# LECTURES ON THE WAVE EQUATION

ANDREW HASSELL

ABSTRACT. These are notes from five lectures on the wave equation given at the 2005 AMSI Summer School held at ANU.

## 1. Introduction, Distributions, Fourier transform

The wave equation is the PDE

$$\frac{\partial^2 u}{\partial t^2} = \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2} \tag{1}$$

for a function $u = u(x,t)$ of $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. For brevity this will usually be written

$$u_{tt} = \Delta u.$$

This is usually supplemented by initial conditions, specified at $t = 0$. Since it is a second order equation in time the initial value and the initial time derivative are usually specified:

$$u(x,0) = f(x), \quad u_t(x,0) = g(x) \tag{2}$$

where $f$ and $g$ are given functions in $\mathbb{R}^n$.

Note that the wave equation is definitely not an elliptic PDE, since the sign of the second $t$-derivatives is opposite to the sign of the space derivatives. We shall soon see that the solutions to the wave equation have a very different character to those of elliptic equations.

Let $h(s)$ be a $C^2$ function of one variable, and let $\omega$ be a unit vector in $\mathbb{R}^n$. Then $u(x,t) = h(x \cdot \omega - t)$ is a solution to the wave equation.

*Exercise* 1. Prove this.

This is a 'travelling wave' solution to the equation, and is the reason for the name 'wave equation'. Does $h$ really have to be $C^2$?

*Exercise* 2. Show that if $u(x,t)$ solves the wave equation, then $u(x,-t)$ does too. That is, the wave equation has the property of time reversibility.

We can also consider the wave equation on domains other than $\mathbb{R}^n$ — for example, domains in $\mathbb{R}^n$ (in which case we need to supplement the equation with boundary conditions) or manifolds (in which case we need to choose a metric on the manifold and replace the Laplace operator on the right hand side with the Laplace-Beltrami operator determined by the metric). However, we will avoid dealing with manifolds in these lectures, with the single exception of the flat torus. The flat torus of dimension $n$ and length $L$ is $\mathbb{R}^n$ quotiented by the group of translations along a lattice of side length $L$. Functions on the torus can be thought of as functions on $\mathbb{R}^n$ which are invariant under all such translations.

However, for the first two lectures we shall stick to the wave equation on $\mathbb{R}^n$. In that case we can use the Fourier Transform to help solve the wave equation.

1.1. **Fourier Transform.** The wave equation is a constant coefficient equation, and as such the Fourier Transform is the appropriate tool to use to solve it. I will not assume that you already know about the Fourier Transform, but on the other hand time is short, so this is an extremely brief and incomplete introduction to it.

The Fourier transform on $\mathbb{R}^n$ takes (sufficiently nice) functions on $\mathbb{R}^n$ and maps them to functions on $\mathbb{R}^n$. It is defined by

$$\mathcal{F}f(\xi) \equiv \hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, dx. \tag{3}$$

Here, 'sufficiently nice' means that the integral converges, i.e. $f$ decays sufficiently fast at infinity. It can be thought of as a continuum limit of Fourier series on an interval $[-L, L]$ as $L \to \infty$. The Fourier coefficients here would be defined by (restricting to one dimension for simplicity)

$$f_n = \int_{-L}^{L} e^{-in\pi x/L} f(x)\, dx, \quad f \in L^2([-L, L]). \tag{4}$$

Then Bessel's identity is

$$\frac{1}{2L} \sum_n |f_n|^2 = \int_{-L}^{L} |f(x)|^2\, dx.$$

*Exercise* 3. For every $N$, show that

$$\frac{1}{2L} \sum_{n=-L}^{L} |f_n|^2 \le \int_{-L}^{L} |f(x)|^2\, dx.$$

Now let $\xi = n\pi/L$. In the continuum limit, $L \to \infty$, at least heuristically $\xi$ becomes a continuous variable and (4) tends to the expression (3). Bessel's identity also has a continuum limit, since

$$\frac{1}{2L} \sum_n |f_n|^2 \to \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2\, d\xi.$$

The basic property of the Fourier Transform which makes it useful for analyzing constant coefficient PDE is that it transforms derivative operators (in $x$) into polynomials (in $\xi$). Let $D_j$ stand for the operator $-i\partial/\partial x_j$. Then integrating by parts in the integral (3) shows that

$$\mathcal{F}(D_j f) = \xi_j \mathcal{F} f.$$

Also, the 'opposite' is true (with a change of sign). Consider the Fourier transform of the function $x_j f$. By writing $x_j e^{-ix\cdot\xi} = i\partial/\partial\xi_j e^{-ix\cdot\xi}$, we see that

$$\mathcal{F}(x_j f) = -D_j(\mathcal{F}f).$$

More generally, let $\alpha = (\alpha_1, \dots, \alpha_n)$, $\alpha_i$ nonnegative integers, be a multi-index and let $D^\alpha$ denote $D_{x_1}^{\alpha_1} \dots D_{x_n}^{\alpha_n}$ and $\xi^\alpha$ denote $\xi_1^{\alpha_1} \dots \xi_n^{\alpha_n}$. Then we have (assuming that $f$ is small enough at infinity so that the integrals are all convergent, and hence integrations by parts may be performed)

$$\mathcal{F}(D^\alpha f) = \xi^\alpha \mathcal{F}f, \quad \mathcal{F}(x^\alpha f) = (-1)^{|\alpha|} D^\alpha \mathcal{F}f. \tag{5}$$

Now assume that we have a solution, $u(x,t)$, to the wave equation on all of $\mathbb{R}^n \times \mathbb{R}$. Then the Fourier transform satisfies an algebraic equation

$$\tau^2 \hat{u}(\xi,\tau) = -|\xi|^2 \hat{u}(\xi,\tau), \tag{6}$$

which is in principle much easier to solve. (Although, it looks like a strange sort of equation, since it implies that $\hat{u}$ vanishes except where $\tau^2 = |\xi|^2$, which is a set of measure zero! What is going on here will be explained shortly, in terms of distributions.) Actually, it is more useful to take the Fourier transform in the $x$ variables only. That is, we think of $u$ as a function of $t$ with values in functions of $x$, and then the Fourier transform is a function of $t$ with values in functions of $\xi$. If we do this then we end up with the equation

$$\hat{u}_{tt}(\xi,t) = |\xi|^2 \hat{u}(\xi,t), \tag{7}$$

which is an *ordinary* differential equation in $t$ for each value of $\xi$, which is simple to solve. This approach is more suited to dealing with initial conditions.

1.2. **Test functions and Distributions.** The analysis of the wave equation via the Fourier transform leads inexorably to distribution theory. We've had two hints of it already in this lecture. Let's approach this first by asking the question: what is a class of functions on which the Fourier transform acts nicely? If we want the integral to be convergent, then we should ask that $f$ is in $L^1(\mathbb{R}^n)$. However, it is very difficult to characterize the *range* of $\mathcal{F}$ on $L^1(\mathbb{R}^n)$, so we shall deal with a much nicer (i.e. more restrictive) class of functions, namely *Schwartz functions* or *functions of rapid decrease*. Let $\mathcal{S}(\mathbb{R}^n)$ denote the class of Schwartz functions on $\mathbb{R}^n$.

By definition, a Schwartz function on $\mathbb{R}^n$ is one such that for any two multi-indices $\alpha$, $\beta$,

$$\sup_{x\in\mathbb{R}^n} |x^\beta D^\alpha f(x)| = C_{\alpha,\beta} < \infty.$$

Then $f$ decreases rapidly at infinity in the sense that it is smaller than the reciprocal of any polynomial, and this is true not just for $f$ but all its derivatives as well. Then

$$f(x) \leq C(1+|x|^2)^{-n} \implies f \in L^1(\mathbb{R}^n),$$

and this is also true for all derivatives of $f$. So the Fourier Transform is well-defined on Schwartz functions and (5) holds for all Schwartz $f$.

*Exercise* 4. Show that (5) implies that the Fourier transform of a Schwartz function is also Schwartz[1].

Now we come to one of the main properties of the Fourier transform — the inversion formula. Let us define the map $\mathcal{G}$, which is almost the same as $\mathcal{F}$ but with two slight changes:

$$\mathcal{G}f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} f(\xi)\, d\xi. \tag{8}$$

The differences are the factor of $(2\pi)^{-n}$ and the change of sign in the phase of the exponential.

**Theorem 1.** *The map $\mathcal{G}$ on $\mathcal{S}(\mathbb{R}^n)$ is a two-sided inverse to $\mathcal{F}$.*

*Proof.* We first compute the Fourier transform of the Schwartz function $e^{-ax^2/2}$ (in one dimension). This function satisfies the differential equation

$$\frac{\partial u}{\partial x} = -axu.$$

Since, by (5), the Fourier transform swaps the operations of multiplication by polynomials and differentiation, the Fourier transform of $u$ satisfies the differential equation

$$a\frac{\partial \hat{u}}{\partial \xi} = -\xi \hat{u}.$$

This implies that $\hat{u}$ is a multiple of $e^{-\xi^2/2a}$. To see which multiple, we compute

$$\hat{u}(0) = \int_{-\infty}^{\infty} e^{-ax^2/2}\, dx = \sqrt{\frac{2\pi}{a}}.$$

(Here we use the standard result $\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$.) Hence

$$\mathcal{F}(e^{-ax^2/2}) = \sqrt{\frac{2\pi}{a}} e^{-\xi^2/2a}.$$

Thus the more peaked the original Gaussian $u$ is, the more spread out the Fourier transform is, and vice versa. We can compute the Fourier transforms of Gaussians of several variables in the same way (one variable at a time), and we get in $\mathbb{R}^n$

$$\mathcal{F}(e^{-a|x|^2/2}) = \big(\frac{2\pi}{a}\big)^{n/2} e^{-|\xi|^2/2a}. \tag{9}$$

Now to treat the general case, take a Schwartz function $\phi(x)$, $x \in \mathbb{R}^n$. The composition $\mathcal{G} \circ \mathcal{F}\phi$ is given by the integral

$$(2\pi)^{-n} \int_{\mathbb{R}^n} d\xi e^{ix\cdot\xi} \Big[ \int_{\mathbb{R}^n} e^{-iy\cdot\xi} \phi(y)\, dy \Big].$$

[1] There is also a *topology* on $\mathcal{S}(\mathbb{R}^n)$ which makes $\mathcal{F}$ a continuous bijection on $\mathcal{S}(\mathbb{R}^n)$. However, we shall not discuss the topology due to time constraints.

This cannot be regarded as an integral over $\mathbb{R}^{2n}$ because the integrand is not convergent there; there is no decay in the $\xi$ directions at all, only in the $y$ direction, due to rapid decrease of the function $\phi$. To remedy this we introduce artificially a function decaying as $|\xi| \to \infty$:

$$\lim_{\epsilon \to 0} (2\pi)^{-n} \int_{\mathbb{R}^n} d\xi e^{ix\cdot\xi} \Big[ e^{-\epsilon|\xi|^2/2} \int_{\mathbb{R}^n} e^{-iy\cdot\xi} \phi(y)\, dy \Big].$$

The limit as $\epsilon \to 0$ gives the original integral we want. (This ought to be true, since the pointwise limit of $e^{-\epsilon|\xi|^2/2}$ as $\epsilon \to 0$ is 1. Rigorously: since the Fourier transform of $\phi$ is a Schwartz function of $\xi$, and therefore $L^1$ in $\xi$, the result follows by the dominated convergence theorem.)

Now the integrand is an $L^1$ function in $\mathbb{R}^{2n}$. This entitles us to apply Fubini's theorem and switch the order of integration; i.e. we can do the $\xi$ integral first. This produces for us the Fourier transform of the Gaussian, which fortunately we just computed:

$$\lim_{\epsilon \to 0} (2\pi)^{-n} \big( \frac{2\pi}{\epsilon} \big)^{n/2} \int_{\mathbb{R}^n} e^{-|x-y|^2/2\epsilon} \phi(y)\, dy.$$

Now $(2\pi\epsilon)^{-n/2} e^{-|x-y|^2/2\epsilon}$, regarded as a function of $y$ with $x$ held fixed, is a positive function with integral 1, which is concentrating at the point $x$. This gives us a 'weighted average' of the function $\phi$ which is concentrating at the point $x$ as $\epsilon \to 0$, so in the limit we get just $\phi(x)$.

*Exercise* 5. Show this carefully, by changing variable to $z = (x - y)/\epsilon$ and using the dominated convergence theorem again.

Thus, we have shown that

$$\mathcal{G}\mathcal{F}\phi(x) = \phi(x),$$

which proves that $\mathcal{G} \circ \mathcal{F}$ is the identity on $\mathcal{S}(\mathbb{R}^n)$. An essentially identical argument proves that $\mathcal{F} \circ \mathcal{G}$ is the identity on $\mathcal{S}(\mathbb{R}^n)$, so this proves that $\mathcal{F}$ is a bijection (one-to-one and onto) on $\mathcal{S}(\mathbb{R}^n)$ with inverse $\mathcal{G}$. □

A similar argument shows that

**Theorem 2.** *The Fourier transform preserves the $L^2$ norm of Schwartz functions in $\mathbb{R}^n$ (up to a factor $(2\pi)^{n/2}$):*

$$\|\mathcal{F}f\|_{L^2} = (2\pi)^{n/2} \|f\|_{L^2}, \quad f \in \mathcal{S}(\mathbb{R}^n).$$

*Proof.* If we write out the square of the $L^2$ norm of $\hat{f}$ longhand we get

$$\int d\xi \Big[ \int e^{-iy\cdot\xi} f(y)\, dy \Big] \Big[ \int e^{ix\cdot\xi} \overline{f(x)}\, dx \Big].$$

Similar to before, this integrand is not $L^1$ on $\mathbb{R}^{3n}$ because we have no decay at infinity in the $\xi$ variable. So we introduce a decaying Gaussian factor as before and take a limit:

$$\lim_{\epsilon \to 0} \int d\xi e^{-\epsilon|\xi|^2/2} \Big[ \int e^{-iy\cdot\xi} f(y)\, dy \Big] \Big[ \int e^{ix\cdot\xi} \overline{f(x)}\, dx \Big].$$

Now the integrand is $L^1$ on $\mathbb{R}^{3n}$ and we may invoke Fubini's theorem and do the $\xi$ integral first. We get

$$\lim_{\epsilon\to 0} \big(\frac{2\pi}{\epsilon}\big)^{n/2} \int\int e^{-|x-y|^2/2\epsilon} f(y)\overline{f(x)}\, dx\, dy.$$

Using similar reasoning as above, the limit is equal to

$$(2\pi)^n \int |f(x)|^2\, dx.$$

□

We now introduce *distributions*, or more precisely *tempered distributions.* The point of distribution theory is to be able to rigorously define and analyse objects like the delta function. The delta function in one dimension, centred at zero is, intuitively at least, given by a limit of functions such as

$$\phi_\epsilon(x) = \frac{1}{\epsilon}\phi\big(\frac{x}{\epsilon}\big) \text{ as } \epsilon \to 0.$$

Here $\phi$ is a Schwartz function with integral 1. For example, $\phi(x)$ could be $(2\pi)^{-1/2}e^{-x^2/2}$ — a function that we have used above. Notice that $\phi_\epsilon$ has the property that the integral is 1 for every value of $\epsilon$, but $\phi_\epsilon$ is concentrating at the origin. The limit certainly does not exist in a pointwise sense, but it does in a weaker sense. Namely, we can integrate against any bounded continuous function $g$, and we find that

$$\lim_{\epsilon\to 0} \int_{-\infty}^{\infty} \phi_\epsilon(x) g(x)\, dx = g(0).$$

(To show this, let $y = x/\epsilon$, change variables and use DCT.) Motivated by this we *define* distributions to be abstract linear functionals defined on Schwartz space, i.e. linear maps from $\mathcal{S}(\mathbb{R}^n)$ to the complex numbers[2]. The delta 'function' is then the linear functional

$$\phi \mapsto \phi(0).$$

We will write the action of a distribution $u$ on a Schwartz function $\phi$ by $\langle u, \phi\rangle$. Thus $\langle \delta, \phi\rangle = \phi(0)$. Every function $f$ on $\mathbb{R}^n$ with at most polynomial growth is a distribution, with the action being

$$\langle f, \phi\rangle = \int_{\mathbb{R}^n} f(x)\phi(x)\, dx.$$

It is a simple and worthwhile exercise to show that there is no such function whose action agrees with the delta 'function'. Thus the delta function is not a function at all, and therefore the class of distributions is larger than that of functions (they are sometimes called generalized functions).

[2]To be precise, they are continuous linear functionals, but since I ducked the question of the topology on Schwartz space earlier, I can't say here what continuity means, so we will just ignore this issue for now. You can read about the topology on $\mathcal{S}(\mathbb{R}^n)$ in any standard book on distributions, or most PDE books.

Using this definition we can define various operations on distributions. For example, the operations of scalar multiplication and pointwise addition have obvious extensions to distributions. More interestingly we can define the operations of differentiation and the Fourier transform on distributions. These are defined 'by duality'. First consider differentiation. Suppose that $f$ is a $C^1$ function of polynomial growth. Then

$$\int_{\mathbb{R}^n} (\partial_{x_j} f)(x)\phi(x)\,dx = -\int_{\mathbb{R}^n} f(x)(\partial_{x_j}\phi)(x)\,dx.$$

Hence, for functions, we have the formula

$$\langle \partial_j f, \phi \rangle = \langle f, \partial_j \phi \rangle.$$

Thus we *define*, for any distribution, the partial derivative $\partial_j u$ in the $x_j$ direction by

$$\langle \partial_j u, \phi \rangle = \langle u, \partial_j \phi \rangle. \tag{10}$$

This definition then extends the partial derivative operators from functions to distributions.

*Exercise* 6. Show that partial derivatives commute on distributions.

*Exercise* 7. My calculus text (Adams) gives the following example of a function for which $\partial_x \partial_y f \neq \partial_y \partial_x f$ at the origin:

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}.$$

This function may be regarded as a distribution. Explain the apparent paradox with the previous exercise.

A similar definition works for the Fourier transform. If $f$ is a Schwartz function, then

$$\langle \mathcal{F} f, \phi \rangle = \int \int f(\xi) e^{-ix\cdot\xi} \phi(x)\,dx\,d\xi = \langle f, \mathcal{F}\phi \rangle.$$

(Here we have used the symmetry of the function $e^{-ix\cdot\xi}$ in $x$ and $\xi$.) Thus we *define*, for any distribution, the Fourier transform $\mathcal{F}u$ by

$$\langle \mathcal{F} u, \phi \rangle = \langle u, \mathcal{F}\phi \rangle. \tag{11}$$

As an example of this, let's compute the derivative and the Fourier transform of the delta function (in one dimension). The derivative is given by

$$\langle \delta', \phi \rangle = -\langle \delta, \phi' \rangle = -\phi'(0).$$

Thus $\delta'$ is the distribution that takes a Schwartz function to minus its derivative at zero. The Fourier transform is defined by

$$\langle \mathcal{F}\delta, \phi \rangle = \langle \delta, \hat{\phi} \rangle = \hat{\phi}(0) = \int_{-\infty}^{\infty} \phi(x)\,dx = \langle 1, \phi \rangle.$$

So the Fourier transform of the delta function is the function 1.

*Exercise* 8. What is the Fourier transform of the function $x_j$ (regarded as a distribution)?

One final property of the Fourier transform we need concerns convolutions. The convolution of two Schwartz functions $\phi$ and $\chi$ is the function $\psi = \phi * \chi$ given by

$$\psi(x) = \int_{\mathbb{R}^n} \phi(x-y)\chi(y)\,dy.$$

It turns out that if $\phi$ and $\chi$ are Schwartz then so is $\phi * \chi$. The Fourier transform of a convolution is the product of the convolutions:

$$\mathcal{F}(\phi * \chi) = \hat{\phi} \cdot \hat{\chi}.$$

This is a straightforward calculation: Write the Fourier transform of the convolution

$$\int dx e^{-ix\cdot\xi} \int_{\mathbb{R}^n} \phi(x-y)\chi(y)\,dy.$$

Now write $e^{-ix\cdot\xi} = e^{-i(x-y)\cdot\xi}e^{-iy\cdot\xi}$ and change variables from $y$ to $z = x-y$. Applying the inverse Fourier transform, we deduce that

$$\phi * \chi = \mathcal{G}\big(\hat{\phi} \cdot \hat{\chi}\big). \tag{12}$$

Similarly, the Fourier transform of a product is the convolution of the Fourier transforms, times $(2\pi)^n$. We can define the convolution of a distribution with a Schwartz function by duality.

*Exercise* 9. What should the definition be?

## 2. Wave equation

**2.1. Wave equation in $1+1$ dimensions.** To warm up, we will analyse the solution to the wave equation in one space dimension,

$$u_{tt} = u_{xx}$$

together with the initial conditions

$$u(x,0) = f(x), \quad u_t(x,0) = g(x).$$

We will find an explicit formula for the solution in terms of $f$ and $g$. To do this, we make use of a factorization of the wave equation (this works only in one space dimension). Let us write $v = u_t - u_x$. Then the wave equation is equivalent to $v_x + v_t = 0$, since $v_x + v_t = (u_{tx} - u_{xx}) + (u_{tt} - u_{tx}) = u_{tt} - u_{xx}$.

The equation $v_x + v_t = 0$ for $v$ is a constant-coefficient *transport equation* and is easy to solve. Geometrically the equation means that the directional derivatives of $v$ vanish along the direction $e_x + e_t$ in the $(x,t)$-plane. Integrating, we find that $v$ is constant along lines parallel to the vector $e_x + e_t$. This means that $v(x,t)$ is equal to $v(x-t,0)$ which is given in terms of the initial conditions by $g(x-t) - f'(x-t)$. So we have found that

$$v(x,t) = g(x-t) - f'(x-t). \tag{13}$$

This is all very well but we want to find $u(x,t)$ rather than $v(x,t)$. So we need to solve the equation $u_t - u_x = v$. This equation says that the directional derivative of $u$ in the direction $e_t - e_x$ is equal to $v$. So, using the fundamental theorem of calculus along lines $x+t =$ constant, we find that

$$u(x,t) = u(x+t,0) + \int_0^t v(x+t-s,s)\,ds.$$

Now substitute in (13) and we find that

$$u(x,t) = f(x+t) + \int_0^t \big(g(x+t-2s) - f'(x+t-2s)\big)\,ds.$$

A more symmetrical expression for $u$ is given by

$$u(x,t) = \frac{f(x+t)+f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(s)\,ds. \tag{14}$$

*Exercise* 10. Using a similar method, write down an explicit solution in terms of $f, g, h$ of the inhomogeneous wave equation in $1+1$ dimensions:

$$u_{tt} - u_{xx} = h(x,t), \quad u(x,0) = f(x), \quad u_t(x,0) = g(x).$$

From this simple formula we can deduce several things. Firstly, if we assume that $f$ is $C^2$ and $g$ is $C^1$, then the solution given by (14) is $C^2$. Moreover the $C^2$ norm of the solution is no bigger than the sum of the $C^2$ norms of $f$ and $g$. We say that the PDE is 'well-posed' in $C^2$. Furthermore, the solution at $(x,t)$ only depends on the initial data in the interval $[x-t, x+t]$, so there is a 'finite speed of propagation' property — the initial data at a point $y$ can only affect the solution at $(x,t)$ if $|x-y| \le t$. Finally, there is no analogue of elliptic regularity for the wave equation; a solution to the wave equation does not have to be $C^\infty$, since we can certainly take $f$ to be $C^2$ but not $C^3$ and this gives a perfectly good solution to the wave equation. All these properties distinguish the wave equation from second order elliptic equations; the Cauchy problem (i.e. the initial value problem where the solution and its first derivative are specified at time $t=0$) is not well-posed for an elliptic equation, and there is no finite speed of propagation.

2.2. **Solution via Fourier transform.** For space dimensions $n \ge 2$ there is no analogue of the above argument. We shall instead move to the Fourier transform.

Let $\hat{u}(\xi,t)$ denote the Fourier transform of $u$ in the space variables only. Then the initial value problem is transformed into

$$\hat{u}_{tt} = -|\xi|^2\hat{u}, \quad \hat{u}(\xi,0) = \hat{f}(\xi), \quad \hat{u}_t(\xi,0) = \hat{g}(\xi).$$

For each fixed $\xi$ this is an ODE in $t$ which is easily solved. The solution is

$$\hat{u}(\xi,t) = (\cos|\xi|t)\hat{f}(\xi) + \frac{\sin|\xi|t}{|\xi|}\hat{g}(\xi).$$

Thus the solution of the original PDE is

$$u(x,t) = \mathcal{G}\Big( (\cos|\xi|t)\hat{f}(\xi) + \frac{\sin|\xi|t}{|\xi|}\hat{g}(\xi) \Big) = G_1^{(n)}(t) * f + G_2^{(n)}(t) * g, \tag{15}$$

where

$$G_1^{(n)}(t) = \mathcal{G}(\cos|\xi|t), \quad G_2^{(n)}(t) = \mathcal{G}\big(\frac{\sin|\xi|t}{|\xi|}\big).$$

Notice that $\cos|\xi|t$ does not decay as $\xi \to \infty$, and therefore the inverse Fourier transform is not defined as a convergent integral. Instead, we must interpret it as a distribution. The Fourier transform of it is then another distribution. Let us consider this problem in dimensions 1, 2 and 3 only. First, dimension 1. In this case, we have already solved the problem! We need only interpret our solution (14) in terms of convolutions. Note that

$$\frac{f(x+t)+f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(s)\,ds = G_1^{(1)}(t) * f + G_2^{(1)}(t) * g,$$

with

$$G_1^{(1)}(t) = \frac{\delta(x-t)+\delta(x+t)}{2}, \; G_2^{(1)}(t) = \frac{1}{2} I_{[-t,t]} = \frac{1}{2}\big(H(x+t) - H(x-t)\big).$$

Now for dimension 3, which is the next easiest. We need to find the inverse Fourier transform of $\sin(|\xi|t)/|\xi|$, which is $E_2$. (Once we have done that, it is immediate to find $E_1(t)$ since it is just the time derivative of $E_2(t)$.) Let's try to guess what the answer is based on some heuristics. First, since $\sin(|\xi|t)/|\xi|$ is spherically symmetric, it makes sense to look for a spherically symmetric function.

*Exercise* 11. Show that the Fourier transform of a spherically symmetric function is also spherically symmetric. What about the Fourier transform of a spherically symmetric distribution? (First define spherical symmetry for a distribution!)

In one dimension, $\delta(x-t)$ is a solution to the wave equation. In 3 dimensions, a spherically symmetric solution $k(r,t)$ to the wave equation satisfies the equation

$$\Big(\partial_r^2 + \frac{2}{r}\partial_r - \partial_t^2\Big) k(r,t) = 0.$$

This is different from the equation in one dimension, where the second term is absent. However, we can use a trick: the equation above is equivalent to

$$r^{-1}\Big(\partial_r^2 - \partial_t^2\Big)(rk(r,t)) = 0.$$

The operator $\partial_r^2 - \partial_t^2$ is just the one-dimensional wave operator. So we can take $rk(r,t) = \delta(r-t)$. This makes $k(r,t) = r^{-1}\delta(r-t) = t^{-1}\delta(r-t)$. Now let's see what the Fourier transform of this is. Let's *assume* that we can compute the Fourier transform as

$$t^{-1}\int_{\mathbb{R}^3} e^{-ix\cdot\xi}\delta(|x|-t)\,dx = t\int_{S^2} e^{-it\xi\cdot\omega} d\omega$$

where $S^2$ is the unit 2-sphere and $d\omega$ is the standard measure on it. Now we use angular coordinates $(\theta, \phi)$ on the sphere, relative to the $\xi$ axis, and we get $d\omega = \sin\theta d\theta d\phi$ and $\xi \cdot \omega = |\xi| \cos\theta$. The integral becomes

$$t \int_0^\pi d\theta \int_0^{2\pi} d\phi e^{it|\xi|\cos\theta} \sin\theta.$$

Let $u = \cos\theta$, then this is

$$2\pi t \int_{-1}^1 e^{it|\xi|u} du = 2\pi t \Big( \frac{e^{it|\xi|} - e^{-it|\xi|}}{it|\xi|} \Big) = 4\pi \frac{\sin t|\xi|}{|\xi|}.$$

Therefore, assuming all this can be rigorously justified, the inverse Fourier transform of $\sin t|\xi|/|\xi|$ is[3]

$$G_2^{(3)}(t) = \frac{1}{4\pi} t^{-1} \delta(|x| - t), \ t > 0.$$

The solution to the wave equation in $\mathbb{R}^3$ can therefore be written

$$u(x,t) = \frac{1}{4\pi} \partial_t \Big( t \int_{S^2} f(x + \omega t) d\omega \Big) + \frac{t}{4\pi} \int_{S^2} g(x + \omega t) d\omega, \quad x \in \mathbb{R}^3.$$

Note that we get a similar 'domain of dependence' property here: the initial data $f, g$ at $x$ only affect the solution $u(y,t)$ at places where $|x - y| \le t$. In fact we have an even better result here: initial data $f, g$ at $x$ only affect the solution $u(y,t)$ at places where $|x - y|$ is *equal* to $t$. This remarkable result is called the 'strong Huygen's principle' for the wave equation in $\mathbb{R}^3$. It shows the value of working out the inverse Fourier transform of $G_i^{(3)}(t)$, since the Huygen's property was not at all evident from the form of its Fourier transform[4].

*Exercise* 12. Show that the wave equation in $3+1$ dimensions is not well-posed in $\mathbb{R}^3$. Do this by finding, for each $N$, some initial data $f_N, g_N$ whose $C^2$ norms are uniformly bounded, but with the solution $u_N$ satisfying $|u_N(x,t)| \ge N$ at some point $(x,t)$. Hint: consider a smoothed version of the fundamental solution.

The wave equation for the case $n = 2$ is a bit more difficult to derive, but it can also be obtained by a trick. Namely, it turns out that the distribution $E_2(t)$ for $n = 2$ is obtained from the distribution $E_2(t)$ for $n = 3$ by integrating out one of the variables. This gives the solution for $n = 2$:

$$u(x,t) = \frac{1}{4\pi} \partial_t \Big( t \int_{|z| \le 1} \frac{f(x + zt)}{\sqrt{1 - |z|^2}} dz \Big) + \frac{1}{4\pi} t \int_{|z| \le 1} \frac{f(x + zt)}{\sqrt{1 - |z|^2}} dz, \quad x, z \in \mathbb{R}^2.$$

This only satisfies a weak Huygen's principle: the initial data $f, g$ at $x$ only affect the solution $u(y,t)$ at places where $|x - y| \le t$, but we cannot replace the $\le$ sign with equality.

[3]The reason this is restricted to $t > 0$ is that we implicitly assumed this above, due to the fact that $r > 0$ in polar coordinates.

[4]However, it does follow from a theorem called the Paley-Wiener theorem, without having to compute the inverse Fourier transform explicitly.

The distributions $E_1^{(n)}(t)$, $E_2^{(n)}(t)$ are called fundamental solutions of the wave equation because they are themselves solutions to the wave equation, and they can be used to express all solutions to the equation via the convolution formula (15). The distributions $E_1^{(n)}(t)$, $E_2^{(n)}(t)$ can be regarded as the solutions to the wave equation when the initial datum $f$, respectively $g$ is a delta function and the other vanishes.

2.3. **Energy inequality and uniqueness.** Our work on fundamental solutions proves that solutions to the wave equation exist. What about uniqueness? We can prove uniqueness when $f$ and $g$ are Schwartz functions, as follows: define the 'energy' $E(t)$ of any solution of the wave equation $u(x,t)$ at time $t$ to be

$$E(t) = \int_{\mathbb{R}^n} \sum_{j=1}^n |\partial_j u(x,t)|^2 + |\partial_t u(x,t)|^2 \, dx.$$

Now compute the time derivative of $E(t)$:

$$\begin{aligned}
\partial_t E(t) &= \operatorname{Re} \int_{\mathbb{R}^n} \sum_{j=1}^n \Big( \partial_j u(x,t) \partial_t \partial_j \overline{u(x,t)} + \partial_t \partial_j u(x,t) \partial_j \overline{u(x,t)} \Big) \\
&\qquad + \partial_t^2 u(x,t) \partial_t \overline{u(x,t)} + \partial_t u(x,t) \partial_t^2 \overline{u(x,t)} \, dx \\
&= \operatorname{Re} \int_{\mathbb{R}^n} \sum_{j=1}^n \Big( \partial_j u(x,t) \partial_t \partial_j \overline{u(x,t)} + \partial_t \partial_j u(x,t) \partial_j \overline{u(x,t)} \\
&\qquad + \partial_j^2 u(x,t) \partial_t \overline{u(x,t)} + \partial_t u(x,t) + \partial_j^2 \overline{u(x,t)} \Big) \, dx \\
&= 0,
\end{aligned} \tag{16}$$

by integrating by parts. So $E(t) = E(0)$. But

$$E(0) = \int \sum_{j=1}^n |\partial_j f(x,t)|^2 + |g(x,t)|^2 \, dx$$

can be expressed in terms of the initial data. Now suppose that $u_1$ and $u_2$ are two solutions with the same initial data. Then $u_1 - u_2$ is a solution of the wave equation with zero initial data. Therefore the energy for the solution $u_2 - u_2$ is zero. But this implies that $(u_1 - u_2)_t \equiv 0$, so integrating from $t = 0$ we find that $u_1 = u_2$ identically. This proves uniqueness. So we have

**Theorem 3.** *Let $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then there is a unique solution to the Cauchy problem* (1), (2) *and it is given by the formula* (15).

2.4. **Weak solutions.** A strong, or classical, solution to the wave equation is a $C^2$ function which satisfies the equation (1) pointwise. As we have already hinted several times and you might suspect based on the lectures on elliptic equations, there are weaker notions of solution.

First we give the notion of a weak, or finite energy, solution. By definition, $u(x,t)$ is a finite energy solution to the wave equation if $u_t$ and $\partial_j u$ are continuous functions of $t$ with values in $L^2(\mathbb{R}^n)$, $1 \le j \le n$, and if for every $\phi \in C_c^\infty(\mathbb{R}^n \times \mathbb{R})$, the identity

$$\int_{\mathbb{R}^n \times \mathbb{R}} \big(u_t \phi_t - \nabla u \cdot \nabla \phi\big)\, dx\, dt = 0 \tag{17}$$

is satisfied. This is the analogue of a weak ($W^{1,2}$) solution to an elliptic equation.

*Exercise* 13. Show that every strong solution to the wave equation is a weak soln.

**Theorem 4.** *Let $n \ge 3$. For every pair of functions $(f, g)$ on $\mathbb{R}^n$ with $\nabla f$ and $g \in L^2$, there is a weak solution to the wave equation with initial data $(f, g)$. Moreover, the energy $E(t)$ of the solution is finite and constant in time.*

*Proof.* We already know that if $f$ and $g$ are Schwartz, then there is a solution to the equation given by the fundamental solution. Moreover, it is not hard to show that for each time $t$, $u(\cdot, t)$ is a Schwartz function, as is every time-derivative $(\partial_t)^k u$.

*Exercise* 14. Prove this.

Thus it is a strong solution, and the calculation above shows that the energy is constant in time. (Notice that the calculation above requires that $u$ has *two* derivatives in $L^2$.)

Now $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2$, so given $(f, g)$ with $\nabla f$ and $g \in L^2$, we can find a sequence of pairs of Schwartz functions $(f_n, g_n)$, such that $\nabla f_n \to \nabla f$ and $g_n \to g$ in $L^2$.

*Exercise* 15. Prove this density statement. Note that compactly supported functions are dense in $L^2$, so it is enough to consider compactly supported $L^2$ functions $f$. Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ with $\int_{-\infty}^{\infty} \phi(x)dx = 1$, and let

$$\phi_\epsilon(x) = \epsilon^{-n}\phi(x/\epsilon).$$

Show that $f_n = f * \phi_{1/n} \in \mathcal{S}(\mathbb{R}^n)$ and $f_n \to f$ in $L^2$. (Use the Fourier transform.)

From each pair we get a strong solution $u_n$ which has finite energy. I claim that the $u_n$ converge, in some sense, to a finite energy solution.

The only tricky thing about proving this claim is determining in what sense the $u_n$ converge. Let's start with the time derivatives; denote $v_n = (u_n)_t$. The sequences $(\nabla f_n)$ and $(g_n)$ are Cauchy, so given $\epsilon > 0$ there is an $N$ such that

$$\int_{\mathbb{R}^n} \sum_{j=1}^{n} |\partial_j f_n - \partial_j f_m|^2 + |g_n - g_m|^2\, dx < \epsilon \text{ if } n, m \ge N.$$

The solution to the wave equation with initial data $(f_n - f_m, g_n - g_m)$ is $u_n - u_m$. The time derivative of this is $v_n - v_m$ and our energy equality implies that, for every positive time,

$$\|v_n - v_m\| < \epsilon \text{ if } n, m \geq N.$$

Thus the $v_n(t)$ form a Cauchy sequence in $L^2$ for each $t$, and therefore converge in $L^2$ to a function $v(\cdot, t)$. This may be regarded as a continuous function of $t$ with values in $L^2$ functions of $x$. Reason: for a fixed time $t_0$ choose an interval $I = (t_0 - \delta, t_0 + \delta)$ so that

$$\|v_N(t) - v_N(t_0)\|_{L^2} \leq \epsilon \text{ if } t \in I.$$

This is possible because $u_N$ is a strong solution. Now, for $t \in I$ we can estimate

$$\|v(t)-v(t_0)\|_{L^2} \leq \|v(t)-v_N(t)\|_{L^2}+\|v_N(t)-v_N(t_0)\|_{L^2}+\|v_N(t_0)-v(t_0)\|_{L^2} \leq 3\epsilon.$$

This proves continuity of $v$ in time. So we can integrate in $t$, and set

$$u(\cdot, t) = f(\cdot) + \int_0^t v(\cdot, s)\, ds = f(\cdot) + \lim_n \int_0^t v_n(\cdot, s)\, ds.$$

Now by taking difference quotients (which approximate a spatial partial derivative) in this equation and taking a limit, we can show that $\partial_j u$ is in $L^2$ and is the limit of $(\partial_j u_n)$ (which we already know is a Cauchy sequence in $L^2$). This proves that $u$ is a finite energy solution. Finally, since $\nabla u = \lim_n \nabla u_n$ and $u_t = \lim_n (u_n)_t$, the energy $E(t)$ for $u$ is the limit of the energy $E_n(t)$ for $u_n$. This is constant in time and tends to $\|\nabla f\|_{L^2}^2 + \|g\|_{L^2}^2$ as $n \to \infty$, which proves the final statement of the theorem. □

2.5. **Stronger solutions.** It is not hard to get the following regularity result from Theorem 4.

**Theorem 5.** *For every pair of functions $(f, g)$ on $\mathbb{R}^n$ with $f \in W^{k,2}(\mathbb{R}^n)$ and $g \in W^{k-1,2}(\mathbb{R}^n)$, the weak solution to the wave equation with initial data $(f, g)$ is a continuous function of $t$ with values in $W^{k,2}(\mathbb{R}^n)$.*

This is proved in a similar way to elliptic regularity results; either by taking difference quotients or by differentiating the solution to the equation. The main point is that if $u$ solves the wave equation, then so does every derivative $D^\alpha u$.

2.6. **Distributional solutions.** We can define a much weaker notion of solution to the wave equation, by noting that the wave equation makes sense for *distributions* (simply because repeated partial differentiation is defined for distributions). Thus, we can regard the wave equation as an equation in distributions, and ask to find distributional solutions, say with given distributional initial data. The theory for this works very nicely; for example, there is an analogue of the theorem above:

**Theorem 6.** *Let $n \geq 3$. For every pair of tempered distributions $(f, g)$ on $\mathbb{R}^n$, there is a distributional solution to the wave equation with initial data $(f, g)$. It may be regarded as $C^\infty$ function of time with values in distributions on $\mathbb{R}^n$.*

We will not prove this theorem here. Earlier we noted that if $h(s)$ is a $C^2$ function of one variable, and $\omega$ a unit vector in $\mathbb{R}^n$, then $u(x,t) = h(x \cdot \omega - t)$ is a solution to the wave equation. If you did the computation you will have noticed that there was no particular property of $h$ required to hold, except that its derivatives existed (in some sense). There is a distributional generalization of this: if $h$ is any *distribution* in one variable, then $u(x,t) = h(x \cdot \omega - t)$ is a distributional solution to the wave equation.

## 3. Variable coefficients — existence theory

In this section we consider existence theory for variable coefficient wave equations, closely following a paper of P. Lax, 'On Cauchy's problem for hyperbolic equations and the differentiability of solutions of elliptic equations', Communications on Pure and Applied Mathematics 8, 1955, 615-633. This is a wonderfully insightful paper that is well worth reading fifty years after its publication.

### 3.1. Variable coefficients.

In this section we will consider variable-coefficient wave equations. The model to keep in mind is

$$u_{tt} = \sum_{i,j} \partial_i(a_{ij}(x)\partial_j u), \tag{18}$$

where $a_{ij}(x)$ are smooth functions which are uniformly elliptic at each point: i.e. there are positive constants $c, C$ such that

$$c|\xi|^2 \leq \sum_{i,j} a_{ij}(x)\xi_i\xi_j \leq C|\xi|^2. \tag{19}$$

Why consider such equations? There is motivation from physics. In electromagnetism, for example, the electric and magnetic fields satisfy wave equations

$$u_{tt} = c^2 \Delta u,$$

where $c$ is the speed of light. However, the speed of light is not constant; it varies depending on the medium, so it is really a function of space (and time). Another motivation, from pure mathematics, comes from Riemannian geometry. There we consider a metric function $g_{ij}(x)$ on a manifold, and there is an associated Laplace operator which looks like the right hand side[5] of (18) with $a_{ij} = g^{ij}(x)$, the inverse matrix of $g_{ij}(x)$.

[5] up to lower order terms, which are less important

3.2. **Uniqueness.** The energy $E(t)$ at time $t$ in this case is defined by

$$E(t) = \int_{\mathbb{R}^n} \Big( |u_t|^2 + \sum_{i,j} a_{ij}(x) u_i(x,t) \overline{u_j(x,t)} \Big) \, dx.$$

A computation similar to the one above shows that

$$\frac{d}{dt} E(t) = 0$$

and therefore shows *uniqueness* of the Cauchy problem.

*Exercise* 16. Do this computation.

An interesting variation of this idea is the following: Choose a point $x_0 \in \mathbb{R}^n$ and a time $T > 0$. Consider the 'local energy'

$$F(t) = \int_{B(x_0, C(T-t))} \Big( |u_t|^2 + \sum_{i,j} a_{ij}(x) u_i(x,t) \overline{u_j(x,t)} \Big) \, dx$$

for $0 \leq t \leq T$. So we integrate in a ball around $x_0$ with a shrinking radius, vanishing at $t = T$ (the $C$ here is the same as in (19)).

*Exercise* 17. Show that $dF(t)/dt \leq 0$.

This shows *finite propagation speed* for this equation; if the initial data $f, g$ are zero in the ball $B(x_0, CT)$ then the solution $u(x,t)$ vanishes in the ball $B(x_0, C(T-t))$ at time $t$.

3.3. **Symmetric hyperbolic systems.** Following Lax, we will actually consider a variation of this setup and consider symmetric hyperbolic systems. By definition, a symmetric hyperbolic system is a first order PDE for a vector of functions

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_m \end{pmatrix}$$

with $u_i(x,t)$ a function on $\mathbb{R}^n \times \mathbb{R}$, of the form

$$Lu \equiv u_t + \sum_{j=1}^{n} A_i(x,t) \frac{\partial u}{\partial x_i} + B(x,t) u = v(x,t), \tag{20}$$

where $A_i$ and $B$ are real symmetric $m \times m$ matrices, depending on $x$ and $t$. Let's show that the wave equation can be put into this form. We let $u$ be the $(n+1)$-column vector

$$\begin{pmatrix} u_t \\ u_{x_1} \\ u_{x_2} \\ \dots \\ u_{x_n} \end{pmatrix}.$$

Label these components $u_0, u_1, \ldots, u_n$. Letting $B = 0$, $v = 0$ and $A_i$ the matrix with 1 in the $(0,i)$ and $(i,0)$ position and zeroes elsewhere, it is easy to check that this reproduces the wave equation $u_{tt} = \Delta u$.

The variable-coefficient wave equation (18) can also be put into symmetric hyperbolic form. This is less obvious, and I leave it as an exercise:

*Exercise* 18. Let $A(x)$ be the matrix with entries $a_{ij}(x)$. Since it is a positive definite matrix, it has a square root $B(x) = (b_{ij}(x))$. Show that the $(n+1)$-vector with entries $u_t$ and $\sum_j b_{ij}\partial_j u$ satisfy a symmetric hyperbolic system.

3.4. **Weak solutions and the adjoint operator.** Let $u(x,t)$ be an $L^2$ function of $x$ and $t$. We shall say that $u$ is a weak solution to the equation (20) if, for all $m$-vectors of functions $\phi(x,t) \in C_c^\infty(\mathbb{R}^n \times \mathbb{R})$, the identity

$$\int_{\mathbb{R}^n\times\mathbb{R}} u \cdot \Big( -\phi_t - \frac{\partial}{\partial x_j}(A_j\phi) + B\phi \Big)\, dx\, dt = \int_{\mathbb{R}^n\times\mathbb{R}} v \cdot \phi \, dx\, dt \tag{21}$$

is satisfied. We can write this equation briefly as

$$\langle u, L^*\psi\rangle = \langle v, \psi\rangle, \qquad \text{for all } \psi \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}), \tag{22}$$

where here $\psi = \overline{\phi}$ and $\langle , \rangle$ is the $L^2$ pairing (and is conjugate linear in the second variable, which explains the need to conjugate $\phi$). Of course this definition is rigged so that every strong (classical) solution to the equation is a weak solution. The operator

$$L^*\phi = -\phi_t - \frac{\partial}{\partial x_j}(A_j\phi) + B\phi$$

is called the adjoint operator and is defined by the property that

$$\langle L\chi, \phi\rangle = \langle \chi, L^*\phi\rangle,$$

for all smooth compactly supported $\chi, \phi$. (Notice that we are using the symmetry of the matrices $A_i$ and $B$ here; in general we would get the transposed matrices $A_i^t$ and $B^t$ in the formula for the adjoint operator.)

3.5. **The periodic problem.** Our aim is to solve the Cauchy problem, i.e. to solve the the initial value problem

$$Lu = 0, \quad u(\cdot, 0) = \phi, \quad \phi(x) \in L^2 \tag{23}$$

for the unknown function $u(x,t)$ (more precisely, an $m$-vector of functions) in terms of the given function $\phi$. However we will approach it in a sneaky way due to Lax (building on an idea of Petrovskii). That is, we shall first solve the *periodic problem*

$$Lu = v \tag{24}$$

on a torus $\mathbb{T}^n \times \mathbb{T}$. We can think of the torus $\mathbb{T}^n$ here as the set $[0,R]^n$ with $x_j = 0$ and $x_j = R$ identified, or alternately as $\mathbb{R}^n$ quotiented by a lattice of translations generated by translations of distance $R$ parallel to any of the coordinate axes. Similarly the torus (circle) $\mathbb{T}$ is the time axis quotiented by a translation of duration $T$, so we identify $t = 0$ and $t = T$. Now, to make our problem 'live on the torus' we shall assume that the matrices $A_i$ and $B$ are periodic functions, so that they may be regarded as matrix-valued functions on $\mathbb{T}^n \times \mathbb{T}$. Isn't this changing the problem? Yes it is, but because of finite propagation speed (which follows from exercise (17), and is valid more generally for symmetric hyperbolic systems), changing the coefficient matrices at some far-distant place has no bearing on the solution locally, at least for some long time. So if we are interested in the Cauchy problem for a short time, we may modify our coefficient matrices so that they live on the torus. Then, given a function $v$ on the torus (i.e. periodic), our immediate goal is to find a solution to (24) with a function $u$ on the torus (i.e. periodic).

This we shall do with a little Hilbert space theory. We need the Riesz representation theorem and some density arguments.

We start with an estimate of the adjoint operator. Before we do this, we make a useful reduction. That is: we can always add any multiple of the identity matrix to $B$, and we get an equivalent problem. For let $L_\lambda = L + \lambda I$ be the operator we get by adding $\lambda I$ to $B$. Then $u(x,t)$ is a solution of $Lu = 0$ iff $u(x,t)e^{-\lambda t}$ is a solution to $L_\lambda u = 0$. Because of this simple one-to-one correspondence between solutions, we can regard these two problems as equivalent, and so we can add any multiple of the identity to $B$ if we feel like it. This turns out to be a very useful thing to do! It allows us to assume that

$$B - \frac{1}{2}\sum_j \partial_j A_j \geq I \text{ on } \mathbb{T}^n \times \mathbb{T}. \tag{25}$$

Assuming this we prove the following estimate (where we write $M = \mathbb{T}^n \times \mathbb{T}$ for simplicity).

**Lemma 7.** *Suppose that* (25) *is satisfied. Then for every $C^1$ function $a(x,t)$, we have*

$$\|a\|_{L^2(M)} \leq \|L^* a\|_{L^2(M)}. \tag{26}$$

*In particular, $L^*$ is one-to-one; if $L^* a = 0$ then $a = 0$.*

*Proof.* We compute the inner product

$$\langle L^* a, a\rangle = \langle -a_t - \partial_j(A_j a) + Ba, a\rangle = \int_M \Big(-a_t - \partial_j(A_j a) + Ba\Big)\cdot \overline{a}\, dx\, dt. \tag{27}$$

We want to show that this is big. It is sufficient to show that the real part is big, so we will study the real part of this expression. Consider first the term

$$\int_M -a_t \cdot \overline{a}\, dx\, dt.$$

The real part of this is

$$\int_M \partial_t\big(a \cdot \overline{a}\big)\, dx\, dt$$

and the integral in time is zero due to periodicity and the fundamental theorem of calculus. The spatial derivative terms work out similarly, but there is a catch because the matrices $A_j$ are not constant. We get

$$\begin{aligned}\operatorname{Re}\int_M \Big(-\partial_j(A_j a)\Big)\cdot\overline{a}\,dx\,dt &= \operatorname{Re}\int_M \Big(-\partial_j\big(A_j a\cdot\overline{a}\big) + A_j a\cdot\partial_j\overline{a}\Big)\,dx\,dt \\ &= \operatorname{Re}\int_M a\cdot A_j\partial_j\overline{a}\,dx\,dt \\ &= \operatorname{Re}\int_M \Big(a\cdot\partial_j(A_j\overline{a}) - a\cdot(\partial_j A_j)\overline{a}\Big)\,dx\,dt.\end{aligned}$$

(A summation convention over repeated indices is in force here.) The first term on the right is minus the term on the left, so we have

$$-2\operatorname{Re}\int_M \Big(-\partial_j(A_j a)\Big)\cdot\overline{a}\,dx\,dt = -\operatorname{Re}\int_M a\cdot(\partial_j A_j)\overline{a}\,dx\,dt.$$

So we get that the real part of (27) is equal to

$$\int_M \Big(a\cdot(B-\frac{1}{2}\partial_j A_j)\overline{a}\Big)\,dx\,dt.$$

Now our assumption (25) comes to the rescue: the matrix $B-\partial_j A_j/2$ is everywhere positive definite, indeed $\geq I$ everywhere. This implies that the absolute value of the real part of (27) is at least

$$\int_M \Big(a\cdot\overline{a}\Big)\,dx\,dt = \|a\|^2_{L^2(M)}.$$

So we have shown that

$$|\langle L^*a, a\rangle| \geq \|a\|^2_{L^2(M)}.$$

On the other hand,

$$|\langle L^*a, a\rangle| \leq \|a\|_{L^2(M)}\|L^*a\|_{L^2(M)}$$

and these two inequalities imply (26). □

Now we can quite quickly show the existence of weak solutions of eqrefperiodic. Given $v$ as in (24), consider the linear functional $a\mapsto\langle a, v\rangle$. Consider the vector space

$$W = \{b\in L^2(M)\mid b = L^*a,\ a\in C^1(M)\}$$

and define a linear functional on $W$ by

$$b\mapsto\langle a, v\rangle \text{ where } b = L^*a.$$

This is well defined since there is at most one $a$ with $L^*a = b$, due to Lemma 7. Moreover,

$$|\langle a, v\rangle| \leq \|a\|\|v\| \leq \|b\|\|v\|,$$

so this is a bounded linear functional on $W$. Since it is bounded on $W$ it extends to the whole space $L^2(M)$, for example by being defined as zero on the orthocomplement of $W$. By the Riesz representation theorem, this linear functional is inner product with some function $u$. So there exists a $u$ such that

$$\langle a, v\rangle = \langle b, u\rangle \implies \langle a, v\rangle = \langle L^* a, u\rangle \text{ for all } a \in C^1(M).$$

But this is equivalent to saying that $u$ is a weak solution to the equation $Lu = v$!

In Lax's paper more is proved: it is shown that if $v$ has a certain number of derivatives in $L^2$, then so does $u$ (the analogue of Theorem 5). This is done using the same argument but in the Sobolev space $W^{k,2}(M)$ rather than $L^2(M)$. We will skip this step because of time constraints, but it is needed below for technical reasons.

3.6. **Solution of the Cauchy problem.** Finally we show that (23) has a solution. Let us consider the vector space $V$ of smooth initial conditions, $\phi \in C^\infty(\mathbb{T}^n)$ such that there exists a smooth solution to (23). At the moment, the only thing we know about $V$ is that it contains the zero vector. However, we shall soon show that $V$ is dense in $L^2(\mathbb{T}^n)$.

We now prove an a priori estimate for solutions of (23) with initial conditions in $V$. Note that we are not assuming that solutions actually exist — the statement will be that *if* a solution exists then it has to satisfy the following inequality.

**Lemma 8.** *Suppose that $u(x,t)$ is a smooth solution to* (23) *with initial condition $\phi \in V$. Then*

$$\|u(\cdot,t)\|^2_{L^2(\mathbb{T}^n)} \le e^{-t}\|u(\cdot,0)\|^2_{L^2(\mathbb{T}^n)}.$$

*Proof.* This is similar to the calculation involving $L^* a$. We compute

$$\partial_t \|u(\cdot,t)\|^2_{L^2(\mathbb{T}^n)} = \partial_t\langle u, u\rangle = \langle -A_j\partial_j u - Bu, u\rangle = \langle u, \partial_j(A_j u) - Bu\rangle$$

Our previous estimates show that this is no bigger than $-\langle u, u\rangle$. So if we call $\|u(\cdot,t)\|^2_{L^2(\mathbb{T}^n)} = k(t)$, then we have established the differential inequality

$$k'(t) \le -k(t)$$

for $k$. This can be integrated up to $k(t) \le k(0)e^{-t}$, which proves the lemma. □

Now let us define the 'solution operator' $S$ on the vector space $V$. This takes $\phi \in V$ and sends it to $S\phi = u(\cdot, T)$, where $u$ is the solution with initial condition $\phi$. We have just shown that

$$\|S\phi\|_{L^2} \le e^{-T}\|\phi\|_{L^2},$$

i.e. the operator norm of $S$ is no bigger than $e^{-T}$. In particular,

(28) The operator norm of $S$ is strictly less than 1.

Now we shall find a large class of solutions to (23), that will prove that $V$ is dense. This is done in the following way: let $\psi \in C^\infty(\mathbb{T}^n)$, and define

$$u_1(x,t) = -\frac{t}{T}\psi(x).$$

Now using the previous section, solve the following equation for $u_2$:

$$Lu_2 = -Lu_1 \text{ with } u_2 \in L^2(M).$$

In particular $u_2$ is periodic, so $u_2(x,T) = u_2(x,0) = \phi$, say. Let $u = u_1 + u_2$. Then we have found a $u$ with $Lu = 0$, and

$$\begin{aligned} u(\cdot, 0) &= \phi \\ u(\cdot, T) &= -\psi + \phi. \end{aligned} \tag{29}$$

By the remark at the end of the previous subsection, since $Lu_1$ is smooth, so is $u_2$. It follows that $\phi$ belongs to $V$, and we have

$$\psi = \phi + S\phi. \tag{30}$$

Now this, together with (28), shows that $V$ is dense. The reason for this is that, since $S$ is bounded on $V$, it extends to a bounded linear map on the whole of $L^2(M)$, for example by being defined as zero on the orthocomplement of $V$. Then we can invert $\mathrm{Id} - S$ on $L^2(M)$, since $\|S\| < 1$ by (28). We then have $\phi = (\mathrm{Id} - S)^{-1}\psi \in V$. Since the set of $\psi$ is dense and $(\mathrm{Id} - S)^{-1}$ is an isomorphism, $V$ is also dense.

*Exercise* 19. Show from first principles that if $S$ is a linear operator on a Hilbert space and $\|S\| < 1$ then $\mathrm{Id} - S$ is one-to-one and onto. (To show that it is onto, use a method of successive approximations.)

So now we know that the Cauchy problem can be solved for a dense set of data. Now we are in the same situation as for the standard wave equation in subsection 2.4. The proof that there are solutions for every $\phi$ proceeds in the same way, by a density argument.

## 4. Weyl's asymptotic formula

In this section we'll do something completely different and look at what the wave equation can tell us about eigenvalues and eigenfunctions of the Laplace operator on manifolds.

I will not assume any knowledge of Riemannian geometry in this lecture, but I will make a few comments to motivate what we're about to do. A Riemannian manifold $(M, g)$ is a manifold $M$ together with a Riemannian metric on it. The metric determines, in particular, two important things: a *measure* on $M$, with which we can define the Hilbert space $L^2(M)$, and a *Laplace-Beltrami operator* $\Delta_g$ on $M$. I won't say how $\Delta_g$ is defined, but when $M = \mathbb{R}^n$ and $g$ is the standard metric (i.e. inner product) on $\mathbb{R}^n$ then $\Delta_g f = \sum_j \partial_j^2 f$ is the usual Laplacian.

Suppose that $M$ is a compact manifold — for example, the torus $\mathbb{T}^n$ from the previous section. Then $-\Delta_g$ is an (unbounded) self-adjoint operator on $L^2(M)$ with discrete spectrum tending to infinity. This means that there is an orthonormal basis $u_1, u_2, \dots$ of $L^2(M)$ consisting of eigenfunctions: $-\Delta_g u_j = \lambda_j u_j$ for some sequence of real numbers (eigenvalues) $\lambda_j \to \infty$.

Let's just give the simplest example here: if $M$ is a circle and $g$ is a metric that gives $M$ length $2\pi$, then the eigenfunctions of $-\Delta_g$ are

$$1, \cos\theta, \sin\theta, \cos 2\theta, \sin 2\theta, \ldots$$

and the sequence of eigenvalues $\lambda_j$ is $0, 1, 1, 4, 4, 9, 9, \ldots$.

The wave equation on $M$ can be used to get information about the eigenvalues and eigenfunctions of $-\Delta_g$. In particular, we can get information about the asymptotics of the *counting function* $N(\lambda)$, which by definition is the number of $\lambda_j$ which are less than $\lambda^2$. This is an increasing, integer-valued function and the growth rate of $N(\lambda)$ as $\lambda \to \infty$ is of considerable interest.

For the remainder of this lecture we shall consider the example of the $n$-torus $\mathbb{T}^n$ as in the previous lecture, and the wave equation $u_{tt} = \sum_j \partial_j^2 u$ on the torus. However, bear in mind that this technique works for far less symmetric settings.

4.1. **Operator Theory.** Recall that if $A$ is a diagonal matrix, with entries $\mu_i$ down the diagonal, then we define $p(A)$ to be the diagonal matrix with diagonal entries $p(\mu_i)$. We do the same for $\Delta$ on the torus; the operator $p(-\Delta)$ is defined to be the operator that takes $u_j$ to $p(\lambda_j)u_j$, which defines a unique operator since $(u_j)$ is a basis. It will be useful below to write $P$ for oft-occuring operator $\sqrt{-\Delta}$. Thus, $P$ maps $u_j$ to $\sqrt{\lambda_j}u_j$.

Now, let's choose an function $\chi(t)$ which is supported in the interval $[-R, R]$. Let $\hat{\chi}$ be its Fourier transform. The operator $\hat{\chi}(P)$ is by definition the operator that maps $u_j$ to $\hat{\chi}(\sqrt{\lambda_j})u_j$. Now

$$\hat{\chi}(\sqrt{\lambda}) = \int_{-\infty}^{\infty} \chi(t) e^{-it\sqrt{\lambda}}\, dt.$$

It follows from this that

$$\hat{\chi}(P) = \int_{-\infty}^{\infty} \chi(t) e^{-itP}\, dt.$$

Here of course $e^{-itP}$ is the operator that takes $u_j$ to $e^{-it\sqrt{\lambda_j}}u_j$. Now, the point of the strange-looking operator $e^{-itP}$ is that it is the *solution operator for the wave equation* with initial conditions

$$u_{tt} = \Delta u, \quad u(\cdot, 0) = f, \quad u_t(\cdot, 0) = -iPf.$$

To prove this, take $f = u_j$. Then it can be checked explicitly that the solution of this PDE is $u(x,t) = e^{-it\sqrt{\lambda_j}}u_j(x) = e^{-itP}f$. Since this is true for a basis of $f$'s, it is true for all $f$ (the wonderful property of linearity!).

Now here is a powerful idea. Since the wave equation has the property of finite propagation speed, for $t < R$ the fundamental solution of the wave equation on the torus is 'the same as' the fundamental solution for the wave equation on $\mathbb{R}^n$. (Intuitively, if you have travelled distance $< R$ then you can't tell whether you are on the torus of length $R$ or on $\mathbb{R}^n$.) Notice that the operator $e^{-itP}$ on $\mathbb{R}^n$ has the form

$$\mathcal{G}e^{-it|\xi|}\mathcal{F},$$

and therefore acts on functions according to

$$f \mapsto (2\pi)^{-n} \int e^{ix\cdot\xi} e^{-it|\xi|} e^{-iy\cdot\xi} f(y)\, dy\, d\xi.$$

So this is true also on the torus (in an appropriate sense — we are abusing notation here). The operator $\hat{\chi}(P)$ therefore acts according to

(31)

$$f \mapsto (2\pi)^{-n} \int_{-\infty}^{\infty} dt\, \chi(t) \int e^{ix\cdot\xi} e^{-it|\xi|} e^{-iy\cdot\xi} f(y)\, dy\, d\xi = (2\pi)^{-n} \int e^{i(x-y)\cdot\xi} \hat{\chi}(|\xi|)\, d\xi.$$

We can now prove

**Lemma 9.** *Let $\lambda_j$ be the sequence of eigenvalues of the Laplacian on the $n$-torus of length $R$, and let $\chi$ be a smooth function supported in $[-R, R]$. Then*

$$\sum_j \hat{\chi}(\sqrt{\lambda_j}) = \big(\frac{R}{2\pi}\big)^n \int_{\mathbb{R}^n} \hat{\chi}(|\xi|)\, d\xi. \tag{32}$$

*Proof.* Consider the operator $\hat{\chi}(P)$ as above. It can be represented as a kernel $K(x, y)$, that is, acting as

$$f \mapsto \hat{\chi}(P) f = \int K(x, y) f(y)\, dy$$

in two different ways. Using the fact that it maps $u_j$ to $\hat{\chi}(\sqrt{\lambda_j})u_j$, its kernel can be represented

$$K(x, y) = \sum_j \hat{\chi}(\sqrt{\lambda_j}) u_j(x)\overline{u_j(y)}.$$

(To check this, test against each $u_k$ and see that it gives the right thing.) On the other hand, the kernel is represented by (31). Now compute $\int K(x, x)\, dx$. This is given by

$$\int \sum_j \hat{\chi}(\sqrt{\lambda_j}) |u_j(x)|^2\, dx = \sum_j \hat{\chi}(\sqrt{\lambda_j}).$$

On the other hand, using (31),

$$K(x, x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{\chi}(|\xi|)\, d\xi.$$

Integrating this over the torus $\mathbb{T}^n$ gives the right hand side of (32). This completes the proof. □

4.2. **The counting function.** Now we are going to investigate the asymptotics of the counting function. One problem with the counting function is that it is quite rough — it is integer-valued, with jumps — and it is easier to analyse a smoothed version. We will smooth it by convolving it with the function $\hat\rho$, the Fourier transform of a smooth function $\rho$ of compact support. We will assume that $\int \hat\rho = 1$, and will also need a couple of extra conditions, explained later. So consider

$$\hat\rho * \frac{d}{d\lambda} N(\lambda) = \sum_j \hat\rho(\lambda - \sqrt{\lambda_j}).$$

We are going to do some analytic manipulations with this, and will soon need some 'extra decay', so we introduce another decaying function $\hat\chi$ as before. We shall assume that $\hat\chi$ is the Fourier transform of a smooth function supported in $[-R/2, R/2]$, and write the above expression as a limit

$$\lim_{\epsilon\to 0} \sum_j \hat\rho(\lambda - \sqrt{\lambda_j}) \hat\chi(\epsilon\sqrt{\lambda_j}).$$

Now, writing $\hat\rho$ as a Fourier transform, this is

$$\lim_{\epsilon\to 0} \int_{-\infty}^{\infty} e^{-it\lambda} \rho(t) \sum_j e^{it\sqrt{\lambda_j}} \hat\chi(\epsilon\sqrt{\lambda_j}).$$

Now we assume in addition that $\rho$ is supported in the interval $[-R/2, R/2]$. Then we only need to consider values of $t$ in this interval. The function $e^{it\sqrt{\lambda_j}} \hat\chi(\epsilon\sqrt{\lambda_j})$ is the Fourier transform of $\epsilon^{-1}\chi(\cdot/\epsilon - t)$ which for $\epsilon < 1$ and $t \in [-R/2, R/2]$ has support in $[-R, R]$. Therefore we can apply Lemma 9 above, and get

$$\hat\rho * \frac{d}{d\lambda} N(\lambda) = \big(\frac{R}{2\pi}\big)^n \lim_{\epsilon\to 0} \int_{-\infty}^{\infty} e^{-it\lambda} \rho(t) \int_{\mathbb{R}^n} e^{it|\xi|} \hat\chi(\epsilon|\xi|)\, d\xi\, dt.$$

Now in the $\xi$ integral we can use polar coordinates because the integrand only depends on $|\xi|$. Integrating in the angular variables gives us $n\omega_n$, the volume of the unit $(n-1)$-sphere and we get

$$\big(\frac{R}{2\pi}\big)^n n\omega_n \lim_{\epsilon\to 0} \int_{-\infty}^{\infty} e^{-it\lambda} \rho(t) \int_0^{\infty} e^{it\tau} \hat\chi(\epsilon\tau) \tau^{n-1}\, d\tau\, dt.$$

Performing the $t$ integral we get

$$\big(\frac{R}{2\pi}\big)^n n\omega_n \lim_{\epsilon\to 0} \int_0^{\infty} \hat\rho(\tau - \tau) \hat\chi(\epsilon\tau) \tau^{n-1}\, d\tau\, dt.$$

Let us write $\tau_+^{n-1}$ for the function that is $\tau^{n-1}$ when $\tau \geq 0$ and 0 when $\tau \leq 0$. We can now take the limit as $\epsilon \to 0$ and get

$$\hat\rho * \frac{d}{d\lambda} N(\lambda) = \big(\frac{R}{2\pi}\big)^n n\omega_n \int_{-\infty}^{\infty} \hat\rho(\tau - \tau) \tau^{n-1}\, d\tau = \big(\frac{R}{2\pi}\big)^n n\omega_n \hat\rho * \tau_+^{n-1}.$$

So we have proved the equality

$$\hat\rho * \frac{d}{d\lambda} N(\lambda) = n\omega_n \big(\frac{R}{2\pi}\big)^n \hat\rho * \tau_+^{n-1}. \tag{33}$$

We can integrate this equation to get

$$\hat\rho * N(\lambda) = \omega_n \big(\frac{R}{2\pi}\big)^n \hat\rho * \tau_+^{n}.$$

Now intuitively if we convolve the rather smooth function $\tau_+^n$ by $\hat\rho$, we get something quite similar to $\tau_+^n$. In fact, it isn't hard to show that you get $\tau^n + O(\lambda^{n-1})$ as $\lambda \to \infty$.

*Exercise* 20. Show that

$$\hat\rho * \tau_+^n(\lambda) = \lambda^n + O(\lambda^{n-1}) \text{ as } \lambda \to \infty.$$

(Hint: as a warm-up, show that $\hat\rho * \lambda^n$ is a polynomial of degree $n$ in $\lambda$.)

So we have shown that

$$\hat\rho * N(\lambda) = \big(\frac{R}{2\pi}\big)^n \omega_n \lambda^n + O(\lambda^{n-1}) \text{ as } \lambda \to \infty. \tag{34}$$

This is nice — it says that if we smooth the counting function a bit, it behaves as $(R/2\pi)^n\omega_n\lambda^n$ for larger $n$. So there are about $(R/2\pi)^n\omega_n\lambda^n$ eigenvalues less than or equal to $\lambda^2$ for the Laplacian on the torus.

What about the counting function itself — can we get rid of the convolution? It turns out that you can, due to a very cunning argument due to Hörmander.

**Theorem 10.** *The counting function*

$$N(\lambda) = \big(\frac{R}{2\pi}\big)^n \omega_n \lambda^n + O(\lambda^{n-1}). \tag{35}$$

*Proof.* We already know that this asymptotic is true for $\hat\rho * N$, so we need to show that

$$N = \hat\rho * N = O(\lambda^{n-1}). \tag{36}$$

To do this we need to make a further assumption on $\hat\rho$; as well as our assumptions so far (namely, that its support of $\rho$ is contained in $[-R/2, R/2]$ and that $\int \hat\rho = 1$), we need to assume that $\hat\rho$ is strictly positive.

*Exercise* 21. Show that such a function $\hat\rho$ exists.

We can write

$$(N * \hat\rho)(\lambda) - N(\lambda) = \int_{-\infty}^{\infty} \hat\rho(\lambda')\big(N(\lambda-\lambda') - N(\lambda)\big)\, d\lambda'. \tag{37}$$

To get a handle on this, we estimate $N(\lambda - \lambda') - N(\lambda)$. We can write

$$N(\lambda+1) - N(\lambda) = \int_{\lambda}^{\lambda+1} \frac{d}{d\lambda} N(\lambda')\, d\lambda'$$

and, if $|\hat{\rho}(\lambda')| \geq c > 0$ for $\lambda' \in [-1, 1]$, then this is bounded by

$$c^{-1} \int_{-\infty}^{\infty} \frac{d}{d\lambda} N(\lambda') \rho(\lambda - \lambda') \, d\lambda' = c^{-1} \hat{\rho} * \frac{d}{d\lambda} N(\lambda)$$

which, by (33), we know is $O(\lambda^{n-1})$. So

$$N(\lambda + 1) - N(\lambda) = O(\lambda^{n-1}), \lambda \to \infty.$$

By induction, then we have

$$N(\lambda + k) - N(\lambda) = O(|k|(\lambda + |k|)^{n-1}), \lambda \to \infty$$

for all integers $k$. Now for a noninteger $\lambda'$, using the monotonicity of $N$ we get

$$N(\lambda + \lambda') - N(\lambda) = O((1 + |\lambda'|)(\lambda + |\lambda'| + 1)^{n-1}), \lambda \to \infty$$

by replacing $\lambda'$ by the nearest integer larger in absolute value. If we plug this into (37) then we get a bound

$$\big|(N * \hat{\rho})(\lambda) - N(\lambda)\big| \leq C \int_{-\infty}^{\infty} \hat{\rho}(\lambda')(1 + |\lambda'|)(1 + |\lambda| + |\lambda'|)^{n-1}\big) \, d\lambda'. \tag{38}$$

If we expand the term $(1 + |\lambda| + |\lambda'|)^{n-1}$ using the binomial series then we get a series of terms each of which is $O(\lambda^{n-1})$ (the integrals all converge since $\hat{\rho}$ is rapidly decreasing). So we have shown (36), and this completes the proof. □

*Exercise* 22. For the Laplacian on the torus of length $\pi$, show that the eigenvalues are given by $m^2 + n^2$ for every integer lattice point $(m, n)$ in the plane. Hence prove the asymptotic (35) for counting function $N(\lambda)$ directly.

This might seem a bit disappointing in view of all the work we just did! However, the point is that the method here works generally for Riemannian metrics on manifolds. You need to replace the Fourier transform calculation here with something called the 'Hadamard parametrix' for the wave equation on such manifolds. The general result is that for a Riemannian manifold $(M, g)$, the counting function satisfies

$$N(\lambda) = (2\pi)^{-n} \omega_n \operatorname{vol}(M) \lambda^n + O(\lambda^{n-1}). \tag{39}$$

A quick application of this result: the dimension of the $\lambda$ eigenspace is no bigger than $C\lambda^{n-1}$, since otherwise this would contradict the asymptotics above.