Lecture 1 & background: The wave equation throughout physics

1 The aim of the course:

1.1 What we'll cover:

This course will cover some mathematical methods for solving the wave equation, with a particular emphasis on the interaction of waves with materials. We will use a mixture of pen–and–paper methods and simple computer programs we'll write ourselves. My hope is that the course gives you a box of mathematical tricks that you didn't get from your undergraduate lectures, which you can use to solve problems in your research. I also hope to convey how wonderful waves are: slight variations of one equation describe an enormous number of physical phenomena and lie at the heart of mysterious topics such as special relativity and quantum mechanics. You also get to experience the consequences of the wave equation every day, with nearly all of your senses.

After this lecture, the topics covered will be:

- **Sources:** Green functions and the mathematics of wave generation.
- **Pulses:** Wave–packets as a sum over modes, and the effect of material dispersion on reshaping pulses.
- **Beams:** The paraxial approximation, its connection with the Schrodinger equation and an approximate theory of diffraction.
- **Boundary conditions:** Different kinds of boundary condition, and their relationship to the power carried by a wave.
- **Scattering—one obstacle:** Some exact and approximate theories of scattering from a single object.
- **Scattering—an infinite number of obstacles (part 1):** An introduction to basic diffraction in 1D, including an exactly soluble model. The transfer matrix method.
- **Scattering—an infinite number of obstacles (part 2):** Perturbation theory applied to two and three dimensional periodic structures. The importance of symmetry groups.
- **Inhomogeneous media:** The WKB approximation, geometrical optics and some exact solutions in spatially changing materials.
- **Curiosities:** A collection of interesting wave phenomena (including supersymmetric quantum mechanics and reflectionless potentials)

At the end of some lectures I will set homework problems. The idea is to get you to implement something covered in the lectures that you will hoperfully use again.

1.2 Why do we do maths, can't we just use COMSOL?:

Put simply, using commercial packages does not teach you very much about physics. They are easy to operate, and useful tools for particularly complicated problems. But after completing a PhD in physics, I believe you should be able to do physics on your own. It's bad news for everyone if you get to the end of your research degree and lose your ability to do physics simply because you lose access to powerful computers and expensive modelling software.

2 The wave equation in several areas of physics:

Probably the wave equation you are most familiar with is that for a scalar wave ϕ (perhaps a sound wave) propagating through a homogeneous medium at a speed c

$$\boxed{\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}.} \tag{1}$$

It's useful to try to read an equation without solving it, and we can read equation (1) as follows. The left hand side is the Laplacian acting on the field, which measures how the value of ϕ curves in space. For instance, basic calculus tells us that if $\phi(\boldsymbol{x})$ is at a maximum then $\nabla^2\phi(\boldsymbol{x})$ will be negative, and vice versa for a minimum. Meanwhile the right hand side of (1) is the acceleration in the value of ϕ. Therefore (1) says that in general where ϕ is a maximum in space, its value will accelerate so as to reduce (at a rate determined by c^2), and where ϕ is a minimum it will accelerate so as to increase. So with very little work, we can already see that (1) generically has solutions that must undulate in space and time—these are waves!

Now I want to show how equations like (1) arise in most areas of physics. It's not essential that you understand all of these sections and we certainly won't cover them all in the lecture. I just want to give you the impression of the common threads that tie together different areas of physics, and how the wave equation is an important connecting strand.

2.1 Electromagnetic waves:

Electromagnetic radiation is emitted due to the acceleration of electric charges, and propagates through space as polarized waves of interconnected electric and magnetic fields. We can derive the behaviour of these waves directly from Maxwell's equations. In a material medium (a piece of glass, a sheet of metal, a bath of water...) Maxwell's equations take the form (see [6] and [13])

$$\begin{aligned} \nabla\cdot\boldsymbol{D} &= 0 \\ \nabla\cdot\boldsymbol{B} &= 0 \\ \nabla\times\boldsymbol{E} &= -\frac{\partial\boldsymbol{B}}{\partial t} \\ \nabla\times\boldsymbol{H} &= \frac{\partial\boldsymbol{D}}{\partial t} \end{aligned} \tag{2}$$

where $\boldsymbol{E}$ is the average electric field and $\boldsymbol{B}$ is the average magnetic field. Note that these equations usually assume that the microscopic details of the material such as the electrons and the nuclei have been averaged out. The vector $\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P}$ includes the electric field and the polarization $\boldsymbol{P}$ of the material, and $\boldsymbol{H} = \boldsymbol{B}/\mu_0 - \boldsymbol{M}$ includes the magnetic field and the magnetization $\boldsymbol{M}$. When the average fields are relatively weak the polarization and magnetization can be taken as linear functions of the fields $\boldsymbol{P} \sim \varepsilon_0\chi_E\boldsymbol{E}$ and $\boldsymbol{M} \sim \mu_0\chi_H\boldsymbol{H}$ so that we can write $\boldsymbol{D}$ and $\boldsymbol{B}$ as functions of $\boldsymbol{E}$ and $\boldsymbol{H}$

$$\begin{aligned} \boldsymbol{D} &= \varepsilon_0\boldsymbol{E} + \varepsilon_0\chi_E\boldsymbol{E} &&= \varepsilon_0\varepsilon\boldsymbol{E} \\ \boldsymbol{B} &= \mu_0\boldsymbol{H} + \mu_0\chi_H\boldsymbol{H} &&= \mu_0\mu\boldsymbol{H} \end{aligned} \tag{3}$$

These are the *constitutive relations* of the material. Taking the curl of the third of (2), applying both the constitutive relations (3), and the last of (2) we find

$$\boxed{\begin{aligned} \nabla\times(\nabla\times\boldsymbol{E}) &= -\frac{\partial\nabla\times\boldsymbol{B}}{\partial t} = -\mu_0\mu\frac{\partial\nabla\times\boldsymbol{H}}{\partial t} = -\mu_0\mu\frac{\partial^2\boldsymbol{D}}{\partial t^2} \\ &= -\frac{\mu\varepsilon}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2} \end{aligned}} \tag{4}$$

where we used $\varepsilon_0 \mu_0 = c^{-2}$, where c is the speed of light in vacuum. This is the *vector wave equation*. For a fixed frequency $\boldsymbol{E} = \boldsymbol{E}_\omega e^{-i\omega t}$ this reduces to

$$\boxed{\nabla \times (\nabla \times \boldsymbol{E}_\omega) - \frac{\mu\varepsilon\omega^2}{c^2}\boldsymbol{E}_\omega = 0} \tag{5}$$

which is the *vector Helmholtz equation*. Applying the triple vector product $\nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}$ we can see that in the cases where $\nabla \cdot \boldsymbol{E} = 0$ (i.e. in the absence of any charge) (4) reduces to

$$\boxed{\nabla^2 \boldsymbol{E} = \frac{\varepsilon\mu}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2}} \tag{6}$$

which is the *wave equation* (1), but now for the electric field inside a material. For a field of a fixed frequency $\boldsymbol{E} = \boldsymbol{E}_\omega e^{-i\omega t}$ (6) reduces to

$$\boxed{\nabla^2 \boldsymbol{E}_\omega + \frac{\varepsilon\mu\omega^2}{c^2}\boldsymbol{E}_\omega = 0} \tag{7}$$

which is the *Helmholtz equation*. Notice that the speed of the wave in (6) is modified from the free space value, and equals $c/\sqrt{\varepsilon\mu}$. For typical bulk materials (not metamaterials!) $\varepsilon\mu > 1$, so that the light wave is slowed down.

Note that of the two equations (6) and (7), equation (7) is actually the more accurate in practical situations. How can this be so? Because in real life ε and μ are always functions of frequency (an effect known as *dispersion*), and equations (3) and (6) ignore this frequency dependence, while it can be readily included in (7). When we wrote down equation (3) we implicitly assumed that the polarization does not depend on the value of the electric field in the past, which in general it does. We'll come back to the phenomenon of dispersion in lecture 3.

2.2 Acoustic pressure waves:

Sound propagates through a gas or a liquid as a pressure wave. Perhaps unsurprisingly, the behaviour of this pressure wave can be derived from fluid dynamics. You may not be familiar with much fluid dynamics, so I'll briefly derive the fundamental equation that governs the motion of a non–viscous fluid (for more details see [4]).

Fluid dynamics can be derived from two simple physical principles. The first principle is the conservation of mass: if the total mass of fluid in a volume of space V changes, it's because some fluid flowed in or out through the boundary[1] ∂V.

Rate of change of mass in V =Rate of mass flowing in through the boundary ∂V

If the fluid mass density is ρ and the flow velocity is $\boldsymbol{v}$ then the above statement can be given mathematically as

$$\begin{aligned} \frac{\partial}{\partial t}\int_V \rho dV &= -\int_{\partial V} \rho \boldsymbol{v} \cdot d\boldsymbol{S} \\ \rightarrow \quad \frac{\partial \rho}{\partial t} &= -\nabla \cdot (\rho \boldsymbol{v}) \end{aligned} \tag{8}$$

where in the final step I used the divergence theorem $\int_V \nabla \cdot \boldsymbol{V} dV = \int_{\partial V} \boldsymbol{V} \cdot d\boldsymbol{S}$ and the fact that the volume V is arbitrary (so the integrand itself must be zero). The minus sign is because the mass increases when $\boldsymbol{v}$ points into the volume, and $d\boldsymbol{S}$ by convention points outwards. Equation (8) is known as the *continuity equation*.

1. Here we use the symbol '∂' to mean 'boundary of'.

The second principle is the conservation of momentum: if the total momentum in a volume V changes then this is the sum of the momentum flowing through the boundary ∂V and the force acting on the boundary ∂V.

$$\textit{Rate of change of momentum in } V = \textit{Rate of momentum flowing in through } \partial V + \textit{Force on } \partial V$$

Given that the momentum per unit volume is $\rho\boldsymbol{v}$, and the force on ∂V is due to the pressure p in the fluid, the above statement is equivalent to

$$\frac{\partial}{\partial t}\int_V \rho\boldsymbol{v} dV = -\int_{\partial V} \rho\boldsymbol{v}\,(\boldsymbol{v}\cdot d\,\boldsymbol{S}) - \int_{\partial V} p\,d\,\boldsymbol{S} \tag{9}$$

Applying the divergence theorem to the right hand side, and grouping all terms together (9) becomes

$$\int_V \left[\left(\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\boldsymbol{v})\right)\boldsymbol{v} + \rho\frac{\partial\boldsymbol{v}}{\partial t} + \rho\boldsymbol{v}\cdot\boldsymbol{\nabla}\,\boldsymbol{v} + \boldsymbol{\nabla}p\right] dV = 0 \tag{10}$$

Combining the conservation of mass (8) and the conservation of momentum (10) we find what we're looking for

$$\boxed{\frac{\partial\boldsymbol{v}}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\,\boldsymbol{v} = -\frac{1}{\rho}\boldsymbol{\nabla}p} \tag{11}$$

which is known as *Euler's equation*. This is the basic equation of invicid (non–viscous) fluid dynamics. It's one of the simpler equations of fluid dynamics, but is non–linear and still bloody hard to solve in almost all cases.

Now we can move on to what we're interested in: sound waves are typically small disturbances that travel through the fluid. We write the density, pressure and velocity as background values (subscript '0') plus this small perturbation (subscript '1').

$$\begin{aligned} \rho &= \rho_0 + \rho_1 \\ p &= p_0 + p_1 \\ \boldsymbol{v} &= \boldsymbol{v}_1 \end{aligned} \tag{12}$$

Using the fact that the perturbation is small we can re–write (11) as (squares of small numbers are really small, and we neglect them)

$$\rho_0\frac{\partial\boldsymbol{v}_1}{\partial t} = -\boldsymbol{\nabla}p_1 \tag{13}$$

Similarly when written in terms of the quantities defined in (12), the continuity equation (8) becomes

$$\frac{\partial\rho_1}{\partial t} = -\boldsymbol{\nabla}\cdot(\rho_0\boldsymbol{v}_1) \tag{14}$$

we now use the isentropic compressibility β_s that comes from thermodynamics [4] to relate the change in density to the change in pressure

$$\frac{\partial\rho_1}{\partial t} = \left(\frac{\partial\rho_1}{\partial p_1}\right)_S \frac{\partial p_1}{\partial t} = \rho_0\beta_s\frac{\partial p_1}{\partial t} \tag{15}$$

(β_s tells us how easy it is to change the volume of a fluid by applying a pressure, assuming no change in entropy). Equation (15) can be thought of as the constitutive relation of pressure acoustics, analogous to the relationship between $\boldsymbol{D}$ and $\boldsymbol{E}$ (3). Applying (15) to (14) and differentiating with respect to time

$$\begin{aligned} \rho_0\beta_s\frac{\partial p_1}{\partial t} &= -\boldsymbol{\nabla}\cdot(\rho_0\boldsymbol{v}_1) \\ \rightarrow\quad \rho_0\beta_s\frac{\partial^2 p_1}{\partial t^2} &= -\boldsymbol{\nabla}\cdot\left(\rho_0\frac{\partial \boldsymbol{v}_1}{\partial t}\right) \end{aligned}$$

Finally using (13) we have

$$\boxed{\nabla^2 p_1 = \rho_0\beta_s\frac{\partial^2 p_1}{\partial t^2}} \tag{16}$$

which shows that small disturbances to the pressure propagate in a fluid according to the wave equation, with the wave speed given by $\sqrt{1/\rho_0\beta_s}$. The sound you can hear at this moment is this kind of wave, and in air (standard atmospheric conditions) $\rho_0 = 1.225$ kg m^{-3} and $\beta_s = 7.04\times 10^{-6}\text{m}^2/\text{N}$ so that $c = 340\,\text{m}\,\text{s}^{-1}$ Again, for constant frequency $p_1 = p_\omega e^{-i\omega t}$ this reduces to the Helmholtz equation

$$\boxed{\nabla^2 p_\omega + \rho_0\beta_s\omega^2 p_\omega = 0} \tag{17}$$

Note that in electromagnetism we can change the refractive index through changing either ε or μ, while to change the refractive index for an acoustic wave we must change the background density ρ_0, or the isentropic compressibility β_s.

2.3 Elastic waves:

When a sound wave passes through a fluid, the local forces within the fluid are described by a pressure p. This simple description is no longer possible in a solid. A pressure pushes with the same magnitude on the boundary of a volume element dV, whatever the orientation of the boundary. This serves to compress or expand dV. However, even within an isotropic solid there are more complicated local forces that do not simply locally compress or expand the solid. These are shear forces, and arise when e.g. one tries to slide two crystalline planes over one another.

To describe the forces within the body we use the *stress* rather than the pressure. Like the pressure, the stress is the local force per unit area, but unlike the pressure which is a scalar, stress is a tensor with two indices σ_{ij} which we shall write as the bold symbol $\boldsymbol{\sigma}$. The force $d\boldsymbol{f}$ pushing on a given surface element within the solid $d\boldsymbol{S}$ is given by $d\boldsymbol{f} = \boldsymbol{\sigma}\cdot d\boldsymbol{S}$, which in index notation is $df_i = \sum_j \sigma_{ij} dS_j$. We can immediately see that using the stress tensor means that the orientation of the surface element can make a difference to the size of the force pushing it. In the special case where $\sigma_{ij} = -p\delta_{ij}$ (δ_{ij} is the Kronecker delta), the force $d\boldsymbol{f}$ reduces to that due to a pressure as given in (10).

To derive the wave equation for sound propagating through a solid we again consider the change of momentum in a volume of solid of mass density ρ. In this case we characterize the elastic deformation of the solid using the vector $\boldsymbol{u}(\boldsymbol{x})$, which tells us how much the point $\boldsymbol{x}$ has been moved due to the deformation. The momentum density is thus $\rho\partial\boldsymbol{u}/\partial t$ and momentum conservation requires that

Rate of change of momentum in $V =$ *Total force acting on the boundary* ∂V

which mathematically is written as

$$\frac{\partial}{\partial t}\int_V \rho\frac{\partial \boldsymbol{u}}{\partial t}dV = \int_{\partial V}\boldsymbol{\sigma}\cdot d\boldsymbol{S} \tag{18}$$

Using the divergence theorem and the fact that the volume V is arbitrary (18) becomes

$$\rho\frac{\partial^2\boldsymbol{u}}{\partial t^2} = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma} \tag{19}$$

As in the previous section we now need to specify a constitutive relation to complete the description of the problem. In this case the constitutive relation links the stress $\boldsymbol{\sigma}$ to the displacement of the points in the solid $\boldsymbol{u}$. For small deformations we can use *Hooke's law* (stress is proportional to strain), which sounds simple, but in tensor form is quite a complicated law

$$\sigma_{ij} = -\sum_{kl} c_{ijkl}\,\epsilon_{kl} \tag{20}$$

where c_{ijkl} is the four index *elasticity tensor* and ϵ_{kl} is the *strain tensor*, defined as

$$\epsilon_{kl} = \frac{1}{2}\left[\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right]. \tag{21}$$

Note that strain depends on derivatives of the displacement $\boldsymbol{u}$. This is because strain is differential displacement: moving the body as a whole does not count as strain. In isotropic solids the elasticity tensor c_{ijkl} takes a relatively simple form such that Hooke's law can be written as

$$\sigma_{ij} = K\delta_{ij}(\boldsymbol{\nabla}\cdot\boldsymbol{u}) + 2G\left[\varepsilon_{ij} - \frac{1}{3}(\boldsymbol{\nabla}\cdot\boldsymbol{u})\delta_{ij}\right] \tag{22}$$

where K is known as the *bulk modulus* and G the *shear modulus*. This particular form of Hooke's law comes from the assumption that the material responds differently to volume changing deformations (compression) and deformations that leave the volume fixed (shear). Note that the change in volume due to the deformation of the solid is equal to $\delta V = \int_{\partial V}\boldsymbol{u}\cdot d\boldsymbol{S} = \int_V \boldsymbol{\nabla}\cdot\boldsymbol{u}\, dV$, so that $\boldsymbol{\nabla}\cdot\boldsymbol{u} = \mathrm{Tr}[\boldsymbol{\epsilon}]$ measures the change in the volume of an infinitesimal element of the solid. Meanwhile, the traceless part of the strain tensor $\boldsymbol{\epsilon}$, $\varepsilon_{ij} - \frac{1}{3}(\boldsymbol{\nabla}\cdot\boldsymbol{u})\delta_{ij}$ is not associated with any volume change.

Applying (22) to our equation of motion (19) then gives

$$\rho\frac{\partial^2\boldsymbol{u}}{\partial t^2} = \left(K + \frac{1}{3}G\right)\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{u}) + G\boldsymbol{\nabla}^2\boldsymbol{u} \tag{23}$$

writing the displacement $\boldsymbol{u}$ as the sum of two parts $\boldsymbol{u} = \boldsymbol{u}_L + \boldsymbol{u}_T$ where $\boldsymbol{\nabla}\times\boldsymbol{u}_L = 0$ and $\boldsymbol{\nabla}\cdot\boldsymbol{u}_T = 0$ we find (23) reduces to two separate wave equations

$$\boxed{\begin{array}{lll} \frac{\rho}{K+\frac{4}{3}G}\frac{\partial^2\boldsymbol{u}_L}{\partial t^2} & = \boldsymbol{\nabla}^2\boldsymbol{u}_L & \text{Compression waves} \\ \frac{\rho}{G}\frac{\partial^2\boldsymbol{u}_T}{\partial t^2} & = \boldsymbol{\nabla}^2\boldsymbol{u}_T & \text{Shear waves} \end{array}} \tag{24}$$

Therefore, once again we see that the same wave equation as (6) and (16) also governs the motion of elastic waves propagating through a solid. One difference is that there are now two different kinds of waves that can propagate through the solid, longitudinal compression waves with velocity $c_L = \sqrt{(K + 4G/3)/\rho}$, and transverse shear waves with velocity $c_T = \sqrt{\rho/G}$. These two velocities can be quite different, for example aluminium has $c_L \sim 6000\,\mathrm{ms}^{-1}$ and $c_T \sim 3000\,\mathrm{ms}^{-1}$.

For a fixed frequency (e.g. $\boldsymbol{u}_L = \boldsymbol{u}_{L,\omega} e^{-i\omega t}$) these two equations reduce to two separate Helmholtz equations

$$\boxed{\begin{array}{rll} \boldsymbol{\nabla}^2 \boldsymbol{u}_L + \dfrac{\rho\omega^2}{K + \frac{4}{3}G} \boldsymbol{u}_L & = 0 & \text{Compression waves} \\ \\ \boldsymbol{\nabla}^2 \boldsymbol{u}_L + \dfrac{\rho\omega^2}{G} \boldsymbol{u}_L & = 0 & \text{Shear waves} \end{array}} \tag{25}$$

For more details on elastic waves see [5].

2.4 Matter waves:

Quantum mechanics is also a wave theory. We can obtain the equations of quantum mechanics from those of classical mechanics through applying the de–Broglie relations relating momentum and energy to wave–vector and frequency: $\boldsymbol{p} = \hbar \boldsymbol{k}$ and $E = \hbar\omega$.

In classical non–relativistic mehanics the energy and momentum are related as follows

$$\frac{\boldsymbol{p}^2}{2m} + V = E \tag{26}$$

which in terms of the wave–vector and frequency of the quantum wave is

$$\frac{\hbar^2 \boldsymbol{k}^2}{2m} + V = \hbar\omega$$

This equation applies for a wave $\psi = \exp[i(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t)]$ and in terms of this wave can be written as

$$\boxed{-\frac{\hbar^2}{2m} \boldsymbol{\nabla}^2 \psi + V\psi = i\hbar \frac{\partial \psi}{\partial t}} \tag{27}$$

which holds for any wave ψ and is known as the *time dependent Schroedinger equation*. For a fixed energy $\psi = \psi_E e^{-iEt/\hbar}$ (27) becomes

$$\boxed{\boldsymbol{\nabla}^2 \psi + \frac{2m}{\hbar^2}(E - V)\psi = 0} \tag{28}$$

which is the t*ime independent Schroedinger equation*. Note that the time independent Schroedinger equation is actually the Helmholtz equation that we've seen above in electromagnetism (7), pressure acoustics (17), and elasticity (25). Therefore, for a fixed energy E and with a spatially varying potential $V(\boldsymbol{x})$ the quantum wave–function actually behaves in exactly the same way as (for example) a sound wave in a fluid with a spatially varying compressibility $\beta_s(\boldsymbol{x})$.

While quantum mechanics is weird (most notably in entangled systems and the collapse of the wavefunction during measurement), there is a general modern approach to the subject that encourages us to give up on intuition and retreat into Hilbert space[2]. In many simple cases I don't think we need to do this: the wavefunction is a wave like any other.

2.5 Gravitational waves:

Gravitational waves are something predicted by general relativity, and it's a lot of work to introduce general relativity. There obviously isn't space here. But—just for fun—let's have a go at deriving the gravitational wave equation. You can always pretend this part didn't happen.

2. "*I feel that Einstein's intellectual superiority over Bohr, in this instance, was enormous; a vast gulf between the man who saw clearly what was needed, and the obscurantist.*" J. S. Bell

Without justifying it (for this see e.g. [2]) we'll start from Einstein's field equation, which is a tensor equation

$$R_{ij} = \frac{8\pi G}{c^4}\left(T_{ij} - \frac{1}{2}g_{ij}T\right) \tag{29}$$

where G is Newton's gravitational constant and c is the speed of light in vacuum. If you managed to get through the section on elasticity which was full of two index tensors, this equation shouldn't be so bad. The only fundamental mathematical difference is that the indices of these tensors range over four values rather than three, e.g. $i \in [0, 1, 2, 3]$ where the '0' index refers to the time direction, and '1', '2' and '3' are the three spatial directions.

The left hand side of (29) is written in terms of something called the *Ricci tensor* R_{ij}. This tells us how space–time is deformed, with R_{ij} quantifying aspects of the curvature of space–time. The Ricci tensor depends on the metric tensor g_{ij} which also appears on the right hand side of (29).

The metric tensor tells us how to measure distance and time relative to the coordinates x_i we use to cover space–time. For example, the elapsed time τ measured by a clock that moves between space–time points a and b is given by a formula containing the metric

$$\tau = \frac{1}{c}\sum_{i,j}\int_a^b \sqrt{g_{ij}dx_i dx_j} \tag{30}$$

In a flat space–time covered by Cartesian coordinates the metric tensor reduces to

$$g_{ij} = \eta_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and the elapsed time on our moving clock (30) reduces to the result we know from special relativity

$$\tau = \frac{1}{c}\int_a^b \sqrt{(dx_0)^2 - (d\boldsymbol{x})^2} = \int_a^b dt\sqrt{1 - \left(\frac{\boldsymbol{v}}{c}\right)^2}$$

where I used $dx_0 = c dt$, where c is the speed of light in vacuum.

Returning our attention to Einstein's field equation, on the right hand side of (29) we have the *energy–momentum tensor* T_{ij} and its trace T both of which tell us about the distribution and flow of energy and momentum in space–time. We can now see that in words Einstein's field equation says: *space–time is deformed at each point in proportion to the amount of energy and momentum present at that point*. The Ricci tensor R_{ij} is a non–linear function of the metric tensor g_{ij}, and to find a solution to Einstein's field equation we must find the metric tensor for a particular configuration of energy and momentum. Once we know the metric tensor then we can work out how bodies move through space–time.

The situation we face in solving (29) is not too different from that we face in solving the Euler equation of fluid dynamics (11). The Ricci tensor is a non–linear function of the metric tensor (the metric tensor is what we're after), just as the change in momentum in a fluid is a non–linear function of the flow velocity $\boldsymbol{v}$. Just as we did there, we'll derive our wave equation through considering small perturbations to the metric of space–time. In this case we'll take empty flat space–time as our background metric and consider a small perturbation $g_{ij}^{(1)}$ on top of this

$$g_{ij} = \eta_{ij} + g_{ij}^{(1)}$$

Einstein's field equation then becomes

$$R_{ij} - \eta_{ij}R = 0 \tag{31}$$

As in the previous cases covered above, to proceed we need a constitutive relation connecting the Ricci tensor with the metric tensor. For small deformations of space–time this consitutive relation is given by (see [2] for where this comes from)

$$R_{ij} = \frac{1}{2}\sum_{i,k} \eta_{ik}\left[\frac{\partial^2 g_{il}^{(1)}}{\partial x_k \partial x_j} + \frac{\partial^2 g_{jk}^{(1)}}{\partial x_l \partial x_i} - \frac{\partial^2 g_{jl}^{(1)}}{\partial x_k \partial x_i} - \frac{\partial^2 g_{ik}^{(1)}}{\partial x_l \partial x_j}\right] \tag{32}$$

Notice that the value of R_{ij} is unchanged if we make the following substitution

$$g_{jl}^{(1)} \to g_{jl}^{(1)} + \frac{\partial \alpha_j}{\partial x_l} + \frac{\partial \alpha_l}{\partial x_j}$$

where α_j is any vector. We use this freedom in the form of $g_{ij}^{(1)}$ to choose α_j such that

$$\sum_{i,k} \eta_{ik}\left[\frac{\partial^2 g_{il}^{(1)}}{\partial x_k \partial x_j} + \frac{\partial^2 g_{jk}^{(1)}}{\partial x_l \partial x_i} - \frac{\partial^2 g_{ik}^{(1)}}{\partial x_l \partial x_j}\right] = 0 \tag{33}$$

Substituting the constitutive relation (32) and the condition (33) into Einstein's field equation (31) then gives us

$$\boxed{\sum_{i,k} \eta_{ik}\frac{\partial^2 g_{jl}^{(1)}}{\partial x_k \partial x_i} = 0 \to \frac{1}{c^2}\frac{\partial^2 g_{jl}^{(1)}}{\partial t^2} = \boldsymbol{\nabla}^2 g_{jl}} \tag{34}$$

and we therefore see that perturbations to the metric tensor travel through space–time as waves moving at the speed of light in vacuum, obeying the exact same equation as for electromagnetic waves (6) and acoustic waves (16,24). The wave has a tensorial amplitude, which is another level of complication above the vector amplitudes encountered in electromagnetism and elasticity, but otherwise it's dependence on position and time will be very similar.

2.6 Summary:

The basic point I've tried to make here is that almost wherever you look in physics, there is the same wave equation governing the propagation of small disturbances. In the next few lectures we'll cover techniques for solving this wave equation. There are many other areas of physics I could have included, such as the spin waves that propagate through magnetic materials, or waves that propagate through plasmas.

It's worth bringing out the common features that led to the wave equations we obtained above. The first is the need for a constitutive relation, which is different for every material supporting the wave and determines simple things like the wave speed (or refractive index). Here the message is that through changing the material supporting the wave, you can change the wave equation. This is what metamaterials do. Even gravitation has something like a constitutive relation, linking the Ricci curvature to the metric tensor. It is only the matter wave that doesn't require any kind of constitutive relation.

Another common feature is that when the wave amplitude becomes too large, then in general the simple linear wave–equation no–longer applies and we have to deal with non–linear wave equations (except for matter waves, the amplitude of which is constrained by the normalization condition $\int |\psi|^2 dV = 1$). Non–linear waves are fascinating, but we won't cover them in this course. If you're interested see e.g. [6,14].

3 Reference:

3.1 Programming:

Mathematica is a superb piece of software, and Matlab is tolerable I suppose. But they don't really fit with the philosophy of this course (they are also commercial packages). Where I give example programs I shall use Python (a beginners guide can be found here). It is an easy to learn, freely available and open source programming language. I write and run Python programs in the Jupyter notebook, which is very much like the mathematica notebook, but is based in your web browser. Below I give a list of links to the software you'll need to set up python for solving physics problems.

Unless you use Linux (where you can get all the necessary software using something like '*apt-get*'), the Anaconda distribution is a good bundle of Python software which will include Jupyter and most of the libraries you need. You can find the Windows, Mac and Linux flavours of this distribution here:

Anaconda distribution: https://www.continuum.io/downloads

In case you want to install them yourself, or for some reason any of them are not included within the current Anaconda distribution, the following Python libraries (roughly in order of importance) are essential or useful to write physics programs:

Scipy (Linear algera, integration, special functions...): https://www.scipy.org/

Matplotlib (2D and 3D plotting): https://matplotlib.org/

Numpy (Basic numerical functions, required for everything else): http://www.numpy.org/

Mpmath (Very good special functions library, with much else): http://mpmath.org/

Mayavi (Advanced 3D plotting/visualization): http://docs.enthought.com/mayavi/mayavi/

Sympy (Will do symbolic manipulations for you): http://www.sympy.org/en/index.html

Tqdm (useful for timing slow functions): https://pypi.python.org/pypi/tqdm

Fenics (not included in Anaconda, but may be worth a look. This is a brilliant finite element library that allows you to set up and solve most differential equations yourself in the same way Comsol does): https://fenicsproject.org/

Gmsh (Fenics mesh generation is not up to much, but it can import from Gmsh): http://gmsh.info/

All the notebooks used in the course, plus some general examples to help you are available on my personal website.

3.2 Summary of important mathematical results:

I'm still writing this part, and will add to it as the course develops.

3.3 Vector identities:

The following identities of vector algebra and vector calculus are used in the course

$$\begin{aligned}
\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) &= \boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{c})-\boldsymbol{c}(\boldsymbol{a}\cdot\boldsymbol{b})\\
\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{V}) &= \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{V})-\boldsymbol{\nabla}^2\boldsymbol{V}\\
\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) &= \boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a})\\
&= \boldsymbol{c}\cdot(\boldsymbol{a}\times\boldsymbol{b})\\
\boldsymbol{\nabla}\cdot(\boldsymbol{V}\times\boldsymbol{W}) &= \boldsymbol{W}\cdot(\boldsymbol{\nabla}\times\boldsymbol{V})-\boldsymbol{V}\cdot(\boldsymbol{\nabla}\times\boldsymbol{W})
\end{aligned}$$

3.4 The divergence and circulation theorems:

There are two very important theorems of vector calculus that can be used to translate the integral of a vector field over a boundary of a volume or an area into the integral of the derivatives of this field inside that volume or area. These are

$$\int_{\partial V} \boldsymbol{v} \cdot d\boldsymbol{S} = \int_V \boldsymbol{\nabla} \cdot \boldsymbol{v}\, dV \qquad \text{Divergence theorem}$$

$$\int_{\partial A} \boldsymbol{v} \cdot d\boldsymbol{l} = \int_A \boldsymbol{\nabla} \times \boldsymbol{v} \cdot d\boldsymbol{S} \qquad \text{Circulation theorem}$$

where V indicates a three dimensional volume, and A a two dimensional area. One can think of these two formulae as generalizations of the one dimensional result of basic calculus that the integral of a derivative of a function equals the difference in the value of the function at the end points: $\int_a^b dx\, df(x)/dx = f(b) - f(a)$.

3.5 Cauchy's integral formula:

I'm not sure how many of you have been taught complex analysis, but it can be very useful for evaluating integrals. This section is supposed to just describe the very basics of a part of complex analysis. For more information see [7]. One of the most important results is Cauchy's integral formula, which we'll use a few times in the course. The formula is as follows:

$$\frac{1}{2\pi i}\oint_C \frac{f(z)}{z - z_0} dz = \begin{cases} f(z_0) & \text{Contour encloses } z_0 \\ 0 & \text{Otherwise} \end{cases} \tag{35}$$

We first need to explain what this is all about. In basic calculus we integrate both real and complex valued functions, but the integrals are always carried out along the real line. For example $\exp(ix)$ is a complex valued function, and its standard integral

$$\int_a^b \exp(ix) dx = -i[\exp(ib) - \exp(ia)] \tag{36}$$

is carried out along the real x axis. But broaden your mind, we don't need to restrict ourselves to integating along the real line! There is no obstacle to us integrating functions along *any trajectory* in the complex plane, letting x becoming $z = x + iy$. This enables us, for example to replace the real numbers a and b in (36) with complex numbers, and then the integral is taken over some path in the complex plane that joins a and b.

But having broadened our mind we might start to worry: there is more than one possible path that connects a and b. How do we know whether the value of the integral depends on the path we decide to choose? If it does then we need to start being careful about telling people which path our integral is taken over. Answering this question is what motivates Cauchy's theorem (35), which concerns the integral of a function over a closed contour C in the complex plane.

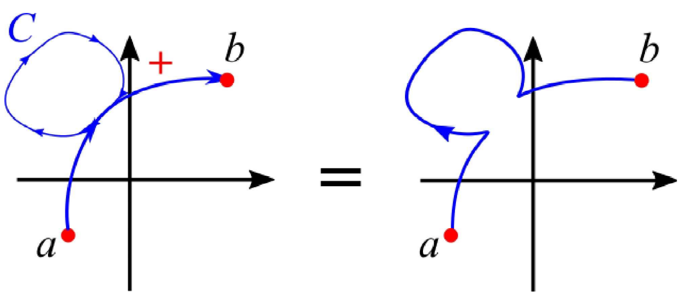

Figure 1. Adding a closed contour C to a path between two points can be used to deform the path. To show that the value of an integral between a and b is independent of the chosen path then we can equivalently show that the integral around a closed loop is zero.

For the value of an integral in the complex plane to be independent of the path joining a and b we need the integral to be zero when taken over a closed path. This is because we can always change the path joining a and b by adding a closed contour (see figure 1). If the contribution of a closed contour is zero then deforming the path of integration does not affect the value of the integral.

Let's consider the integral of a function $g(z)$ over a closed contour. We write $g(z) = g_x(z) + i g_y(z)$ (g_x and g_y being the real and imaginary parts of the function g), and $dz = dx + i dy$

$$\text{Integral} = \oint g(z)\, dz = \oint (g_x(z)\, dx - g_y(z) dy) + i \oint (g_x(z) dy + g_y(z) dx) \tag{37}$$

we can think of the two integrals on the right of (37) as line integrals of two dimensional vectors $\boldsymbol{w} = g_x \hat{\boldsymbol{x}} - g_y \hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{z}} \times \boldsymbol{w}$. Applying the circulation theorem of vector calculus $\oint_{\partial A} \boldsymbol{w} \cdot d\,\boldsymbol{l} = \int_A \boldsymbol{\nabla} \times \boldsymbol{w} \cdot d\,\boldsymbol{S}$ then gives us

$$\text{Integral} = -\int \left[\frac{\partial g_y}{\partial x} + \frac{\partial g_x}{\partial y} \right] dx\, dy + i \int \left[\frac{\partial g_x}{\partial x} - \frac{\partial g_y}{\partial y} \right] dx\, dy$$

Both real and imaginary parts of this integral are zero when

$$\begin{aligned} \frac{\partial g_y}{\partial x} &= -\frac{\partial g_x}{\partial y} \\ \frac{\partial g_x}{\partial x} &= \frac{\partial g_y}{\partial y} \end{aligned} \tag{38}$$

conditions which are known as the *Cauchy–Riemann* conditions. *When these conditions are satisfied then the value of the integral of $g(z)$ between two points is independent of the path joining the points.* Functions satisfying the Cauchy–Riemann conditions in a given portion of the complex plane are said to be *analytic* in that region of the complex plane. The Cauchy–Riemann conditions are not generally satisfied when $g(z)$ has singular points (poles), or when $g(z)$ is discontinuous (has branch cuts).

Exercise: Show that the Cauchy–Riemann conditions are satisfied for $g(z) = \exp(i z)$.

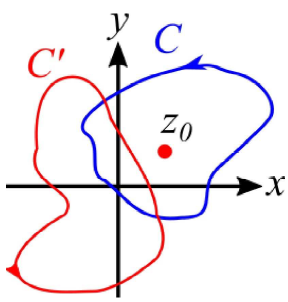

Figure 2. Suppose we have a function $f(z)$ that satisfies the Cauchy − Riemann conditions everywhere in some region of the complex plane. The function $f(z)\,/\,(z - z_0)$ thus satisfies the Cauchy-Riemann conditions everywhere except at z_0. Cauchy's theorem says that an integral of this function along a contour enclosing z_0 (here shown as C) will give the value $2\pi i f(z_0)$, while an integral along a contour that does not enclose z_0 (C') will give zero.

So we have established when the value of a contour integral is independent of the path chosen to join the end points. From looking at (35) we can see that it is a special case where the Cauchy–Riemann conditions are *not* satisfied.

To explore this further, consider the function $g(z) = 1/z = (x - iy)/(x^2 + y^2)$, which is singular at $z = 0$. Plugging this function into the Cauchy Riemann conditions we see that away from the point $z = 0$ we have

$$\begin{aligned} \frac{\partial g_x}{\partial x} &= \frac{\partial}{\partial x}\frac{x}{x^2+y^2} &&= \frac{y^2-x^2}{(x^2+y^2)^2} \\ \frac{\partial g_x}{\partial y} &= \frac{\partial}{\partial y}\frac{x}{x^2+y^2} &&= -\frac{2xy}{(x^2+y^2)^2} \\ \frac{\partial g_y}{\partial x} &= -\frac{\partial}{\partial x}\frac{y}{x^2+y^2} &&= \frac{2xy}{(x^2+y^2)^2} \\ \frac{\partial g_y}{\partial y} &= -\frac{\partial}{\partial y}\frac{y}{x^2+y^2} &&= \frac{y^2-x^2}{(x^2+y^2)^2} \end{aligned} \tag{39}$$

and after a bit of staring we can see that the Cauchy–Riemann conditions (38) are satisfied away from $z = 0$. However, at the single point $z = 0$ (where the above derivatives all become infinite) they are not satisfied. We can see this through taking a small closed circular contour integral around the point $z = 0$

$$\oint \frac{dz}{z} = \int_0^{2\pi} \frac{ire^{i\theta}d\theta}{re^{i\theta}} = 2\pi i$$

where I wrote the complex number z in polar coordinates $z = re^{i\theta}$. The fact that this integral is non–zero means that the Cauchy–Riemann conditions must not be satisfied at the single point $z = 0$ (equation (39) shows that they are satisfied everywhere else). This leads us to the following result

$$\frac{1}{2\pi i}\oint_C \frac{dz}{z - z_0} = \begin{cases} 1 & \text{Contour encloses } z_0 \\ 0 & \text{Otherwise} \end{cases} \tag{40}$$

which is almost Cauchy's theorem (35). The last step just requires us to recognise that if we multiply the integrand of (40) by a function $f(z)$ that itself satisfies the Cauchy–Riemann conditions within C then the contour can be deformed to an infinitesimal circle surrounding z_0, on which $f(z)$ tends towards the constant value $f(z_0)$. Therefore (see figure 2)

$$\frac{1}{2\pi i}\oint_C \frac{f(z)dz}{z - z_0} = \begin{cases} f(z_0) & \text{Contour encloses } z_0 \\ 0 & \text{Otherwise} \end{cases}$$

where the closed contour C is traversed in an anticlockwise sense. There is a Jupyter notebook on my website that you can play with that numerically demonstrates this theorem.

3.6 Special functions, identities and expansions:

In two dimensions a plane wave can be expanded as a series of cylindrical waves with different angular momentum l.

$$e^{ikx} = e^{ikr\cos(\theta)} = \sum_{l=-\infty}^{\infty} i^l J_l(kr) e^{il\theta} \tag{41}$$

The functions $J_l(kr)$ are Bessel functions of the first kind and are the solutions to the following differential equation

$$\left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{l^2}{r^2} + k^2\right] J_l(kr) = 0 \tag{42}$$

We shall use formula (41) several times. A simple proof can be found in [7]. There is also a Jupyter notebook on my website where you can play around with this expansion.

As with every second order differential equation, Bessel's equation (42) has two solutions, the second of which is denoted $Y_l(kr)$. You can think of J_l and Y_l as analogues of the sine and cosine functions that you get in the one dimensional wave equation. Just as we can build left and right travelling waves out of sines and cosines $\exp(\pm ikx) = \cos(kx) \pm i\sin(kx)$, we can build incoming and outgoing cylindrical waves out of the two Bessel functions J_l and Y_l. These are the Hankel functions,

$$H_l^{(1)}(z) = J_l(z) + iY_l(z)$$

$$H_l^{(2)}(z) = J_l(z) - iY_l(z)$$

The Hankel function of the first kind $H_l^{(1)}$ is an outgoing travelling cylindrical wave, and the Hankel function of the second kind $H_l^{(2)}$ is an incoming travelling cylindrical wave.

Reference [9] (section 10.11) gives an identity for zeroth order Bessel functions that we'll make use of in the course:

$$J_0(-z) = J_0(z)$$

$$Y_0(-z) = Y_0(z) + 2i\,J_0(z)$$

Applying this to our formulae for the two Hankel functions we see that

$$H_0^{(1)}(-z) = J_0(-z) + iY_0(-z) = -J_0(z) + iY_0(z) = -H_0^{(2)}(z). \tag{43}$$

3.7 Suggested books:

We won't be following any textbook very closely, but here are some books where you'll find some extra information on the topics covered.

[1] W. C. Elmore and M. A. Heald, "*Physics of Waves*", Dover, New York (1985).

[2] L. D. Landau and E. M. Lifshitz, "*Classical Theory of Fields*", Butterworth–Heinemann, Oxford (2000).

[3] L. D. Landau and E. M. Lifshitz, "*Quantum Mechanics*", Butterworth–Heinemann, Oxford (2003).

[4] L. D. Landau and E. M. Lifshitz, "*Fluid Mechanics*", Butterworth–Heinemann, Oxford (2009).

[5] L. D. Landau and E. M. Liftshitz, "*Theory of Elasticity*", Butterworth–Heinemann, Oxford (1999).

[6] L. D. Landau and E. M. Lifshitz, "*Electrodynamics of Continuous Media*", Butterworth–Heinemann, Oxford (2004).

[7] G. B. Arfken, H. J. Weber and F. E. Harris, "*Mathematical Methods for Physicists*", Academic Press, Oxford (2013).

[8] P. M. Morse and H. Feshbach, "*Methods of Theoretical Physics (Parts 1 and 2)*", McGraw–Hill (1959). [This may be difficult to find in print]

[9] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders, eds. *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.14 of 2016-12-21. [a reference on special functions, I wouldn't suggest a 'cover-to-cover' reading]

[10] J. W. S. Rayleigh, "*The Theory of Sound (Volumes 1 and 2)*", Dover, New York (1998).

[11] J. Billingham and A. C. King , "*Wave Motion*", Cambridge University Press (2000).

[12] J. Lighthill, "*Waves in Fluids*", Cambridge University Press (2005).

[13] J. D. Jackson, "*Classical Electrodynamics*", Wiley (2007).

[14] P. G. Drazin and R. S. Johnson, "*Solitons: an introduction*", Cambridge University Press (1996).

[15] L. Brillouin, "*Wave Propagation and Group Velocity*", Academic Press (1960).

[16] H. C. van der Hulst, "*Light scattering by small particles*" Dover, New York (1981).

[17] M. Tinkham, "*Group theory and quantum mechanics*" McGraw–Hill (1964).

[18] J. Heading, "*An introduction to phase integral methods*", Dover (2013).

[19] U. Leonhardt and T. G. Philbin, "*Geometry and Light: The Science of Invisibility*", Dover (2010).

[20] A. V. Woolf, "*The Waves*", Wordsworth Classics (2000).