Solid** Soft tissue cannot realistically be treated as a gas, so consider a continuous solid and imagine it divided into many cubic 'elements', something like the solids in Fig. 1. The *density* of an element is the mass of that element divided by its volume, and if the element is squashed, so the element volume decreases, the density will go up. Each element will be constantly vibrating with a small amplitude due to thermal motion, and the *temperature* at a point is proportional to the amplitude of the vibration there, ie. how much the element there is 'jiggling' about its mean position. The *entropy* is, as above, a measure of the 'disorder' of the system. When heat flows it spreads the random vibrations of molecules thereby increasing the disorder and therefore the entropy. The linear acoustics, the assumption of constant entropy is equivalent to assuming there is zero heat flow. The pressure - or rather stress, for a solid - is described below.

• **Normal and Shear Stress** Now imagine the elastic solid is distorted, perhaps by twisting or squashing, then cut in half, and one half is taken away. The remaining half will return to its original shape, because it is elastic, unless forces are applied to each point of the freshly-cut surface to replace the forces that the removed part of the mesh was providing. Those forces (per unit area) are called the *stresses* at those points. Before the distorted solid was cut, those stresses existed within the material. For a small cubic element of the elastic solid, Fig. 30, there are two types of stress acting on each face: normal stress (perpendicular to the surface) and shear stress (parallel to the surface, two components). As there are three directions in space, (x, y, z), nine numbers are needed in general to define the stress at any point:

$$\sigma_{ij} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix}. \tag{123}$$

The first subscript refers to the surface (it gives the axis to which the surface is perpendicular) and the second to the direction of the force. (Some authors use the opposite convention: there is no general agreement.) So the stresses on the matrix diagonal, σ_{xx}, σ_{yy} and σ_{zz}, are normal stresses, and σ_{xy}, for example, is a shear stress in the direction y on a surface perpendicular to the x-axis.

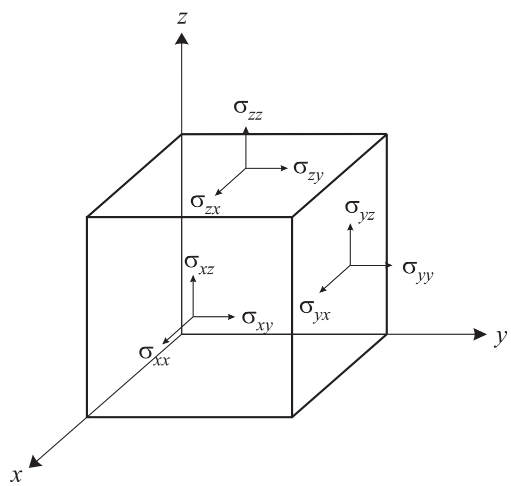


Figure 30: The nine components of a stress tensor, σ_{ij}.

• **Stress and Pressure** A non-viscous fluid cannot support shear stresses, so in this case the off-diagonal terms in the stress tensor σ_{ij} are zero. The *pressure*, p, is defined as the mean of the normal stress components

$$p = -\frac{1}{3}\left(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}\right). \tag{124}$$

The minus sign arises because, by convention, a positive pressure is compressive and a compressive stress is negative (a tensile stress is positive).

• **Shear and Compressional Waves** When a shear wave travels through tissue, it is the shear stress that fluctuates, but when a compressional wave propagates, it is the normal stress, ie. the pressure. In a fluid, only compressional waves can propagate - these are sound waves or acoustic waves.

• **Velocity potential** [m^2/s] In acoustics, it is usually assumed that the fluid is irrotational (no fluid elements are rotating) which allows the acoustic velocity vector $\mathbf{u}_a$ to be written in terms of a velocity potential, ϕ:

$$\mathbf{u}_a = \nabla\phi = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right). \tag{125}$$

The velocity potential is related to the acoustic pressure (in linear acoustics) by $p_a = -\rho_0(\partial\phi/\partial t)$. Some authors use velocity potential to describe acoustic fields. However, the **acoustic pressure** will be the variable usually used in these notes, as it is the one that can be measured most readily.

A.3 Vectors and vector calculus

- A **scalar** is a quantity that can be represented by a single number, such as mass.
- A **vector** is a quantity that has an associated direction, such as space shuttle's velocity, or the position of a point, $\mathbf{x}$, with respect to the origin, and so requires three numbers to represent it in three-dimensional space, $\mathbf{x} = (x, y, z)$, two to represent it in two-dimensional space, etc. Vectors will usually be written in **bold** in these notes, although some authors use other notations, eg. $\vec{u}$.
- A **tensor** is a further generalisation of this idea. A quantity that requires 9 numbers to describe it is a tensor of order 2. (In a sense, a vector is a tensor of order 1 and a scalar is a tensor of order 0.) All second order tensors can be written as 3×3 matrices, but not all 3×3 matrices are tensors. Higher order tensors are possible, but we are rarely concerned with anything more than vectors in this course.
- A **scalar/vector/tensor field** is a function, typically of position $\mathbf{x}$, that takes a scalar/vector/tensor value.
  - Scalar fields: $T(\mathbf{x})$, $p(\mathbf{x})$ which are scalar functions of the position $\mathbf{x}$, ...
  - Vector fields: particle velocity $\mathbf{u}(\mathbf{x}) = (u_x(\mathbf{x}), u_y(\mathbf{x}), u_z(\mathbf{x}))$, ...
  - Tensor fields: stress, $\sigma_{ij}(\mathbf{x})$, strain $\epsilon_{ij}(\mathbf{x})$, ...

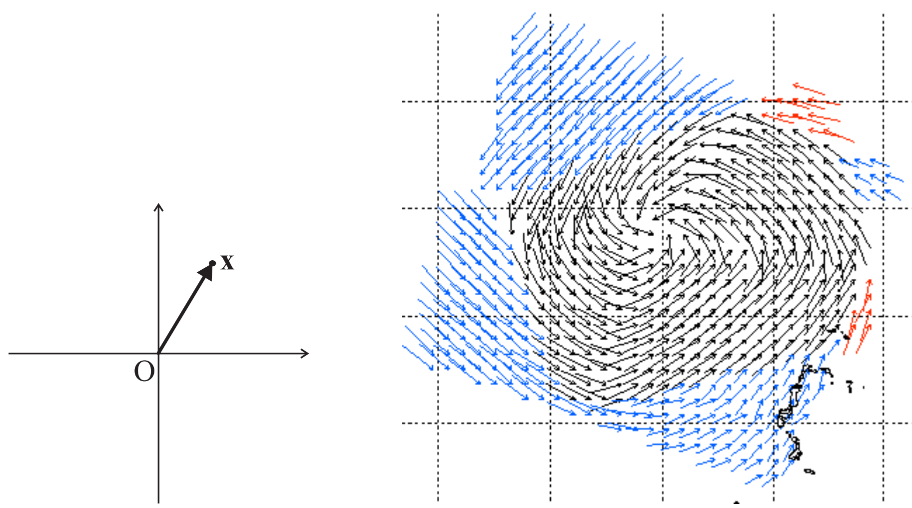
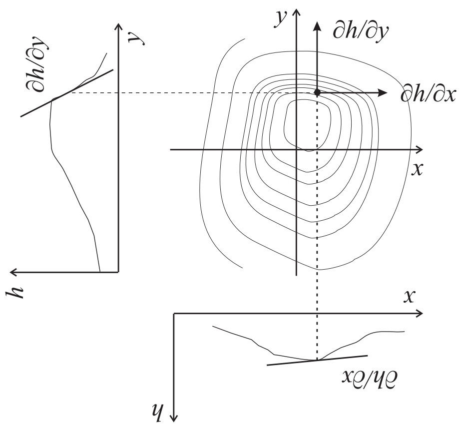


Figure 31: (left) A vector, such as position relative to the origin O, has magnitude and direction. (right) A vector field: the wind velocity in a hurricane.

• The **gradient operator** ∇, sometimes written **grad** and also called *del* or *nabla*, is defined in three dimensions as

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right). \tag{126}$$

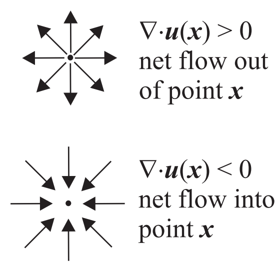
The one-dimensional gradient operator is just $\partial/\partial x$.



• **Gradient of a scalar is a vector**

$$\nabla \phi \equiv \left(\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}, \frac{\partial \phi}{\partial z} \right) = \mathbf{u} \tag{127}$$

Think of a hill, pictured with contour lines below. The height at a point $\mathbf{x} = (x, y)$ on the hill can be written as the scalar field $h(x, y)$. At $\mathbf{x}$ there are two components to the gradient vector, $\nabla h = (\partial h/\partial x, \partial h/\partial y)$.



• **Divergence of a vector is a scalar**

$$\nabla \cdot \mathbf{u} \equiv \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z} \right) \tag{128}$$

Positive divergence indicates dilatation or a *source*, in the following sense. Imagine a submerged water hose pumping water into a tank. If $\mathbf{u}(\mathbf{x})$ is the velocity vector of the flow then the divergence over a small volume surrounding the end of the hose $\int (\nabla \cdot \mathbf{u}) dV$ would be positive, whereas around a plughole it would be negative.

A.4 Fluid element acceleration - material derivative

The **material derivative** is a measure of the *time* rate of change of some quantity when movement in *space* is also involved. It is also called the 'Lagrangian derivative' and 'derivative following the motion' among other things. It is written as the operator D/Dt and defined as

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \tag{129}$$

where $\mathbf{v}$ is a velocity vector.

Example 1 Imagine a large room with a light bulb on at one end, and some light-absorbing curtains, so the amount of light varies throughout the room. Now imagine someone holding a light meter which gives some measure of the amount of light, L, and we are interested in the rate of change of L with time. There are two ways in which the reading on the light meter, L, might change:

1. If the light meter were held still, L would only change if the amount of light being emitted by the bulb changed. This is the usual partial derivative with respect to time:

$$\text{time rate of change of } L \text{ at a fixed point } \mathbf{x} = \frac{\partial L(\mathbf{x})}{\partial t}. \tag{130}$$

2. Now assume the light from the bulb is constant. If the light meter were moved through small distance at a velocity $\mathbf{v}$, the reading L would change because the amount of light in different parts of the room varies with position - the meter has been moved to a brighter or dimmer spot. The change in L with time would be the rate of change of position with *time* $d\mathbf{x}/dt$ multiplied by the rate of change of L with *position* $\partial L/\partial \mathbf{x}$.

$$\text{time rate of change of } L \text{ due to changing position } = \frac{d\mathbf{x}}{dt}\frac{\partial L}{\partial \mathbf{x}} = \mathbf{v} \cdot \nabla L \tag{131}$$

- In general, the total rate of change of L is the sum of these two contributions

$$\text{total time rate of change of } L \ = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right) L \tag{132}$$

Example 2 Imagine an oil-filled tank. The owner of the oil wants to measure its temperature, so puts a tiny controllable temperature sensor into the tank which records T_{sensor}.

- First, when the sensor is stationary, any changes in T_{sensor} must be due to the fact that the temperature of the oil $T(\mathbf{x})$ is changing, eg. perhaps it is being heated by the sun,

$$\text{time rate of change of } T_{\text{sensor}} \text{ for a stationary sensor} = \frac{\partial T(\mathbf{x})}{\partial t} \tag{133}$$

- Second, assume that the temperature field $T(\mathbf{x})$ is in a steady state, so is not changing with time but is different at different points. If the sensor is driven around in the oil, measured changes in T_{sensor} must be due to the fact that it moves into hotter or colder regions of the fluid. The time rate of change of T_{sensor} must depend on how fast the sensor is moving, $\mathbf{v}$, and on how large the temperature gradients are:

$$\text{time rate of change of } T_{\text{sensor}} \text{ due to sensor movement } = \mathbf{v} \cdot \nabla T \tag{134}$$

- In general, both contributions must be taken into account, so the time rate of change of T_{sensor} measured by a sensor moving with velocity $\mathbf{v}$ in a varying temperature field $T(\mathbf{x}, t)$ is

$$\text{total time rate of change of } T_{\text{sensor}} = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) T. \tag{135}$$

Example 3 Suppose the tank in the last example is replaced by a pipe, so the oil can flow, and the temperature sensor is replaced by a velocity sensor (perhaps it can use GPS to work out its velocity vector $\mathbf{v}$). Suppose also that the sensor moves passively with the fluid - it has no motor to power itself along.

- First, assume that the velocity vector field $\mathbf{u}(\mathbf{x})$ does not change with time, it is in a steady state. Although the velocity of the flow is different at different points, eg. where there is a constriction in the pipe the oil will flow faster, the pattern of velocites remains constant with time. In this case, the rate of change of velocity measured by the sensor, $\mathbf{v}$, would be due purely to its changing position:

$$\text{time rate of change of } \mathbf{v} \text{ due to sensor movement } = (\mathbf{v} \cdot \nabla)\, \mathbf{u} = (\mathbf{u} \cdot \nabla)\, \mathbf{u} \tag{136}$$

  Note that, as the sensor is moving with the fluid, the velocity of the sensor $\mathbf{v}$ is just the value of the local velocity vector field $\mathbf{u}$.

- Now suppose that the velocity vector field varies with time $\mathbf{u}(\mathbf{x}, t)$. (Perhaps it is turbulent.) Changes in the velocity vector field $\mathbf{u}(\mathbf{x})$ will directly affect the velocity of the sensor $\mathbf{v}$, so the additional time-dependent term must be added:

$$\text{time rate of change of sensor velocity } \mathbf{v} \ = \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathbf{u} = \frac{D\mathbf{u}}{Dt} \tag{137}$$

- A time rate of change of velocity is an acceleration, so this expression gives the acceleration of the fluid particles that have been flowing along with the sensor.

Another way to think of the **fluid acceleration** is that, as the velocity vector depends on both time t and the position of the fluid particle, $\mathbf{x}$, and as the latter also depends on time, the total derivative with respect to time must include both partial derivatives. Consider one component of the velocity u_x:

$$\text{total rate of change of } u_x \text{ with } t \approx \frac{\partial u_x}{\partial t} + \frac{\partial \mathbf{x}}{\partial t} \cdot \nabla u_x = \frac{\partial u_x}{\partial t} + \mathbf{u} \cdot \nabla u_x \tag{138}$$

Extending this to the other three dimensions gives the acceleration vector of a fluid particle moving with the fluid

$$\frac{D\mathbf{u}}{Dt} = \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathbf{u}. \tag{139}$$

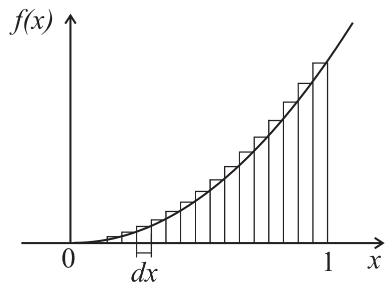
A.5 Integrals

• The integral of a function is the area under the curve. If the function goes negative, such as $\sin(x)$, then the integral might be zero or even negative.

• Integrals are just summations (the integral sign is a distorted 'S' for 'sum'). eg. the area under the curve $f(x)$ between 0 and 1 is

$$A = \int_0^1 f(x)dx \approx \sum_{n=1}^{N} f(n/N)dx \tag{140}$$

where $dx = 1/N$ and the approximation is more accurate the larger N is.

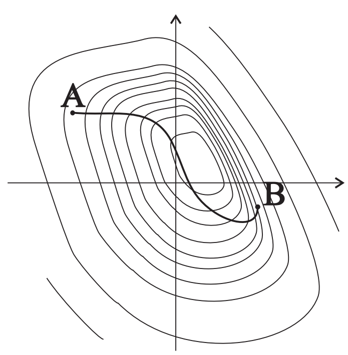


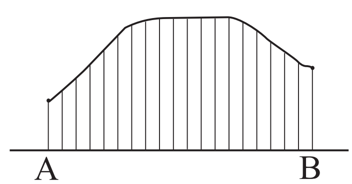
• **Line integrals** A line integral, or path integral

$$\int_l f(\mathbf{x})dl \tag{141}$$

is the summation of the values of a function $f(\mathbf{x})$ that lie on the line or path $\mathbf{x} = l$. eg. l may be the line $x = 1$, or the circle $x^2 + y^2 = 1$, or an arbitrary wiggly line $A \rightarrow B$. In one dimension, there is only one line to follow, so the 1-D line integral $\int f(x)dx$ is just the normal integral described above. A two-dimensional example is shown in the figure. A three dimensional example would be a thin 'pencil-beam' of X-rays passing through the body. The X-rays are absorbed as they go through the body, so the total absorbed is the integral of the body's X-ray absorption coefficient along the path of the beam,

$$\int_{\text{x-ray path}} \alpha(x)dl \tag{142}$$





• **Surface integrals** A surface integral extends the concept of a line integral to an integration over a two-dimensional surface.

$$\iint_S f(\mathbf{x})dS \tag{143}$$

Example: imagine a fan blowing against a wall. What is the *total* pressure on the wall caused by the fan? It is the integral of the pressure at each point on the wall over the surface of the wall. If the local pressure at the wall is denoted $p(x, y)$, then

$$\text{total pressure} = \iint_{\text{wall}} p(x, y)dxdy \tag{144}$$

Another example: in order to calculate the total amount of light leaving a light bulb at a moment in time, imagine a spherical shell enclosing the bulb, and sum the values of the light intensity, $I(\mathbf{x})$ at every point on that spherical shell.

$$\iint_{\text{imaginary sphere}} I(\mathbf{x})dS. \tag{145}$$

The surface of integration could be any arbitrary shape.

• **Volume integrals** A volume integral is an integral over three dimensions

$$\iiint_V f(\mathbf{x})dV \tag{146}$$

For instance, if the density of the earth is given by $\rho_{\text{earth}}(\mathbf{x})$, then the mass of the earth would be

$$\text{mass of the earth } = \iiint_V \rho_{\text{earth}}(\mathbf{x})dV \tag{147}$$

where V is the volume of the earth. If the density of the earth were constant and it were spherical with radius r, then this would simplify to

$$\text{mass of the earth } = \rho_{\text{earth}}(\text{volume of sphere}) = \rho_{\text{earth}}\left(\frac{4}{3}\pi r^3\right) \tag{148}$$

A.6 Complex numbers and trigonometric functions

- A complex number consists of two scalars x and y in the combination $z = x+iy$, where i is defined such that $i^2 = -1$.
- The real and imaginary parts of z are written $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$, which in this case are x and y respectively.
- z can also be written in modulus and argument form, $z = re^{i\theta}$, where $r = |z| = \sqrt{x^2 + y^2}$ is the modulus of z and $\theta = \tan^{-1}(y/x)$.

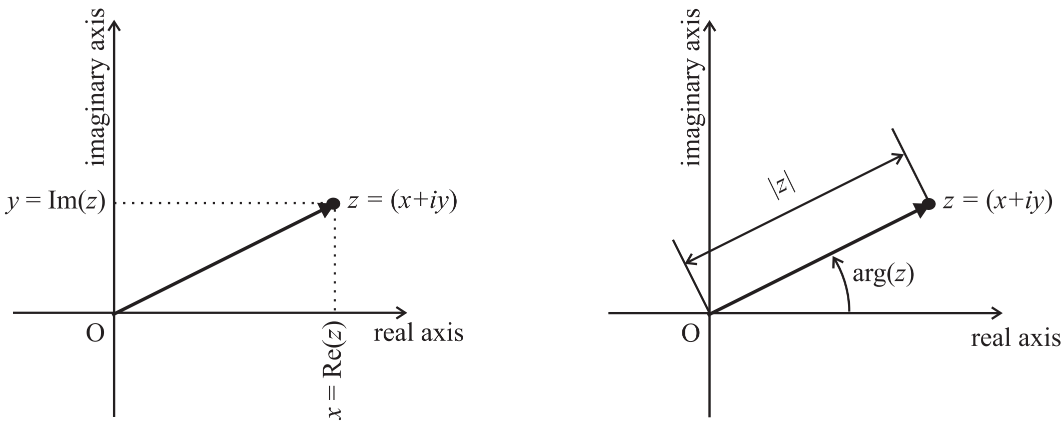


Figure 32: Two ways to represent a complex number z on the complex plane: in real and imaginary parts, $z = x + iy$, or in modulus and argument form, $z = re^{i\theta}$.

- Euler's formula relates the sine, cosine and exponential functions:

$$e^{i\theta} = \cos(\theta) + i\sin(\theta) \tag{149}$$

Putting $\theta = \pi$ gives

$$e^{i\pi} + 1 = 0 \tag{150}$$

one of the most beautiful equations in maths, combining five fundamental numbers in one tidy and surprising formula. Euler's formula can be proved by looking at the Taylor series expansions of the functions $\sin(\theta)$, $\cos(\theta)$ and $\exp(i\theta)$.

A.7 Fourier analysis

• **Periodic signals: Fourier series** A pure tone is a sinusoidal acoustic pressure variation, and so has a well-defined frequency ω in rad/s (or f in Hz) and can be written as $\sin(\omega t)$ (or equivalently $\sin(2\pi f t)$), as shown in Fig. 33A. It is a periodic signal, in that it repeats every $2\pi/\omega$ seconds.

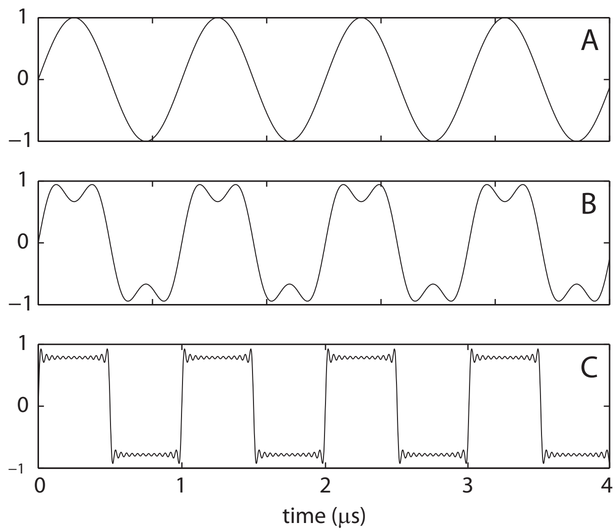


Figure 33: (A) A sine wave, (B) A periodic wave formed by adding two sine waves together, (C) A periodic wave formed by adding 25 sine waves together.

If two sine waves of different frequency are added together, then another periodic waveform results, eg. $y(t) = \sin(\omega_1 t) + \frac{1}{3}\sin(\omega_2 t)$, plotted in Fig. 33B for $\omega_1 = 1$ MHz, $\omega_2 = 3$ MHz.

Can we construct any waveform in this way? A sine wave is an *odd function*, which means it is not symmetrical about zero, but rather $\sin(-x) = -\sin(x)$. Any *odd and periodic* waveform can be constructed by adding togther sine waves. eg. Fig. 33C shows the first 25 terms of the following sum which approximates a square wave:

$$y_{\text{square}}(t) = \sum_{n=1,3,\ldots}^{\infty} \frac{1}{n}\sin(n\omega_0 t) \tag{151}$$

Similarly, we could construct any *even function* by adding together cosines. As any function can be separated uniquely into odd and even parts,[2] we can generalise the ideas above to any *periodic* function (odd, even, or neither) by adding together both sines and cosines. We can write that

$$\boxed{y_{\text{periodic}}(t) = a_0 + \sum_{n=1}^{\infty}\left(a_n\cos(n\omega_0 t) + b_n\sin(n\omega_0 t)\right)} \tag{152}$$

where a_0 is a constant (sometimes called the d.c. component), and ω_0 is called the *fundamental frequency*. Note that all the components have frequencies that are multiples of this fundamental frequency. This representation of a signal is called a **Fourier series** and can be found for any *periodic* wave.

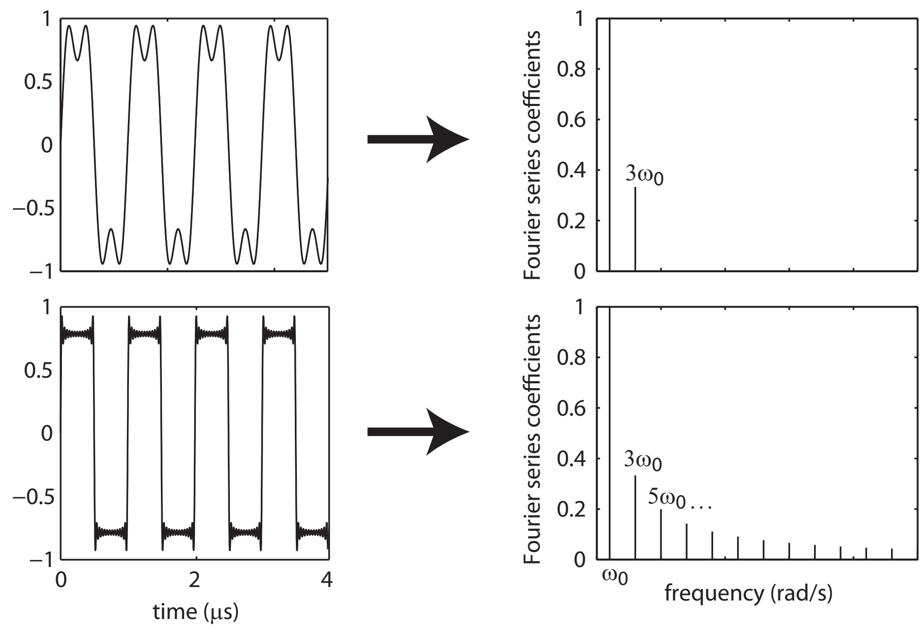
[2]For an arbitrary function $f(x)$, the odd part is $f_{\text{odd}} = (f(x) - f(-x))/2$ and the even part $f_{\text{even}} = (f(x) + f(-x))/2$. Clearly $f_{\text{odd}} + f_{\text{even}} = f(x)$.

Fourier series coefficients Given a periodic signal $y_{\text{periodic}}(t)$, how do we find the coefficients a_n, b_n? They can be found using the following integrals over one period T of the signal:

$$a_0 = \frac{1}{T}\int_0^T f(t)dt, \qquad a_n = \frac{2}{T}\int_0^T f(t)\cos(n\omega_0 t)dt, \qquad b_n = \frac{2}{T}\int_0^T f(t)\sin(n\omega_0 t)dt \quad (153)$$

These integrals are asking, "for a particular frequency $n\omega_0$, how much $\cos(n\omega_0 t)$ is there in the signal $f(t)$?" eg. if $f(t)$ was such that when multiplied by $\cos(m\omega_0 t)$ the resulting function, $f(t)\cos(m\omega_0 t)$, was zero everywhere, then that coefficient a_m would be zero, ie. this particular cosine is not useful when trying to represent $f(t)$ in terms of sin and cos.

Examples Starting with an almost trivial example. If we have the function $f(t) = 2\cos(\omega_0 t)$ then all the coefficient integrals will work out to be zero except for $a_1 = 2$. The only useful term in this instance being $a_1\cos(\omega_0 t)$, and we need $a_1 = 2$ to get the function $f(t)$. Another example: if we have the periodic function $y(t)$ shown in Fig. 33B, and we want to find the Fourier series representation of it, we will find that most of the integrals are zero and only two terms are needed to describe it, $b_1\sin(\omega_0 t)$ and $b_3\sin(3\omega_0 t)$ with $b_1 = 1$ and $b_3 = \frac{1}{3}$. This combination gives $y(t) = \sin(\omega_0 t) + \frac{1}{3}\sin(3\omega_0 t)$, as expected. This example, along with the square wave example above, are shown in Fig. A.7 in both the time and frequency (Fourier) domains. Representing a function as Fourier series can be thought of as *mapping* it into the frequency domain.



• **Non-periodic signals: Fourier transform** The Fourier *series* representation described above is ok for periodic functions, but what about most signals which aren't periodic? Recall that in the Fourier series representation, all the component frequencies were multiples of the fundamental frequency, $n\omega_0$. The fundamental frequency is related to the period of the wave by $T = 2\pi/\omega$. If a signal is not periodic, it never repeats, so it could be argued that its period tends to infinity. As T gets longer and longer, ω_0, which is the spacing between the Fourier components, gets smaller and smaller. Eventually, the frequency components, $n\omega_0$, are so close together, the sum in Eq. 152 becomes an integral, and we can write, for any *non-periodic* signal $f(t)$

$$f(t) = \int_{-\infty}^{\infty} (a'(\omega)\cos(\omega t) + b'(\omega)\sin(\omega t))\, d\omega \tag{154}$$

Euler's formula (see appendix) tells us that $\exp(i\omega t) = \cos(\omega t) + i\sin(\omega t)$, so this can be rewritten in terms of complex exponentials as

$$\boxed{f(t) = \int_{-\infty}^{\infty} F(\omega)\exp(i\omega t) d\omega}. \tag{155}$$

The complex function $F(\omega)$ is the **Fourier *transform*** of the function $f(t)$. $F(\omega)$ 'lives' in the frequency, or Fourier domain, and $f(t)$ in the time domain. Interestingly, and very usefully, the **inverse Fourier transform** is simply

$$\boxed{F(\omega) = \int_{-\infty}^{\infty} f(t)\exp(-i\omega t) dt} \tag{156}$$

Together $f(t)$ and $F(\omega)$ are known as a **Fourier transform pair**.

$F(\omega)$ is complex, so can either be expressed in real and imaginary parts or in term of magnitude and phase:

$$F(\omega) = F_r(\omega) + iF_i(\omega) = |F(\omega)|\exp(i\theta(\omega)) \tag{157}$$

It is more common to express frequency domain functions in terms of magnitude $|F(\omega)|$ and phase $\theta(\omega)$. In fact, it is often only the magnitude of the Fourier transform that is of interest. Fig. 34 shows functions $f(t)$ and the corresponding $|F(\omega)|$.

Fourier transforms could easily form a whole course by themselves, and there is no space here to describe all the properties of Fourier transforms. However, a few properties are so important they are worth mentioning in passing. Here, the Fourier transform will be denoted as $F(\omega) = \mathcal{F}(f(t))$.

- Fourier transforms are linear: $\mathcal{F}(g(t) + f(t)) = \mathcal{F}(g(t)) + \mathcal{F}(f(t)) = G(\omega) + F(\omega)$
- Convolutions in the time domain become multiplications in the frequency domain: $\mathcal{F}(h(t) \otimes f(t)) = \mathcal{F}(h(t))\mathcal{F}(f(t)) = H(\omega)F(\omega)$
- If $f(t)$ is a short pulse, then $F(\omega)$ will be broad in frequency - short pulses contain lots of frequencies.
- Conversely, if $f(t)$ is long, eg. an infinite sinusoid, then $F(\omega)$ is narrow. (See Fig. 34.)

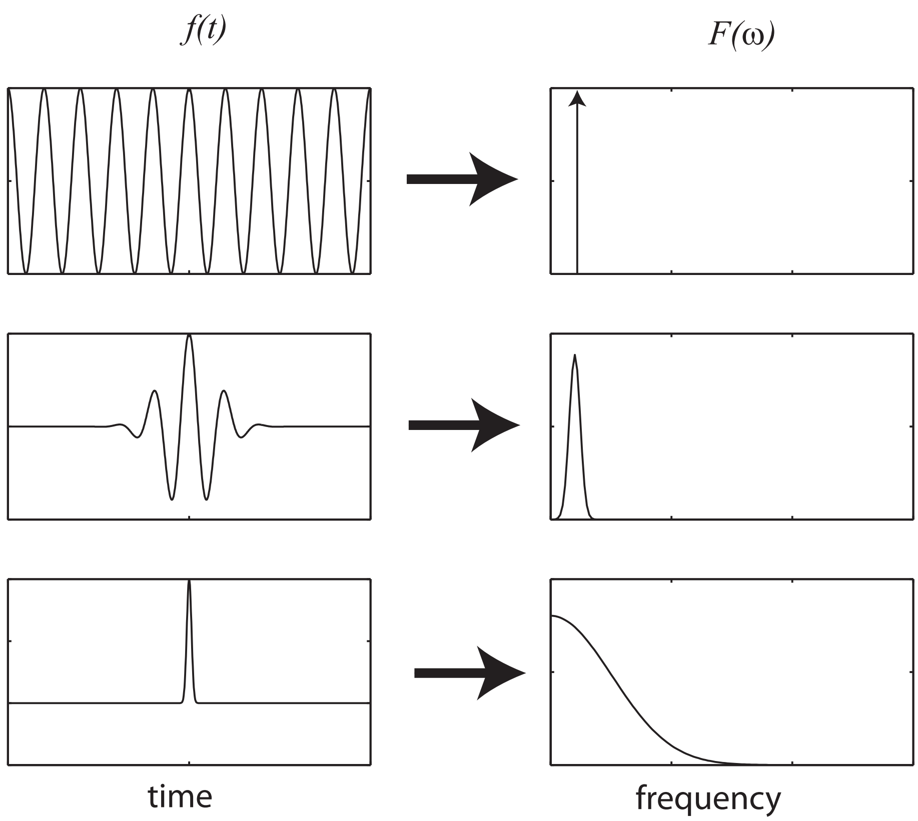


Figure 34: Long signals in time have narrow Fourier transforms. The Fourier transform of an infinite sinusoid is a delta function. Short pulses in time have broad Fourier transforms - they are 'broadband' in that they contain a wide range of frequencies.