Chapter 2

Non-dispersive Waves

We will begin this series of lectures on waves by spending some time discussing non-dispersive waves. Sound waves are the most common example of non-dispersive waves in fluids. Other examples in other contexts include electromagnetic waves, waves on a uniform string, etc .... As we shall see, the defining property of non-dispersive waves is that their *dispersion relation* (i.e. the relationship between a wave's spatial wavenumber k and temporal frequency ω) is linear: $\omega = \alpha k$. As a result of this linear relation, non-dispersive waves have a number of interesting properties, which will be discussed here. Although we will limit our analysis to the case of sound waves, the general properties we shall derive are applicable to all non-dispersive waves.

In the first part of this lecture, we will look at small-amplitude sound waves in a homogeneous and time-independent background medium. Later on, we will relax some of these assumptions and see what other interesting dynamics arise.

2.1 Sound waves in a homogeneous, invariant medium

When sound waves are small in amplitude, we can treat them *perturbatively*, that is, we can view them as small perturbations on a given background state. As given in the title of this Section, we will first consider a background state that is homogeneous (i.e. its properties are independent of position) and time-invariant (i.e. its properties are independent of time). Let us therefore assume that we have a fluid, whose density, pressure and temperature are constant, and equal to ρ_0, p_0 and T_0 respectively. This could be, for instance, the conditions inside a well-insulated room. We shall assume that this fluid has, for instance, an equation of state relating ρ_0, p_0 and T_0. We shall also assume that the background fluid is perfectly still, that is, without any motion at all.

2.1.1 The wave equation

We now want to study how pressure and velocity perturbations propagate in this environment. The equations describing the motion of the fluid, assuming that it is inviscid (which is a good approximation for normal-frequency sound waves in air) are

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}) = 0$$
$$\rho \left(\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} \right) = -\nabla p \tag{2.1}$$

where we have neglected the effect of gravity on the grounds that it is very small – this can be verified a posteriori. Even with this approximation, solving these equations is rather difficult because of their nonlinear nature. However, if we assume that perturbations to the motionless, homogeneous background state are small, then we can linearize the governing equations around that background state, and study the behavior of the linearized equations instead. This approach is in essence very similar to that used in "Dynamical Systems" (AMS214), to study the behavior of solutions near fixed points.

Hence, let $p(x, y, z, t) = p_0 + \tilde{p}(x, y, z, t)$, and similarly $\rho(x, y, z, t) = \rho_0 + \tilde{\rho}(x, y, z, t)$. Since the background state is motionless, we simply have $\boldsymbol{u} = \tilde{\boldsymbol{u}}$. Substituting this into the governing equations, we then have

$$\frac{\partial}{\partial t}(\rho_0 + \tilde{\rho}) + \nabla \cdot [(\rho_0 + \tilde{\rho})\tilde{\boldsymbol{u}}] = 0$$
$$(\rho_0 + \tilde{\rho})\frac{\partial \tilde{\boldsymbol{u}}}{\partial t} + (\rho_0 + \tilde{\rho})\tilde{\boldsymbol{u}} \cdot \nabla \tilde{\boldsymbol{u}} = -\nabla(p_0 + \tilde{p}) \tag{2.2}$$

These equations can be simplified by remembering that the steady state is homogeneous (so $\nabla p_0 = 0$ and $\nabla \rho_0 = 0$) and time-invariant (so $\partial \rho_0 / \partial t = 0$), and neglecting any quadratic term in the perturbations. This implies

$$\frac{\partial \tilde{\rho}}{\partial t} + \nabla \cdot (\rho_0 \tilde{\boldsymbol{u}}) = 0$$
$$\rho_0 \frac{\partial \tilde{\boldsymbol{u}}}{\partial t} = -\nabla \tilde{p} \tag{2.3}$$

Taking the time-derivative of the mass conservation equation, and substituting the momentum equation yields

$$\frac{\partial^2 \tilde{\rho}}{\partial t^2} + \nabla \cdot \left(\rho_0 \frac{\partial \tilde{\boldsymbol{u}}}{\partial t} \right) = \frac{\partial^2 \tilde{\rho}}{\partial t^2} - \nabla \cdot (\nabla \tilde{p}) = 0 \tag{2.4}$$

This implies

$$\frac{\partial^2 \tilde{\rho}}{\partial t^2} = \nabla^2 \tilde{p} \tag{2.5}$$

Clearly, we need another equation to relate $\tilde{\rho}$ and $\tilde{p}$.

So far, we have not used the equation of state, so let's do so now, and linearize it as we did for the momentum and mass conservation equations. Suppose, for instance, that we have a perfect gas, for which $p = R\rho T$. Then, linearizing this, we have

$$p_0 + \tilde{p} = R(\rho_0 + \tilde{\rho})(T_0 + \tilde{T}) \tag{2.6}$$

which yields

$$\tilde{p} = R(\rho_0 \tilde{T} + T_0 \tilde{\rho}) \tag{2.7}$$

This equation has the disadvantage of containing $\tilde{T}$, for which need another equation. However, we have not used the thermal energy equation yet, so this is where the final piece of the puzzle will come from.

2.1.2 Isothermal vs. Adiabatic sound waves

Physically speaking, the appearance of $\tilde{T}$ in the problem is not particularly surprising: pressure waves are – by definition – compressional waves, and in many cases a gas heats up when compressed. The change in the temperature then changes the density, through the equation of state. The proper way to proceed with the problem is to add the thermal energy equation and linearize it as well, to close the system. This is somewhat overkill, however, but a nice exercise. In practice, there are two extreme possibilities.

On the one hand, if the waves are very low frequency, and thermal dissipation is very efficient (either by radiation in an optically thin medium, or by very efficient thermal diffusion), then any local heating of a fluid parcel is immediately dissipated away. In that case, despite the compressional heating, $\tilde{T} = 0$ at all times and $\tilde{p}$ and $\tilde{\rho}$ are now related via $\tilde{p} = RT_0\tilde{\rho}$. Plugging this in the wave equation, we get

$$\frac{\partial^2 \tilde{\rho}}{\partial t^2} = RT_0 \nabla^2 \tilde{\rho} = c_T^2 \nabla^2 \tilde{\rho} \tag{2.8}$$

which now really has the form of a wave equation, and reveals the quantity RT_0 to be the square of the wave speed c_T (see AMS 212A, and below). Given the assumption that the perturbations are isothermal, we see that $c_T = \sqrt{RT_0}$ is the *isothermal* wave speed.

On the other hand, we could assume that we are in an optically thick medium, and/or that the waves are high frequency, so the heat generated by compression is locally trapped – the waves are adiabatic. We saw in the previous chapter that, in the adiabatic case, one can use the entropy equation instead of the thermal energy equation, and the former reduces to

$$\frac{Dp}{Dt} = \gamma \frac{p}{\rho} \frac{D\rho}{Dt} \tag{2.9}$$

for adiabatic motion. Linearizing this equation, we get

$$\frac{\partial}{\partial t}(p_0 + \tilde{p}) + \tilde{\boldsymbol{u}} \cdot \nabla(p_0 + \tilde{p}) = \gamma \frac{p_0 + \tilde{p}}{\rho_0 + \tilde{\rho}} \left[\frac{\partial}{\partial t}(\rho_0 + \tilde{\rho}) + \tilde{\boldsymbol{u}} \cdot \nabla(\rho_0 + \tilde{\rho}) \right] \tag{2.10}$$

so, neglecting all quadratic terms in the perturbations, we get

$$\frac{\partial \tilde{p}}{\partial t} = \gamma \frac{p_0}{\rho_0} \frac{\partial \tilde{\rho}}{\partial t} \tag{2.11}$$

Plugging this into the wave equation, we now have

$$\frac{\partial^2 \tilde{p}}{\partial t^2} = \gamma \frac{p_0}{\rho_0} \nabla^2 \tilde{p} = \gamma R T_0 \nabla^2 \tilde{p} = c_s^2 \nabla^2 \tilde{p} \tag{2.12}$$

Again, we recover a standard wave equation but this time the wave speed is slightly different, and is $c_s = \sqrt{\gamma R T_0} = \sqrt{\gamma} c_T$. The quantity c_s is also a sound speed, but because it was derived for adiabatic perturbations, it is called the *adiabatic* sound speed.

How different are the adiabatic and isothermal sound speeds? The answer is, for normal fluids, not very different – the adiabatic index γ is fairly close to 1, and so its square root is even closer to 1. For instance, for air at ambient temperature, $\gamma \simeq 1.4$ so the difference between the isothermal and adiabatic sound speeds is of the order of 18%. Still, this is not negligible, and was already noted by Laplace. The isothermal sound speed in air is about 290m/s, and the adiabatic sound speed is about 340m/s. For normal frequency sounds (pitch of someone's voice, etc..), it is the adiabatic sound speed that is relevant.

Exercise: If the time between a lightning strike and the roll of thunder is 10 seconds, how far was the strike?

2.1.3 Where are sound waves found?

Sound waves are simply pressure perturbations that propagate in a compressible medium. Of course, they are most commonly thought of in the context of air, or more generally in any terrestrial gases. However, sound waves also exist in any astrophysical object, and are likely to be found everywhere in the universe, from the interior of stars and planets (where they are called *p*-modes), in accretion disks, and in the interstellar and intergalactic medium. They have recently gained popularity via helio- and astero-seismology, where they are observed and their characteristics are used to infer the internal properties of the solar interior and more generally of stellar interiors.

Liquids such as water are in fact also compressible (but much less so than gases), so they also support sound waves. The speed of sound in water at ambient temperature is around 1500 m/s, hence about 4-5 times faster than in air.

2.2 Some 1D solutions of the wave equation

In what follows, for simplicity we will ignore the difference between the isothermal and the adiabatic sound speed, and drop the tildes, to write the wave equation as

$$\frac{\partial^2 p}{\partial t^2} = c^2 \nabla^2 p \tag{2.13}$$

where c is constant and can either be c_T or c_s.

For simplicity, we will also look at solutions in one dimension, so the wave equation becomes

$$\frac{\partial^2 p}{\partial t^2} = c^2 \frac{\partial^2 p}{\partial x^2} \tag{2.14}$$

Different forms of solutions to this equation exist, which all bring a somewhat different insight into the problem. Let us look at them in turn. We will first consider the propagation of sound waves in an infinite (1D) domain, so $x \in (-\infty, +\infty)$ and no boundary conditions are applied (only initial conditions). We will then look at waves in an acoustic cavity (x is limited to a given interval and boundary conditions are applied at the edge of the interval).

2.2.1 D'Alembert's solution in an infinite domain

D'Alembert's solution is derived by noting that it is possible to factor the wave equation in the following way:

$$(\partial_t - c\partial_x)(\partial_t + c\partial_x)p = 0 \tag{2.15}$$

Using this information, let's now re-map the (x, t) space into a (ξ, η) space[1] where

$$\eta = x - ct \text{ and } \xi = x + ct \tag{2.16}$$

With this transformation, we conveniently have that

$$\partial_\eta = \frac{1}{2}\left(\partial_x - \frac{1}{c}\partial_t\right) \text{ and } \partial_\xi = \frac{1}{2}\left(\partial_x + \frac{1}{c}\partial_t\right) \tag{2.17}$$

so that the original equation simply becomes

$$\partial_{\eta\xi} p = 0 \tag{2.18}$$

This has the solutions

$$p(x, t) = f(\eta) + g(\xi) = f(x - ct) + g(x + ct) \tag{2.19}$$

where f and g are two arbitrary constants, that can be determined by applying initial conditions to the problem.

Suppose that $p(x, 0) = p_0(x)$ and $p_t(x, 0) = q_0(x)$. Then we know that

$$\begin{aligned} p_0(x) &= f(x) + g(x) \\ q_0(x) &= -cf'(x) + cg'(x) \end{aligned} \tag{2.20}$$

[1] While this remapping appears to have been pulled out of a hat, note that there is a sound mathematical theory behind it, that can be applied to find such useful transformation for any second-order PDEs (see the theory of canonical forms in AMS212A for detail).

Integrating the second equation between 0 and x yields

$$f(0) - f(x) + g(x) - g(0) = \frac{1}{c}\int_0^x q_0(x')dx' \tag{2.21}$$

Eliminating f yields g, and vice versa so

$$f(x) = \frac{1}{2}\left(p_0(x) - \frac{1}{c}\int_0^x q_0(x')dx' + f(0) - g(0)\right)$$
$$g(x) = \frac{1}{2}\left(p_0(x) + \frac{1}{c}\int_0^x q_0(x')dx' + g(0) - f(0)\right) \tag{2.22}$$

While new arbitrary integration constants appear here, they cancel out in the final solution:

$$p(x,t) = f(x-ct)+g(x+ct) = \frac{1}{2}\left(p_0(x+ct) + p_0(x-ct)\right) + \frac{1}{2c}\int_{x-ct}^{x+ct} q_0(x')dx' \tag{2.23}$$

This solution is called d'Alembert's solution, and uniquely specifies a solution to the sound-wave equation (in the 1D infinite domain) from its initial conditions.

We see that if $q_0(x) = 0$, then the solution is the sum of 2 perturbations that are each half the amplitude of the initial one, and travel respectively left-ward and right-ward away from the source at velocity c, without change of form. If $q_0(x) \neq 0$ on the other hand, the solution becomes asymmetric (Homework), but still contains two *packets* of information that travel as above, one to the left and one to the right.

Worked Problem: What is d'Alembert's solution for a Gaussian wave packet $p_0(x) = p_0 \exp(-x^2/2)$ initially at rest ($q_0 = 0$)?

Using the formula we simply have

$$p(x,t) = \frac{p_0}{2}\exp\left(-\frac{(x-ct)^2}{2}\right) + \frac{p_0}{2}\exp\left(-\frac{(x+ct)^2}{2}\right) \tag{2.24}$$

The solutions are shown in Figure 2.1. After a short while, the initial Gaussian splits into two Gaussians, each with half the original amplitude. One moves to the left at velocity c, one moves to the right at velocity c. Once the two traveling perturbations are sufficiently far away from each other that they no longer overlap, the solution simply looks like two traveling Gaussians that each satisfy the 1st order advection equations

$$(\partial_t + c\partial_x)p = 0 \text{ and } (\partial_t - c\partial_x)p = 0 \tag{2.25}$$

respectively. In this sense, the first-order PDEs

$$\partial_t p \pm c\partial_x p = 0 \tag{2.26}$$

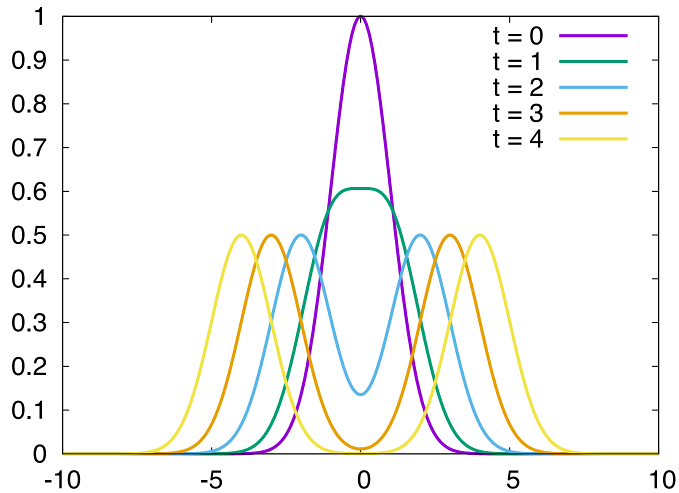

Figure 2.1: d'Alembert's solution for a Gaussian wave packet with $p_0(x) = \exp(-x^2/2)$ with $c = 1$ at various times.

are also often referred-to as wave equations as well, although in my opinion that is somewhat confusing.

D'Alembert's solution thus illustrates one of the most important properties of sound waves (and of all non-dispersive waves), namely that of *propagation without change of form* (at least, in the Cartesian 1D problem). The speed at which each of these packets propagates is called *the group speed*. Here, we see that the group speed is simply equal to $\pm c$ depending on the direction of propagation.

2.2.2 Monochromatic wave solution of the wave equation in an infinite domain

Another type of solution of the wave equation that is also valid in an infinite domain is one of the form

$$p(x,t) = \hat{p}e^{ikx-i\omega t} \tag{2.27}$$

In this expression, $\hat{p}$ can be complex. These are called *monochromatic plane wave solutions*. Note that, written as such, $p(x,t)$ looks like it could be complex. In fact it is (unless $\hat{p}$ is chosen just right), but that's not really a problem. In fact, it will be useful to keep p written as is and allow for complex solutions, as long as we remember at the end of the calculation that only the real part of this function is physical.

Plugging this back into the wave equation, we see that this is indeed a solution, provided $\omega^2 = c^2k^2$. Indeed,

$$\frac{\partial^2 p}{\partial t^2} = -\omega^2 \hat{p}(k) e^{ikx-i\omega t} = -\omega^2 p \tag{2.28}$$

while

$$\frac{\partial^2 p}{\partial x^2} = -k^2 \hat{p}(k) e^{ikx-i\omega t} = -k^2 p \tag{2.29}$$

So, $\partial^2 p/\partial t^2 = c^2 \partial^2 p/\partial x^2$ becomes

$$\omega^2 = c^2 k^2 \rightarrow \omega = \pm ck \tag{2.30}$$

This shows that the dispersion relation for sound waves, that is, the relation between the wave's *frequency* ω and the wave's *wavenumber* k, is linear. This is the defining property of non-dispersive waves.

We also see that there are two branches to the dispersion relation, one with $\omega = ck$ and one with $\omega = -ck$. In all that follows, we will adopt the convention that ω has to be positive, so the two branches correspond to two wavenumber of different sign: $k = \pm\omega/c$. The sign of k then also defines the direction of propagation. The dispersion relation then takes the more common form

$$\omega = c|k| \tag{2.31}$$

The true solution to the wave equation should therefore be written as a linear combination of the solutions in each branch:

$$\begin{aligned} p(x,t) \quad &= \hat{p}_+ e^{ik_+ x - i\omega t} + \hat{p}_- e^{ik_- x - i\omega t} \\ &= \hat{p}_+ e^{i|k|(x-ct)} + \hat{p}_- e^{-i|k|(x+ct)} \end{aligned} \tag{2.32}$$

Recalling that we only care about the real part of the solution, this should really be interpreted as

$$p(x,t) = \Re\left(\hat{p}_+ e^{i|k|(x-ct)} + \hat{p}_- e^{-i|k|(x+ct)}\right) \tag{2.33}$$

If the constants $\hat{p}_\pm$ are real, then

$$p(x,t) = \hat{p}_+ \cos(|k|(x-ct)) + \hat{p}_-(k)\cos(|k|(x+ct)) \tag{2.34}$$

while they are pure imaginary,

$$p(x,t) = -|\hat{p}_+|\sin(|k|(x-ct)) + |\hat{p}_-|\sin(|k|(x+ct)) \tag{2.35}$$

Any $\hat{p}$ that has both a real and an imaginary part will yield a solution for $p(x,t)$ that contains both a sine and a cosine function, but in all cases, we clearly recognize that the solution has two terms, a left-ward traveling wave and a right-ward traveling wave, as in d'Alembert's solution – the two descriptions of the sound waves are of course consistent with one another. In what follows, we now just consider the right-ward traveling wave, for simplicity.

As should now be obvious from their real expressions, these solutions have infinite spatial extent, and oscillate regularly in time and space with a frequency ω (and period $T = 2\pi/\omega$), and a wavenumber k (and wavelength $\lambda = 2\pi/|k|$). They can also be rewritten as

$$p(x,t) = \hat{p}e^{i\theta(x,t)} \tag{2.36}$$

where $\theta(x,t) = kx - \omega t = k(x - ct)$ (for the right-ward wave) is called the phase function. Later, we shall generalize the definition of the phase function, but in the meantime note the interesting relationship between partial derivatives of the phase function, and the frequency and wavenumber of the wave:

$$\omega = -\frac{\partial\theta}{\partial t} \text{ and } k = \frac{\partial\theta}{\partial x} \tag{2.37}$$

Given the expression for $p(x,t)$ in terms of θ, it is clear that p is constant if θ is. For instance, wave-crests and wave-trophs, (the maxima and minima of the function p) are achieved at specific values of θ. We can then ask the question: at which speed do the pressure maxima and minima travel? Since the maxima are given by constant values of θ, say, θ_c, their trajectories are given by $\theta_c = kx - \omega t$ which can be solved for x to yield $x_c(t) = (\theta_c + \omega t)/k$. The velocity of the maxima is the derivative of x_c with respect to time, and is therefore

$$c_p = \frac{\omega}{k} = c \tag{2.38}$$

for the right-propagating wave. This quantity is called the *phase speed* of the wave. The fact that the phase speed is equal to the group speed is another defining property of non-dispersive waves. This is actually the reason why they propagate without change of form. A good way to *visualize* the phase speed is to think about a problem with initial conditions $p_0(x) = cos(x)$. The pressure field later on is given by $p(x,t) = \cos(x - ct)$ (for the right-propagating wave). The solution is shown in Figure 2.2. At $t = 0$, the maxima (in yellow) are at $x = 0$, $x = 2\pi$, etc... Later on these maxima move to the right, at a constant velocity c. These form paths in the (x,t) plane that look like straight lines with slope $1/c$.

2.2.3 Superposition of monochromatic waves in an infinite domain

In general, the initial conditions for a sound wave are localized, not infinite – more like the Gaussian packet example than the infinite sine wave or cosine wave example. Imagine for instance a sound created by someone's vocal chords, by a loud speaker, etc.. As a result, the monochromatic plane waves studied above, which have the same amplitude everywhere in space and exist for all times, is not particularly relevant physically. Instead, the true solution is a linear superposition of monochromatic waves:

$$p(x,t) = \Re\left[\int \hat{p}(k)e^{ikx - i\omega(k)t}dk\right] \tag{2.39}$$

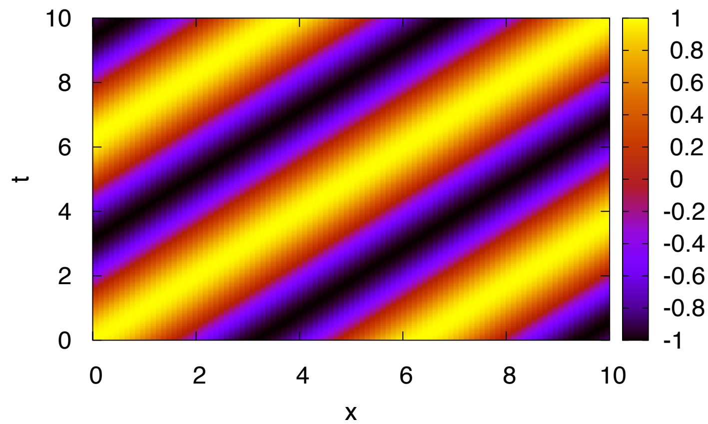

Figure 2.2: Right-ward propagating pressure wave with initial condition $p_0(x) = \cos(x)$, and velocity $c = 1$. The lines constant phase in the (x, t) plane have slope $1/c$.

where $\omega(k) = c|k|$. The $\hat{p}(k)$ is simply obtained from the Fourier transform of the initial conditions:

$$p_0(x) = p(x,0) = \Re\left[\int \hat{p}(k)e^{ikx}\right]$$
$$q_0(x) = p_t(x,0) = \Re\left[-i\int \hat{p}(k)\omega(k)e^{ikx}\right] \tag{2.40}$$

Using properties of Fourier transforms, and a bit of work, it can be shown that this is in fact *exactly* equivalent to d'Alembert's solution (as it should be !)

Superposition of waves thus yields exact solutions of the governing equations, valid at all times and for any initial conditions, and recover d'Alembert's solution. However, while d'Alembert's solution is not obvious to generalize in higher dimensions, wave superposition is, and is therefore the preferred method for general problems. The main difficulty, however, is in the interpretation of the solutions, which isn't particularly intuitive when written in this form (unless you fluently speak Fourier Transforms language).

Later, we shall construct other solutions of the wave equation that have a more intuitive form, and are more easily generalized to the case of non-constant sound-speed. Before we proceed, however, let's look a case we have ignored so far, namely that of waves in a finite domain.

2.2.4 Sound waves in an acoustic cavity

We now move away from the infinite domain case, and consider sound waves in a finite interval, often called an acoustic cavity. To create a 1D acoustic cavity, we have to consider the somewhat artificial problem of sound waves generated in a very narrow tube of finite length L, closed at both ends. The tube is filled with homogeneous fluid, in which the sound speed is c. The equation for the wave is therefore exactly the same as we have looked at earlier, namely (2.14).

This time, however, we need boundary conditions to describe what happens to the waves when they hit the ends of the tube. There are a number of possibilities, but a typical one would be to require that there be no fluid motion across the end of the tube. How do we relate this condition to the pressure? To do so, we need to go back to the original set of equations from which the wave equation was derived, and remember that the velocity perturbations and the pressure perturbations are related via

$$\frac{\partial u}{\partial t} = -\frac{\partial p}{\partial x} \tag{2.41}$$

in one dimension. Disallowing any fluid motion across the boundary implies that $u = 0$ at all times on the boundary, which then requires that

$$\frac{\partial p}{\partial x} = 0 \tag{2.42}$$

at both ends of the tube (e.g. at $x = 0$ and $x = L$).

Eigenmodes and eigenfrequencies

We now have to solve the sound wave equation (2.14) under the boundary conditions given above. To do this, we will use separation of variables, and look for *basic* solutions that can be written as a function of x times a function of t (cf. AMS 212A), namely $a(x)b(t)$. Plugging this into the wave equation, we get

$$a(x)\frac{d^2b}{dt^2} = c^2 b(t)\frac{d^2a}{dx^2} \tag{2.43}$$

Dividing both sides by $a(x)b(t)$ we get

$$\frac{1}{b(t)}\frac{d^2b}{dt^2} = c^2\frac{1}{a(x)}\frac{d^2a}{dx^2} \tag{2.44}$$

The LHS is a function of t only, while the RHS is a function of x only, so the two can only be equal for all x and t if they are both constants. Let that constant be α, for instance. It's quite easy to show that α cannot be positive (Hint: try it, and show that there are no solutions to the spatial problem in that case.). We are left with two cases

- $\alpha = 0$: In that case, we get $d^2a/dx^2 = 0$, whose solution is linear in x. The only solution that also satisfies the boundary conditions, which become $da/dx = 0$ at both ends, is the constant solution: $a_0(x) = C_0$.

- $\alpha < 0$: In that case, let's rewrite $\alpha = -\omega^2$ (which is by definition a negative number). We then have

$$\frac{d^2a}{dx^2} = -\frac{\omega^2}{c^2}a \tag{2.45}$$

which has oscillatory solutions:

$$a(x) = A\cos\left(\frac{\omega}{c}x\right) + B\sin\left(\frac{\omega}{c}x\right) \tag{2.46}$$

In order to satisfy the boundary conditions at $x = 0$, we need to have $B = 0$. In order to satisfy the boundary conditions at L, we need

$$\frac{\omega L}{c} = n\pi \tag{2.47}$$

where n is an integer number. This shows that the constant ω cannot take just any value, it is instead *quantized*, and can only take the values

$$\omega_n = \frac{n\pi c}{L} \tag{2.48}$$

and for these values of ω, the solutions are

$$a_n(x) = A_n\cos\left(\frac{n\pi x}{L}\right) \tag{2.49}$$

Note that since we effectively solved an eigenvalue problem for the eigenfunction a and the eigenvalue ω, the solutions $a_n(x)$ are called the *spatial eigenmodes* of the problem and ω_n are called the *eigenfrequencies*.

Now that we have solved the spatial problem, we now turn to the temporal problem:

$$\frac{d^2b}{dt^2} = -\omega^2 b \tag{2.50}$$

For each value of ω_n there will be a corresponding function b_n. The temporal solutions are

$$b_n(t) = C_n\cos(\omega_n t) + D_n\sin(\omega_n t) \tag{2.51}$$

for $n \geq 1$, while for the case where $\omega = 0$, we simply have $b_0(t)$ constant.

We have therefore constructed a series of basic solutions of the form:

$$\begin{aligned} a_0(x)b_0(t) &\propto \text{constant} \\ a_n(x)b_n(t) &\propto \cos(k_n x)\left[C_n\cos(\omega_n t) + D_n\sin(\omega_n t)\right] \text{ for } n \geq 1 \end{aligned} \tag{2.52}$$

where

$$k_n = \frac{n\pi}{L} \text{ and } \omega_n = \frac{n\pi c}{L} \text{ for } n \geq 1 \tag{2.53}$$

The actual solution for $p(x,t)$ is then a linear combination of all these basic solutions:

$$p(x,t) = C_0 + \sum_{n\geq 1}\left(C_n\cos(\omega_n t) + D_n\sin(\omega_n t)\right)\cos(k_n x) \tag{2.54}$$

The integration constants C_n and D_n are determined from the initial conditions of the system.

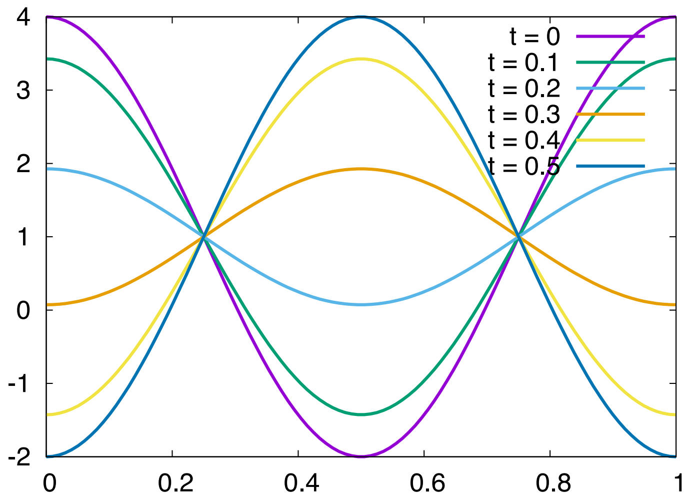

Figure 2.3: Half a period of the solution to the sinusoidal initial condition problem, with $c = 1$.

Worked Example 1: Find the solution to the 1D wave equation in a tube of length 1 with initial conditions $p(x, 0) = 1 + 3\cos(2\pi x)$ and $p_t(x, 0) = 0$.

In this case, we set

$$p(x,0) = 1 + 3\cos(2\pi x) = C_0 + \sum_{n\geq 1} C_n \cos(k_n x)$$
$$p_t(x,0) = 0 = \sum_{n\geq 1} D_n \omega_n \cos(k_n x) \tag{2.55}$$

where $k_n = n\pi$. The second equation implies that $D_n = 0$ for all n, while the first implies that $C_0 = 1$ and $C_2 = 3$, with all other $C_n = 0$, so the complete solution is

$$p(x,t) = 1 + 3\cos(2\pi ct)\cos(2\pi x) \tag{2.56}$$

We see that the solution is a standing wave, i.e. a wave whose nodes remain in place, and whose amplitude changes with time. This is illustrated in Figure 2.3.

This example is somewhat trivial, because the initial condition *is* basically a single eigenmode of the problem. It nevertheless illustrates an important property of the waves in a finite domain, namely that only those eigenmodes that are excited through the initial conditions actually contribute to the solution – all the others that have 0 amplitude at $t = 0$ continue to have 0 amplitude later on. In general, however, a given set of initial conditions will excite the entire spectrum of eigenmodes. The mathematical derivation of the full solution is then a little more complicated but the principle is the same.

Worked example 2: Find the solution to the 1D wave equation in a tube of length 10 with initial conditions $p(x, 0) = \exp(-(x-5)^2/2)$ and $p_t(x, 0) = 0$.

In this case, we have

$$\begin{aligned} p(x, 0) = \exp(-(x-5)^2/2) &= C_0 + \sum_{n \geq 1} C_n \cos(k_n x) \\ p_t(x, 0) = 0 &= \sum_{n \geq 1} D_n \omega_n \cos(k_n x) \end{aligned} \tag{2.57}$$

so we still have $D_n = 0$ for all n, but identifying the C_n is a little more difficult. Here, it's useful to remember some of the orthogonality properties of sine and cosine functions, namely that

$$\begin{aligned} \int_0^L \cos\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) dx &= 0 \text{ for any } n \text{ and } m \\ \int_0^L \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) dx &= \frac{L}{2}\delta_{nm} \\ \int_0^L \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) dx &= \frac{L}{2}\delta_{nm} + \frac{L}{2}\delta_{n0}\delta_{m0} \end{aligned} \tag{2.58}$$

Using these relationships, we then have

$$\int_0^L \exp(-(x-5)^2/2) \cos(k_m x) dx = \int_0^L \left[C_0 + \sum_{n \geq 1} C_n \cos(k_n x) \right] \cos(k_m x) dx \tag{2.59}$$

which shows that

$$\begin{aligned} C_0 &= \frac{1}{L} \int_0^L \exp(-(x-5)^2/2) dx \simeq 0.25066... \\ C_n &= \frac{2}{L} \int_0^L \exp(-(x-5)^2/2) \cos(k_n x) dx \end{aligned} \tag{2.60}$$

where here, $L = 10$. Now, these integrals are not necessarily easy to evaluate analytically, but they can certainly be calculated numerically, so we can do that if we wish to compute $p(x, t)$ for any time $t > 0$. Computing the first few, we get $C_1 = 0$, $C_2 = -0.411523..$, $C_3 = 0$, $C_4 = 0.2276..$, $C_5 = 0$, $C_6 = -0.16077..$, $C_7 = 0$, $C_8 = 0.0213..$, $C_9 = 0$ and $C_{10} = -0.0036..$ etc.. The solution including all the modes up to $n = 10$ is shown in Figure 2.4.

This time, we see that the initial conditions are fairly localized within the interval, and the initial behavior is actually very similar to what happened in the infinite domain case: the initial Gaussian splits into two wave packets that travel in opposite directions. This is not surprising since, for early times, the

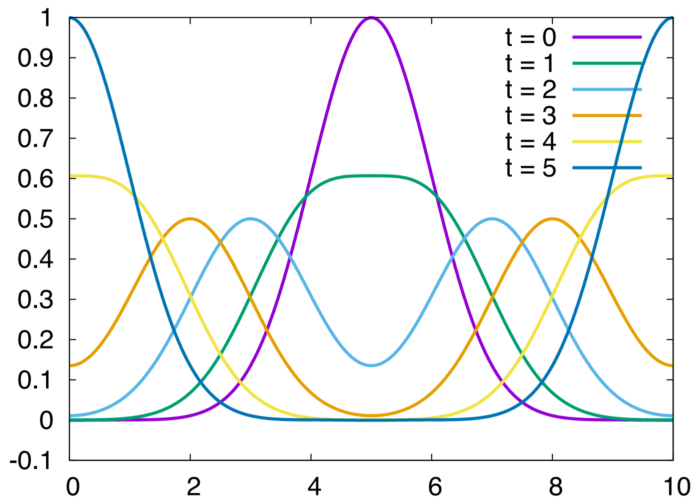

Figure 2.4: Half a period of the solution to the Gaussian initial condition problem in a finite interval, with $c = 1$.

pressure wave does not know yet it's about to encounter boundaries. Later on, however, each packet hits the boundary, piles up there, reflects, and comes back towards the center – and so the pattern repeats. We therefore see a little bit of both behaviors: the solution is indeed a linear combination of standing waves, but these do also add up to form traveling waves.

This is actually a very general statement. Let's recall that *any* standing wave can actually be represented by 2 counter-propagating waves since

$$\cos(kx)\cos(\omega t) = \frac{1}{2}\cos(kx - \omega t) + \frac{1}{2}\cos(kx + \omega t) \tag{2.61}$$

(and similarly for the product of any two sine and cosine function) so it is in fact possible to recast equation (2.54) in the form (2.19), although this time the left-propagating and right-propagating solutions are not expressed as integrals over all k (as in equation (2.39)), but as an infinite discrete sum instead:

$$\begin{aligned} p(x,t) \quad &= \frac{1}{2}\sum_n C_n \cos\left(\frac{n\pi(x+ct)}{L}\right) + C_n \cos\left(\frac{n\pi(x-ct)}{L}\right) \\ &+ D_n \sin\left(\frac{n\pi(x-ct)}{L}\right) + D_n \sin\left(\frac{n\pi(x+ct)}{L}\right) = f(x-ct) + g(x+ct) \end{aligned} \tag{2.62}$$

We also have that, to each eigenmode of wavenumber $k_n = n\pi/L$ corresponds an eigenfrequency of value $\omega_n = n\pi c/L = ck_n$. This recovers the standard dispersion relation for sound waves. We thus see that the only *real* difference between the infinite and finite domain cases is the quantization of the allowable wavenumbers and frequencies. We will come back to this later. Even so, note how as $L \to \infty$, the spacing between the wavenumbers ($\delta k = \pi/L$) and between their corresponding eigenfrequencies ($\delta\omega = c\pi/L$) both go to zero – in other words, although discrete for any strictly finite L, the infinite sum over all modes does converge in a Riemann sense to the integral expression appropriate in the case of an infinite domain.

Properties of eigenmodes and eigenfrequencies of Sturm-Liouville problems

In any case, both examples reveal that the wave field in the tube can be expressed as a finite or infinite sum of *global eigenmodes*, each of which is a standing wave that has a given spatial structure $\cos(n\pi x/L)$, and oscillates with the associated *eigenfrequency* $\omega_n = n\pi c/L$. The eigenmodes are global in the sense that they span the entire acoustic cavity, and obtained by requiring that the sound wave satisfies the boundary conditions on either side of it. Note how the eigenmodes and eigenfrequencies are completely independent of the initial conditions applied – they are an intrinsic property of the tube itself (notably its geometry, and of the boundary conditions applied), and of the equation that characterizes the waves. As we have seen in AMS212A, and as we shall see again later, this statement is also true in 2D and 3D acoustic cavities.

The spatial eigenproblem associated with the 1D wave equation is called a Sturm-Liouville problem. Many types of 1D wave equations, as long as they are linear, lead to such Sturm-Liouville problems so there has been a lot of work done in Applied Mathematics in the last few hundred years to derive general properties of solutions of these problems. See AMS212A for a comprehensive list, and for a more formal introduction to Sturm-Liouville theory. However, for our purposes, there are a few very interesting general properties of the eigenmodes and eigenfrequencies of these problems (and therefore, by extension, of the wave equation), that deserve discussion:

- The eigenvalues of Sturm-Liouville problems form an ordered set that has a smallest eigenvalue, but that does not have a maximum eigenvalue, e.g. $\omega_0 \leq \omega_1 \leq \omega_2 \leq ... < +\infty$. For regular Sturm-Liouville problems, the set is actually strictly ordered (i.e. no two eigenvalues are the same).

- The eigenfunction associated with the smallest eigenvalue does not have a node, i.e. it does not vanish within the domain. It is called the *fundamental eigenmode*. In the case studied here, the fundamental was the constant function.

- The period of the fundamental mode is usually related to the sound-crossing time across the cavity. For instance, in the example shown here, the fundamental mode has a frequency $\pi c/L$, so its period is $2\pi/\omega_1 = 2c/L$, which is the time it takes for the sound to go across the cavity and back. This is the reason why the larger a music instrument, the lower its pitch (e.g. compare the typical sound of the bass to that of the violin).

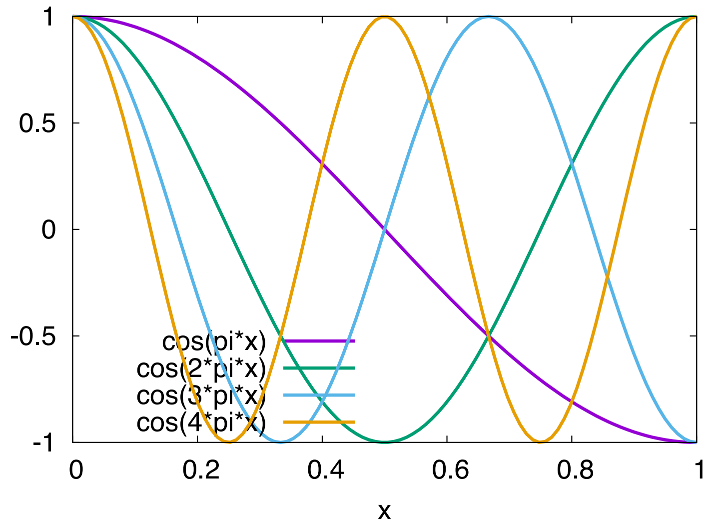

Figure 2.5: Various eigenmodes of the thin tube problem.

- The eigenfunctions associated with successively higher eigenvalues have successively more nodes. In the example we saw here, the one associated with ω_1 is the function $\cos(\pi x)$ and has 1 node, the one associated with ω_2 is $\cos(2\pi x)$ which has 2 nodes, etc... Hence the larger the eigenfrequency, the more spatially complex the eigenmode is, as shown in Figure 2.5

Wave reflection, and a physical interpretation of quantization

Mathematically speaking, the emergence of a discrete set of eigenmodes and eigenfrequencies takes its roots in Sturm-Liouville theory. But physically, why should this be true? To answer this question, we must first understand what happens when a wave reflects off the end of the tube.

To model the reflection of a wave, we must consider the total solution near the wall as the sum of an incoming wave and a reflected wave. If the wall is at $x = 0$, and say, the incoming wave is coming from the right, then the pressure perturbations associated with that incoming wave are

$$p_I(x,t) = \Re\left[\hat{p}_I e^{-i|k|(x+ct)}\right] \tag{2.63}$$

while the reflected wave moves to the right, so

$$p_R(x,t) = \Re\left[\hat{p}_R e^{i|k|(x-ct)}\right] \tag{2.64}$$

The total solution is simply $p = p_I + p_R$ (note that we have assumed here that the frequency and wavenumber of the incident and reflected waves are the same aside from the sign describing the direction of propagation; this will be shown later in the more general case of reflection in multiple dimensions). To guarantee that the real part of $\partial_x p = 0$ at $x = 0$ at all times, we have to have

$$\Re\left[\partial_x p_I(0,t) + \partial_x p_R(0,t)\right] = 0 \to \Re\left[-i|k|\hat{p}_I e^{-i|k|ct} + i|k|\hat{p}_R e^{-i|k|ct}\right] = 0 \tag{2.65}$$

A simple way of guaranteeing that this is true at all times is to take $\hat{p}_R = \hat{p}_I$, or in other words, the amplitude of the incoming wave is simply the same[2] as the amplitude of the outgoing wave so

$$p_R(x,t) = \hat{p}_I e^{i|k|(x-ct)} \tag{2.66}$$

Let's now consider again the question of quantization. Suppose a wave is generated at a point x_0 in the tube, and suppose it travels to the left, reflects off the wall at $x = 0$, then reflects off the wall at $x = L$ and comes back to x_0 at which point it interferes with itself. Since the phase is given by $kx - \omega t$, the phase difference between the original and reflected waves at $x = x_0$ is simply given by

$$\Delta\theta = kx_0 - (kx_0 - \omega T) = \omega T \tag{2.67}$$

where T is the time it took for the wave to travel from x_0 to the two walls and back. Since the phase speed is c, that travel time is simply $2cL$, hence $\Delta\theta = 2\omega cL$.

If $\Delta\theta$ is not a multiple of 2π, the wave interferes destructively with itself. Indeed, each time the wave bounces, it comes back with a different shifted phase. Since the sum of an infinite number of oscillatory functions of different phases adds up to 0, the only way to have a non-zero outcome is to require that $\Delta\theta$ be a multiple of 2π. This, in turn, implies that ω cannot take any values, but instead is restricted to be

$$\omega_n = \frac{n\pi c}{L} \tag{2.68}$$

which is exactly the quantization condition we were looking for.

We therefore see that

- the descriptions of the wave solutions are all consistent with one another.
- it is possible to obtain the spectrum of eigenfrequencies of an acoustic cavity without calculating its eigenmodes, but simply by calculating the phase shift due to the travel time of a wave along its path and requiring that the waves interact constructively.

This second property is particularly useful when we try to calculate eigenfrequencies of oscillations of an acoustic cavity in more than 1 dimension, and for non-constant sound speed.

---

[2]Note that had we chosen a different set of boundary conditions at the end of the tube, as for instance $p = 0$ instead of $\partial_x p = 0$, then the amplitude of the reflected wave would be minus that of the incoming one, effectively resulting in a phase shift of π between the two waves. More on this later.