Lec 1

# Welcome to PHYS 312: Introduction to Mathematical Physics

Instructor: Joanna Karczmarek (karch-MA-rek, she/her)

Format: In person, attendence is required.

However: do not come to class if you are sick.
I will provide accomodations when you have to miss class.

Most of this course is about two connected topics:

- linear (partial) differential equations
  and
- Fourier series.

We will also discuss

- the physical derivations and interpretations of the differential equations we study,
- common special functions (Bessel, Legendre, spherical harmonics ...),
- Fourier transforms.

**Applications:**

- Quantum mechanics (PHYS 304, 402)
- Electromagnetism (PHYS 301, 401)
- Vibrations and Waves, Fluid mechanics, Quantum Field Theory, General Relativity, ...

We will be using the language of **Linear Algebra** to unify the material into a single framework so we can see the patterns more clearly. This is also the language that Quantum Mechanics is written in (quantum states = vectors).
Linear Algebra is in some ways more important than Calculus: it is the basis of most numerical computations, machine learning, image processing, etc...

Some practical details:

- Everything will go through Canvas
- Come to class! We will do a lot of worksheets and similar activities, and there will be a participation component in your grade (10%).
- Please **participate** in class: interrupt me, ask questions, discuss with others.
- There will be weekly HW, due Tuesdays at class time, paper or PDF (15%). The first is due this Tuesday.
- We will have 5 tests (every second Thursday, basically, see Syllabus for dates) and a final exam, for a total of 75%.
- Exam = 3 tests; 5 + 3 = 8, your 6 best count (so each test is 12.5% and the exam is at most 37.5%).

No required textbook. A list of free on-line books, notes and other resources is listed in the Syllabus.

If you cannot come to class, email me in the morning!

Accomodations include a 24 hour window to hand in a completed worksheet via Canvas and a possibility of making a missed. Lectures will not be recorded.

Questions?

The goal of today's lecture is to learn a technique called 'separation of variables', which allows us to take a **partial** differential equation (ie, an equation involving partial derivatives of a function of several variables) and decompose it into several ordinary differential equations (which in principle you know how to solve already).

When this method fails (equation is not separable), this course will have very little to say about the solutions.

We start with the definition of a function.

A function $f$ with domain $X$ and range $Y$ is a rule or a process that associates to every element of $X$ exactly one element of $Y$: $f : X \to Y$.

Example 1: let $X$ and $Y$ be the real numbers. Let $f$ be the rule "for every real number $x \in X$, output half of $x^2$".

This function can be described by a simple formula: $f(x) = \frac{1}{2}x^2$.

Its graph is the set of pairs of numbers $\{(x, y)|y = \frac{1}{2}x^2\}$. When those pairs are represented as points on x-y plane, the result is a parabola (the graph of the function).

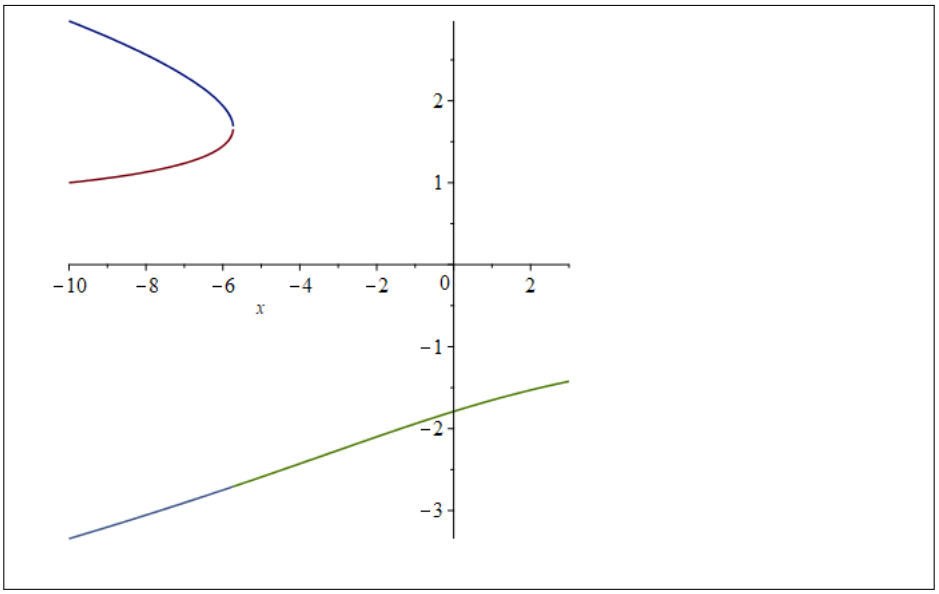
Example 2: Let $X$ and $Y$ be real numbers again. Consider a rule that associated to every number $x$ a solution $y$ of the equation $y^2 - (2 + x^2)y + 1 = 0$. Is this a function?

1. The equation has a solution for every $x$, good. ($\Delta = (2 + x^2)^2 - 4 = 4x^2 + x^4 \geq 0$)

2. For some $x$, there is more than one solution, bad. The rule should not be ambiguous. We will fix it: associate with $x$ the largest solution $y$ to the equation. Now, every $x$ has a corresponding $y$. We can write down a formula, too

$$y = f(x) = \frac{2 + x^2 + \sqrt{4x^2 + x^4}}{2}$$

Example 3: Repeat example 2, but with this equation: $y^5 + xy^3 + 3y^2 - y + 7 = 0$. While this defines a perfectly good function, there is no simple formula we can write down! It is possible to graph it, though.




I could introduce a notation for the function in the last example, say $y = \mathcal{F}(x)$. Here, $\mathcal{F}$ is the function ($\mathcal{F} : \mathbb{R} \to \mathbb{R}$), while $\mathcal{F}(x)$ is the value of the function at $x$.

I can even make it more general and denote the largest (real) root of $y^5+xy^3+ay^2-by+c = 0$ by $\mathcal{F}_{a,b,c}(x)$. This notation stresses that we are considering a large family of functions, $\mathcal{F}_{a,b,c}$ whose domain is the set of real numbers $\mathbb{R}$. Each member of the family is labelled by three real parameters $a$, $b$ and $c$.

Alternatively, I could consider a $\mathcal{G}$ function whose domain is $\mathbb{R}^4$ and write $y = \mathcal{G}(x, a, b, c)$. The line between a parameter and a variable is a matter of definition.

We will define a lot of special functions in this course; it helps to remember that they are just notation for rules that cannot be expressed by a simple formula.

Ok, let's do a worksheet question.

**1.** Consider two functions, $F : \mathbb{R} \to \mathbb{R}$ and $G : \mathbb{R} \to \mathbb{R}$. Given the following properties:

1. for all values of $x$ and $y$, $F(x) = G(y)$ and

2. $F(0) = 3$ ,

find the functions $F$ and $G$.

**2.** What real functions $A$ and $B$ have the property that, for all $x$ and $y$, $A(x) + B(y) = 2$? Make sure you find all of them!

We are now ready for separation of variables. Consider the following differential equation:

$$\frac{\partial}{\partial x}\varphi(x,y) + 2\frac{\partial}{\partial y}\varphi(x,y) = 5\varphi(x,y)$$

To find some solutions of this equation, we will assume that there exist solutions of the form $\varphi(x,y) = f(x)g(y)$. Obviously, that's a big assumption! so maybe we won't find all possible solutions, but maybe we can at least find some. Let's see what happens. Substituting, we get

$$f'(x)g(y) + 2f(x)g'(y) = 5f(x)g(y)$$

Divide both sides by $f(x)g(y)$:

$$\frac{f'(x)}{f(x)} + \frac{2g'(y)}{g(y)} = 5$$

Now, we can think of $\frac{f'(x)}{f(x)} = A(x)$ and $\frac{2g'(y)}{g(y)} = B(y)$. We have an equation similar to the worksheet question 2: $A(x) + B(y) = 5$. We know that $A(x)$ and $B(y)$ must be constant functions! but we don't know what those constants are. Invent some names:

$$\frac{f'(x)}{f(x)} = A(x) = a$$

$$\frac{2g'(x)}{g(x)} = B(x) = b$$

and the original equation becomes $a + b = 5$, or $b = 5 - a$. So, we have two equations to solve:

$$f'(x) = af(x)$$

$$g'(x) = (5-a)g(x)$$

These are ordinary differential equations which you are (supposed to be) able to solve already. We will review that soon.

For now, we want to practice this trick some more. Let's do a slightly more involved question (Laplace equation in 2d in cylindrical coordinates).

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\varphi(r,\theta)\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\varphi(r,\theta) = 0$$

Ansatz: $\varphi(r,\theta) = f(r)g(\theta)$. Substitute:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}f(r)\right)g(\theta) + f(r)\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}g(\theta) = 0$$

Divide by $f(r)g(\theta)$ to try to isolate the variables:

$$\frac{\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}f(r)\right)}{rf(r)} + \frac{1}{r^2}\frac{g''(\theta)}{g(\theta} = 0$$

Following our previous argument, consider the function $G$ with formula $G(\theta) = g''(\theta)/g(\theta)$. Let $G(0) = \lambda$, labelling an unknown value. Consider our above equation at a fixed value of $r$ (say 6) and a different value of $\theta$, say $\theta = 2$. We have that

$$\left.\frac{\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}f(r)\right)}{rf(r)}\right|_6 + \frac{1}{6^2}\lambda = 0$$

$$\left.\frac{\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}f(r)\right)}{rf(r)}\right|_6 + \frac{1}{6^2}G(2) = 0$$

Clearly, we must have $G(2) = G(0) = \lambda$. The argument can be repeated for every value of $\theta$ (not just 2), to show that $G$ must be a constant function, with a formula $G(\theta) = \lambda$. Therefore, the separated equations are

$$g''(\theta) = \lambda g(\theta)$$

and

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}f(r)\right) = -\frac{1}{r^2}\lambda f(r)$$

It's now your turn!

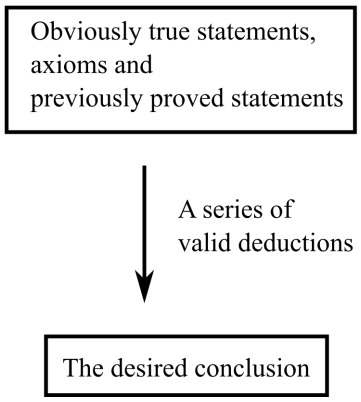
**3.** Separate variables in the following partial differential equation.

$$\frac{\partial^3}{\partial x^3}\varphi(x,y,z) - \frac{\partial}{\partial x}\frac{\partial}{\partial y}\varphi(x,y,z) + \frac{\partial^2}{\partial z^2}\varphi(x,y,z) = 4x^2\varphi(x,y,z)$$

On the first Homework, the last question will asked you to show that a given function is a solution to a differential equation. In other words, you are asked to show that when the function given is substituted into the equation, the equation holds.

Let's clarify what a proof is.

**Proofs 101: What is a proof?**




Example: show that $f(x) = e^{3x}$ solves $f'(x) = 3f(x)$.

Presentation 1: Let $f(x) = e^{3x}$. Clearly, $3e^{3x} = 3f(x)$. Also, $\frac{d}{dx}e^{3x} = 3e^{3x}$ by the usual rules of differentiation. Putting this together, we have

$$f'(x) = \frac{d}{dx}e^{3x} = 3e^{3x} = 3f(x)$$

Therefore, $f'(x) = 3f(x)$ as required.

Presentation 2: Let $f(x) = e^{3x}$. Consider the differential equation $f'(x) = 3f(x)$:
LHS $= f'(x) = 3e^{3x} = 3f(x) =$ RHS $\square$

Hand in your worksheets (for participation grades, we will not grade for correctness).

For HW 1, which is due on Tuesday, you will separate variables in several more equations.

HW 1 is posted on Canvas already.

Lec 2

**1.** Separate the equation

$$\frac{\partial^2}{\partial x^2}u(x,y,z) \; + \; \frac{1}{x}\frac{\partial}{\partial x}\frac{\partial^2}{\partial y^2}u(x,y,z) \; + \; \frac{2}{x^2}\frac{\partial}{\partial z}u(x,y,z) = 0$$

Let's solve these equations one by one. This will cover the most common equations in the course. The first two are linear equations with constant coefficients. The third is a different type of equation, called Cauchy-Euler, or equidimensional equation. It is special because it is homogeneous: every term has the same power of x (derivatives count).

We will use the same approach to both kinds of equations: write down an ansatz (ie, a guess) of the right form with an unknown parameter, substitute this ansatz into the differential equation, simplify to obtain an auxiliary equation, solve auxiliary equation for the parameter.

In general, linear differential equations with constant coefficients are solved by exponentials, the ansatz being of the form $e^{kx}$. For the Cauchy-Euler equation, the ansatz is a power law, $x^p$.

A useful fact about differential equations to remember is first degree equations should have a one parameter family of solutions, second degree equations a two parameter family, etc... We therefore expect one solution to the auxiliary equation for a first order differential equation, two solutions for a second order one, etc...

Let's now solve the three equations in turn.

$$h'(z) = \alpha h(z) \quad \rightarrow \quad h(z) \sim e^{\alpha z}$$

The easiest way to see that is to realize that it's exponential functions that are proportional to their own derivatives, try $e^{kx}$ and get that $k = \alpha$. Notice that this first degree differential equation has a one parameter family of solutions for every $\alpha$: $h(z) = ce^{\alpha z}$, where $c$ is any constant.

It's always good to ask whether there are some special corner cases we are missing. Here, $\alpha = 0$ is special since the solution is then a constant function. This, however, works: the constant function $h(z) = c$ is a solution to the equation $h'(z) = 0$.

$$g''(y) = \beta g(y)$$

This equation has constant coefficients, so we try an exponential: $g(y) = e^{ky}$ (this $k$ is not the same as the previous $k$). Substituting, get $k^2 = \beta$. Looks like we are going to have to consider some cases, because $\beta$ does not need to be a positive number. Keep in mind that we expect two solutions even if $\beta$ is negative!

Case 1: $\beta > 0$, and $k = \pm\sqrt{\beta}$ are the two solutions of the auxiliary equation. Since the differential equation is linear, we can add any two solutions to get new solutions, and we write down the most general solution as $g(y) = c_1 e^{\sqrt{\beta}y} + c_2 e^{-\sqrt{\beta}y}$, clearly a two parameter

family, as expected.

Case 2: $\beta = 0$ (The quadratic equation $k^2 = \beta$ has a double root.) It seems here that we get only one solution, a constant. But that cannot be right, since second order equations must have two solutions. Let's re-examine this from the start: the equation is $g''(y) = 0$, and it has as the most general solution $g(y) = c_1 y + c_2$.

Case 3: $\beta < 0$, $k$ is imaginary: $k = \pm\sqrt{\beta} = \pm\sqrt{-|\beta|} = \pm\sqrt{-1}\sqrt{|\beta|} = \pm i\sqrt{|\beta|} = ib$ where $\beta = -b^2$ and $b$ is positive. Solutions are $g(y) = c_1 e^{iby} + c_2 e^{-iby}$.

When solving physical problems, usually only one of these cases will be relevant given whatever boundary conditions and symmetries have to be applied. But it's good to see all three cases since you never know which one you are going to need.

Looking at Case 3 again, we can ask what is $e^{iby}$? and how come we end up with an imaginary/complex solution to a real problem? does the problem have a real solution? you might remember that there are real solutions to this equation of the form $\sin(ky)$ and $\cos(ky)$. Where are those in this analysis?

To answer all these questions, let's go back and review some things about complex numbers.

## Interlude: complex numbers

Why do we need complex numbers?

When solving real problems, it seems unnecesary to introduce complex numbers at all. But it turns out that they actually simplify (and unify) our thinking.

Let's recall: we have an auxiliary equation $k^2 = \beta$ and we expect two solutions for the original differential equation. It would be convenient, then, to be able to say that '(nearly) all quadratic equations have two solutions'.

This is in fact what complex numbers buy us (and why they were invented). In the field of complex numbers, all polynomial equations have at least one root. And if we count double (and higher) roots as distinct ones, an n-th degree polynomial has n roots.

So, what is the field of complex numbers? it is an extension of the real numbers, where we add a formal new element called $i$ such that $i^2 = i \cdot i = -1$. A complex number has a form $x + iy$ where $x$ and $y$ are real numbers. $x$ is called the real part, $y$ is called the imaginary part. All the standard rules of arithmetic (operations involving +,-,*,/ and integer powers) apply to complex numbers (addition and multiplication are associative and commutative, multiplication distributes over addition, etc...).

Let $z = x + iy$ be a general complex number. Other basic definitions are: the complex conjugate, $\bar{z} = z^* = x - iy$ and the magnitude: $|z| = \sqrt{x^2 + y^2}$, where the positive root is understood.

Let's do some exercises.

Definitions (memorize these)

$$i^2 = -1$$

Let $z = x + iy$. Then,

Complex conjugate:

$$\bar{z} = z^* = x - iy$$

Magnitude:

$$|z| = \sqrt{x^2 + y^2}$$

Note that $|z| = 0$ iff $z = 0$.

**2.**

(a) Show that $z\bar{z} = |z|^2$

(b) Solve the above for $\frac{1}{z}$ and write down a formula for $\frac{1}{z}$ in terms of $x$ and $y$.

(c) Let $\beta$ be a negative real number. Show that $k = \pm i\sqrt{|\beta|}$ are solutions to the equation $k^2 = \beta$. Clearly explain where you assumed that $\beta$ is negative.

We now need some more sophisticated functions of complex numbers. These can be defined from basic arithmetic operations using Taylor expansions. Let's start with

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \ldots \quad = \quad \sum_{k=0}^{\infty} \frac{1}{k!}x^k$$

$$\frac{d}{dx}e^x = \frac{d}{dx}\sum_{k=0}^{\infty}\frac{1}{k!}x^k = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{d}{dx}x^k = \sum_{k=0}^{\infty}\frac{k}{k!}x^{k-1} =$$

$$\sum_{k=1}^{\infty}\frac{1}{(k-1)!}x^{k-1} = \sum_{n=0}^{\infty}\frac{1}{n!}x^n = e^x$$

where on the last step we redefined $n = k - 1$.

$$e^{x+y} = \sum_{n=0}^{\infty}\frac{1}{n!}(x+y)^n = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}x^k y^{n-k} =$$

$$\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{1}{k!j!}x^k y^j = \left(\sum_{k=0}^{\infty}\frac{1}{k!}x^k\right)\left(\sum_{j=0}^{\infty}\frac{1}{j!}y^j\right) = e^x e^y$$

where going from the first to the second line we defined $j = n - k$.

This Taylor expansion definition works for complex numbers just as well. Let $x$ be real (we are not going to differentiate wrt a complex number here) but $\alpha$ be potentially complex. Then

$$e^{\alpha x} = \sum_{k=0}^{\infty} \frac{1}{k!} \alpha^k x^k$$

$$\frac{d}{dx} e^{\alpha x} = \frac{d}{dx} \sum_{k=0}^{\infty} \frac{1}{k!} \alpha^k x^k = \sum_{k=0}^{\infty} \frac{k}{k!} \alpha^k x^{k-1} =$$

$$\alpha \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \alpha^{k-1} x^{k-1} = \alpha \sum_{n=0}^{\infty} \frac{1}{n!} \alpha^n x^n = \alpha e^{\alpha x}$$

Next, as you can imagine, we want to establish a connection to trigonometric functions, and the formula $e^{i\theta} = \cos\theta + i \sin\theta$. There are two different ways to do that. One is to look at the real and imaginary parts of the Taylor expansion of $e^{i\theta}$:

$$e^{i\theta} = 1 + i\theta + \frac{1}{2}(i\theta)^2 + \frac{1}{3!}(i\theta)^3 + \frac{1}{4!}(i\theta)^4 + \dots$$

$$= \left(1 - \frac{1}{2}\theta^2 + \frac{1}{4!}\theta^4 + \dots\right) + i\left(\theta - \frac{1}{3!}\theta^3 + \dots\right)$$

$$= \cos\theta + i \sin\theta$$

or, more formally:

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{1}{k!} i^k \theta^k =$$

$$\sum_{n=0}^{\infty} \frac{1}{(4n)!} i^{4n}\theta^{4n} + \sum_{n=0}^{\infty} \frac{1}{(4n+1)!} i^{4n+1}\theta^{4n+1} +$$

$$\sum_{n=0}^{\infty} \frac{1}{(4n+2)!} i^{4n+2}\theta^{4n+2} + \sum_{n=0}^{\infty} \frac{1}{(4n+3)!} i^{4n+3}\theta^{4n+3} =$$

where we broke the single sum over integers into four sums based on the reminder when k is divided by 4.
Next, we will use that $i^4 = 1$, which implies that $i^{4n} = 1$, $i^{4n+1} = i$, $i^{4n+2} = -1$ and $i^{4n+3} = -i$.

$$= \left( \sum_{n=0}^{\infty} \frac{1}{(4n)!} \theta^{4n} - \sum_{n=0}^{\infty} \frac{1}{(4n+2)!} \theta^{4n+2} \right) +$$

$$i \left( \sum_{n=0}^{\infty} \frac{1}{(4n+1)!} \theta^{4n+1} - \sum_{n=0}^{\infty} \frac{1}{(4n+3)!} \theta^{4n+3} \right) =$$

$$\left( \sum_{m=0}^{\infty} \frac{1}{(2m)!} (-1)^m \theta^{2m} \right) +$$

$$i \left( \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} (-1)^m \theta^{2m+1} \right) =$$

$$\cos\theta + i \sin\theta$$

where we used the Taylor expansions of sin and cos.

A simpler way to think about this relies on the fact that $\cos(\theta)$ is the unique solution of the differential equation $f''(\theta) = -f(\theta)$ with $f(0) = 1$ and $f'(0) = 0$. This is quite instructive, so let's go over it.

**3.** Solve the problem $f''(\theta) = -f(\theta)$, $f(0) = 1$, $f'(0) = 0$ using an exponential ansatz $f(\theta) = e^{k\theta}$.

Now, recall that the unique solution to this problem is $\cos(\theta)$. Therefore, $\cos(\theta) = \frac{e^{i\theta}+e^{-i\theta}}{2}$.

Similarly, we can prove that $\sin(\theta) = \frac{e^{i\theta}-e^{-i\theta}}{2i}$. Then, $\cos\theta + i\sin\theta = e^{i\theta}$.

One particular value of $\theta$ makes a neat little fact: $e^{i\pi} = -1$, or $e^{i\pi} + 1 = 0$. All the most important numbers tied together into one equation.

Equations to memorize:

$$\boxed{e^x = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{3!}x^3 + \ldots}$$

and $\boxed{e^{i\theta} = \cos\theta + i\sin\theta}$

Also: $\boxed{1 + x + x^2 + x^3 + \ldots = \tfrac{1}{1-x}}$

and $\boxed{\ln(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \ldots}$

**end of digression on complex numbers**

So, what about the solutions to our equation? we obtained the following general, complex looking solution to our real differential equation $g''(y) = \beta g(y)$ for $\beta$ negative: $g(y) = c_1 e^{iby} + c_2 e^{-iby}$ with $\beta = -b^2$. Now, use
$e^{iby} = \cos by + i\sin by$ and $e^{-iby} = \cos by - i\sin by$.
We can combine these solutions to get $\frac{1}{2}(e^{iby} + e^{-iby}) = \cos bx$ and $\frac{1}{2i}(e^{iby} - e^{-iby}) = \sin bx$
Oh, good. We recover a different class of solutions to $g''(y) = \beta g(y)$ for negative $\beta$, which are $\sin(by)$ and $\cos(by)$ for $\beta = -b^2$. The general solution is $c_1\sin(by) + c_2\cos(by)$. If we chose $c_1$ and $c_2$ to be real, this solution is clearly real, even though it came from complex solutions.

$$f''(x) + \frac{\beta}{x}f'(x) + \frac{2\alpha}{x^2}f(x) = 0$$

The coefficients in this equation are not constant. This is a different type of equation, called Cauchy-Euler, or equidimensional equation. It is special because it's homogeneous: every term has the same power of $x$ (derivatives count). Solutions are power law: $f(x) = x^p$. Substituting we get the associated polynomial:

$$p(p-1) + \beta p + 2\alpha = 0$$

$$p^2 + (\beta - 1)p + 2\alpha = 0$$

Solutions are given by the quadratic formula:

$$p = \frac{1 - \beta \pm \sqrt{(1-\beta)^2 - 4\alpha}}{2}$$

Again, we have square roots and perhaps we have to analize some cases since for $4\alpha > (1-\beta)^2$, the solution is complex. Let's not do that (since it's messy and not useful from a learning perspective). We do however need to understand what happens when the discriminant is

zero: $4\alpha > (1-\beta)^2$. In this case, we only have one solution $x^p$ and need two (second degree equation). The second solution can be found using the reduction of order method and is $x^p \ln x$. The general solution is $f(x) = (c_1 + c_2 \ln x)x^p$.

To write any specific solution $u(x, y, z)$, we multiply all these together. For example, we have a family of solutions for any positive number $b$ and for any $\alpha < (1+b^2)^2/4$ given by

$$e^{\alpha z}(c_1 \sin(by) + c_2 \cos(by)) \left(c_3 x^{\frac{1}{2}\sqrt{(1+b^2)^2-4\alpha}} + c_4 x^{-\frac{1}{2}\sqrt{(1+b^2)^2-4\alpha}}\right) x^{\frac{1+b^2}{2}}$$

Terribly messy, is it not? physical problems will often turn out to work out a bit nicer.

Often, we will have boundary conditions which will pick out a particular solutions out of the infinite families we get from all the unknown constants from separation of variables.

This will be the topic of Thursday's lecture.

Lec 3

Last lecture we discussed the Taylor expansion definitions of $e^z$ and the connection between $e^ix$, $\cos(x)$ and $\sin(x)$. Let's consider this again, but from perspective of differential equations, which is quite instructive.

$\cos(\theta)$ is the unique solution of the differential equation $f''(\theta) = -f(\theta)$ with $f(0) = 1$ and $f'(0) = 0$. This is due to a general theorem about uniqueness of solutions to differential equations: an $n$-th order ODE with $n$ initial conditions for the function and its first $n-1$ derivatives has a unique solution (at least locally).

**1.** Solve the problem $f''(\theta) = -f(\theta)$, $f(0) = 1$, $f'(0) = 0$ using an exponential ansatz $f(\theta) = e^{k\theta}$.

Now, recall that the unique solution to this problem is $\cos(\theta)$. Therefore, $\cos(\theta) = \frac{e^{i\theta}+e^{-i\theta}}{2}$.

Similarly, we can prove that $\sin(\theta) = \frac{e^{i\theta}-e^{-i\theta}}{2i}$. Then, $\cos\theta + i\sin\theta = e^{i\theta}$. We have obtain this fact previously using Taylor expansions.

One particular value of $\theta$ makes a neat little fact: $e^{i\pi} = -1$, or $e^{i\pi} + 1 = 0$. All the most important numbers tied together into one equation.

## Homogeneous boundary conditions

Often, in addition to the differential equation itself, physics will dictate some sort of restriction along lower dimensional subspace. As an example, the laplace equation in 3-d describes conditions of thermal equilibrium in a heat-conducting material (such as a metal block). At the boundaries of the material (sides of the block), mathematical restrictions representing physical conditions present at the boundary have to be applied. These can take place of, for example, fixing the boundary temperature at a certain value.

We will encounter many types of boundary conditions in this course. They are generally classified as either boundary or initial conditions, an as either homogeneous or inhomogeneous. We will see why these are (mathematically) different types of restrictions that need conceptualy different treatements. In this lecture, we begin with homogeneous boundary conditions. **Homogeneous** here means setting certain values to zero. We will see that this generically leads to restricting the allowed values of separation constants. The whole thing is best understood through an example.

**2.** Find the most general solution to the wave equation

$$c^2\frac{\partial^2}{\partial x^2}u - \frac{\partial^2}{\partial t^2}u = 0$$

for $u(x,t)$ on the region $0 \leq x \leq L$ and $-\infty < t < \infty$, with boundary conditions $u(0,t) = u(L,t) = 0$. Hint: The most general solution is simply a sum of all possible solutions of the form $f(x)g(t)$.

As a first step, make a plan.

1.

2.

3.

4.

Done early? You can get started on HW 2.

$$u(x,t) = \sum_{n=1}^{\infty} \left( b_{n,1} \cos\left(\frac{n\pi ct}{L}\right) + b_{n,2} \sin\left(\frac{n\pi ct}{L}\right) \right) \sin\frac{\pi nx}{L}$$

Looking forward: For concretness, let's take c=1. Let's consider the problem we just solved with inital condition

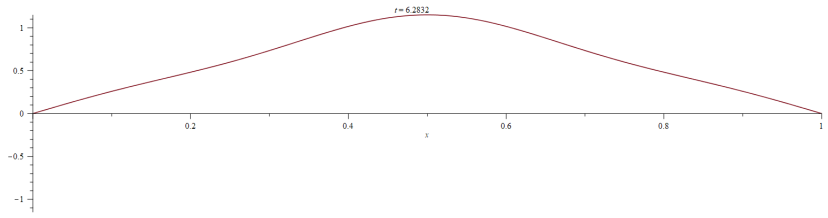
$$u(x,0) = \sin(\pi x) - \frac{1}{9}\sin(3\pi x) + \frac{1}{25}\sin(5\pi x) \ ,$$

and

$$\frac{\partial}{\partial t} u(x,0) = 0 \ ,$$

which is a reasonable approximation to the string being plucked in the middle.

Here is the shape of the string at t=0:




When we solve the wave equation, we obtain the evolution of the string's shape in time. You can see it animated on Canvas.

The second animation on Canvas includes a small exponetial dampening term, changing the general solution to

$$u(x,t) = \sum_{n=1}^{\infty} \left( b_{n,1} \cos\left(\frac{n\pi ct}{L}\right) + b_{n,2} \sin\left(\frac{n\pi ct}{L}\right)\right) \times$$

$$\times \exp\left(\frac{-\alpha n\pi ct}{L}\right) \sin\frac{\pi nx}{L} .$$

$\alpha$ is a small parameter; in the animation, $c = 1$ and $\alpha = 0.015$.

To obtain the time-dependent solution, we must substitute the initial conditions into our general solution and find the coefficients:

$$u(x,0) = \sum_{n=1}^{\infty} b_{n,1} \sin\frac{\pi nx}{L} = \sin(\pi x) - \frac{1}{9}\sin(3\pi x) + \frac{1}{25}\sin(5\pi x)$$

$$\frac{\partial}{\partial t}u(x,0) = \sum_{n=1}^{\infty} b_{n,2}\frac{n\pi c}{L}\sin\frac{\pi nx}{L} = 0$$

The coefficients are, by inspection, $b_{1,1} = 1$, $b_{2,1} = -\frac{1}{9}$ and $b_{3,1} = \frac{1}{25}$, with the rest being zero.

Lec 4

Jupyter Worksheet - read lecture4-worksheet.pdf
Solutions to Worksheet 5 include some solutions to the PDEs in this lecture.

Lec 5

Jupyter Worksheet - read lecture5-worksheet.pdf

Lec 6

The differential equations we have been focusing on are linear (ie, when $u$ is the unknown function, every term in the equation has no more than one factor of $u$ in it, there are no $u^2$ terms, etc...) which is why the principle of supperposition applies to their solutions.

Algebraic linear equations should be familiar to you from Math 221. The theory of linear differential equations has a lot in common with systems of algebraic linear equations. The framework that encompasses both is called Linear Algebra. Linear Algebra is fundamental for lots of physics, most notably Quantum Mechanics. It's also the language of Machine Learning and lots of other modern topics.

We need to start with the basic definitions.

1.18 **Definition** ***addition, scalar multiplication***

- An ***addition*** on a set $V$ is a function that assigns an element $u+v \in V$ to each pair of elements $u, v \in V$.
- A ***scalar multiplication*** on a set $V$ is a function that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbf{F}$ and each $v \in V$.

$\mathbf{F}$ is called a field, which is a math word for 'a bunch of things that can be added, multiplied, divided by, etc...'.
For us, $\mathbf{F}$ is either the set of real numbers $\mathbb{R}$ or (sometimes) complex numbers $\mathbb{C}$.
Reference: This and other definitions are from 'Linear Algebra Done Right' by Sheldon Axler, PDF available for free from Springer for UBC students (see link in the syllabus).

There are two main examples of vectors we will consider, and we want to make sure we understand this definition in both cases:

- Example 1: $V_1$ is the set of d-dimensional column vectors (ie, $\mathbb{R}^d$)
- Example 2: $V_2$ is the set of all real functions $f : \mathbb{R} \to \mathbb{R}$.

In both cases, we have a natural way to add vectors and a natural way to multiply a vector by a number (a scalar).

For $V$ to be a vector space, the addition and scalar multiplication must have some particular properties:

1.19 **Definition** ***vector space***

A ***vector space*** is a set $V$ along with an addition on $V$ and a scalar multiplication on $V$ such that the following properties hold:

**commutativity**
$u + v = v + u$ for all $u, v \in V$;

**associativity**
$(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and all $a, b \in \mathbf{F}$;

continued on next slide...

**Vector space** part two:

**additive identity**
there exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$;

**additive inverse**
for every $v \in V$, there exists $w \in V$ such that $v + w = 0$;

**multiplicative identity**
$1v = v$ for all $v \in V$;

**distributive properties**
$a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b \in \mathbf{F}$ and all $u, v \in V$.

Certain 'obvious' things are easily proven. The additive identity is unique, because if there were two such identities, say $0 \in V$ and $\tilde{0} \in V$, then

$$0 = 0 + \tilde{0} = \tilde{0} + 0 = \tilde{0} \ .$$

The two identities turn out to be the same!

The addivite inverse is unique, because if there were two such inverses $w$ and $\tilde{w}$, with $v + w = 0$ and $v + \tilde{w} = 0$, then

$$w = w + 0 = w + (v + \tilde{w}) = (w + v) + \tilde{w} = 0 + \tilde{w} = \tilde{w}$$

The two inverses turn out to be the same! The usual notation is for the additive inverse of $v$ is $-v$.

Moreover, when any vector $v$ is multiplied by the number 0, the result is the (unique!) additive identity $0 \in V$. Proof: for all $v$ in $V$, $0v + 0v = (0 + 0)v = 0v$. Now, add the

negative inverse of $0v$, denoted with $w_0$, to both sides: $w_0+0v+0v = w_0+0v \;\Rightarrow\; 0+0v = 0 \;\Rightarrow\; 0v = 0$.

Finaly, we can show that $(-1)v = -v$, by computing

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0 \ .$$

This gives you a taste of how this object is studied abstractly. For more, check out the textbook.

It turns out that we won't have to check all these properties every time though, because most of the vector spaces we will need to construct will 'sit inside' another vector space for which we already know that these properties work. You met this approach in Math 221 when you defined a vector space as a subset of $V_1 = \mathbb{R}^d$.

1.32 **Definition** ***subspace***

A subset $U$ of $V$ is called a ***subspace*** of $V$ if $U$ is also a vector space (using the same addition and scalar multiplication as on $V$).

Any subset $U$ of a vector space $V$ is guaranteed to be a subspace as long as it is **nonempty**, and addition and scalar multiplication are still well defined. This basicaly means that the set $U$ is:

- closed under scalar multiplication:
  $au \in U$ if $a \in F$ and $u \in U$
- closed under addition:
  $u + \tilde{u} \in U$ if $u \in U$ and $\tilde{u} \in U$

We will take the same approach for $V_2$: set of real functions of one real variable. A subset of these functions satisfying the above closure conditions is also a vector space.

Examples: Set of functions which are 0 when evaluated at 0, $\{f : \mathbb{R} \to \mathbb{R} \mid f(0) = 0\}$, is a vector subspace.

Set of functions which are 1 at 0, $\{f : \mathbb{R} \to \mathbb{R} \mid f(0) = 1\}$, is not.

Set of even functions is a subspace, as is the set of odd functions. The subspaces of even and odd functions intersect at exactly one vector: the zero function.

Worksheet Question: **1.** In the list below, all but five sets are examples of vector spaces. Spot the five sets that are not vector spaces.

(A) Functions of one real variable that are continuous

(B) Functions of one real variable that are differentiable everywhere

(C) Functions of one real variable that have a continuous derivative

(D) Functions of one real variable that are bounded (ie, for each function $f(x)$ there is a number $C$ such that $|f(x)| < C$ for any $x$).

(E) Functions of one real variable that have a limit of 0 for $x \to \pm\infty$.

(F) Functions of one real variable that go to zero faster than $x^{-1}$ for large $x$ (ie, for each function $f(x)$ there are positive numbers $C$ and $x_0$ such that $|f(x)| < |Cx^{-1}|$ for any $x$ with $|x| > x_0$). Can substitute any function that asymptotes to 0 for $x^{-1}$, such as $x^{-2}$, $e^{-|x|}$ or $1/\ln|x|$.

(G) Functions of one real variable that grow faster than $x^2$ for large $x$ (ie, for each function $f(x)$ there are positive numbers $C$ and $x_0$ such that $|f(x)| > Cx^2$ for any $x$ with $|x| > x_0$).

(H) Functions $f$ of one real variable $x$ such that $f(0) = 0$

(I) Functions $f$ of one real variable $x$ such that $f(1) = 0$

(J) Functions $f$ of one real variable $x$ such that $f(0) = 1$

(K) Functions $f$ of one real variable $x$ such that $f(1) = 0$, $f(2) = 0$, $f(4) = 0$.

(L) Functions $f$ of one real variable $x$ such that $\lim_{x\to 0}|f(x)| = \infty$.

(M) Periodic functions: functions $f$ of one real variable $x$ such that $f(x) = f(x + 2\pi)$

(N) A subspace of the last space: functions $f$ of one real variable $x$ such that $f(x) = f(x + \pi)$. This can be continued with smaller and smaller periods.

(O) Even/odd functions of one real variable

(P) Positive/negative functions of one real variable

(Q) Increasing/decreasing functions of one real variable

(R) Functions of $x$ of the form $\sum_n^N c_n x^n$, where $c_n$ are real numbers.

(S) Polynomial functions

(T) Functions defined on the interval $[0, 1]$

(U) Functions defined on integer numbers only

(V) Functions defined on the unit sphere

(W) Functions of two or three variables

(X) Functions of three variables $(x, y, z)$ that vanish on evaluated on points lying on the sphere $x^2 + y^2 + z^2 = 1$.

(Y) Functions of three variables $(x, y, z)$ that vanish when evaluated on points with $x = y = 3$.

(Z) Vector fields (in three dimensions, for example, a vector field is a function that assigns a 3-vector to every point in space).

In math 221, you wrote systems of linear equations in matrix form, as

$$Ax = b \ .$$

You then learned how to solve these equations in various cases, including with $A$ invertible ($x = A^{-1}b$) and non-invertible (more complicated).

We want to think of linear differential equations as equations of the form $Ax = b$. We want $x$ and $b$ to be elements of some vector space of functions. What should $A$ be then?

The generalization of a matrix to the abstract vector space is called a linear map:

> 3.2 **Definition** ***linear map***
>
> A ***linear map*** from $V$ to $W$ is a function $T : V \to W$ with the following properties:
>
> **additivity**
> $T(u + v) = Tu + Tv$ for all $u, v \in V$;
>
> **homogeneity**
> $T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$.
>
> A linear map is a function from one vector space to another that respects the structure of the vector spaces.

Let's consider some examples, again. We might take $L : V_1 \to V_2$ which defines an arrow in $V_2$ using the top two numbers of a vector in $V_1$ as cartesian coordinates of the arrow (and simply ignores the bottom number).

This map is clearly linear. However, if we change it a bit so that the two top numbers define an arrow using cylindical coordinates: r and $\theta$, respectively, the map is no longer linear. (It's still a function, but it's not linear.).

Another worksheet question: **2.** Consider these two vector spaces:

- $V_1$ is the set of 3-dimensional column vectors;
- $V_2$ is the set of all functions $f : \mathbb{R} \to \mathbb{R}$.

Let's define three maps from $V_1$ to $V_2$, by

$$M_1 : V_1 \to V_2 , \quad M_1 \begin{bmatrix} a \\ b \\ c \end{bmatrix} = ax^b + c$$

$$M_2 : V_1 \to V_2 , \quad M_2 \begin{bmatrix} a \\ b \\ c \end{bmatrix} = ax^2 + bx + c$$

$$M_3 : \ V_1 \ \rightarrow \ V_2 \ , \quad M_3 \begin{bmatrix} a \\ b \\ c \end{bmatrix} = ax + 4$$

Which of these are linear maps?

**3.** Since this is a calculus course, of special interests are linear maps from a vector space of functions (a function space) to another function space. Give some examples of linear maps $L : \ V_2 \ \rightarrow V_2$.

**4.** Consider a differential equation:

$$\frac{\partial u}{\partial t} = 2\frac{\partial^2 u}{\partial k^2}$$

with boundary conditions:

$$u(0,t) = u(\pi,t) = 0$$

Write this differential equation in the form $Ax = b$. Identify: the linear map $A$ and the vector space that $x$ and $b$ live in.

A linear map is a function from one vector space to another that respects the structure of the vector spaces. It can also be called an **operator** (especially in Quantum Mechanics).

3.2 **Definition** ***linear map***

A ***linear map*** from $V$ to $W$ is a function $T : V \rightarrow W$ with the following properties:

**additivity**
$T(u + v) = Tu + Tv$ for all $u, v \in V$;

**homogeneity**
$T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$.

One key property of all linear maps $T$ is that $T0 = 0$ (0 here is a vector). The proof goes as follows: Take any vector $v$. Then $T0 = T(0v) = 0T(v) = 0$. $\square$

We have already seen that operations such as differentiation are linear maps. Thil allow us to think about differential equations in the same way as matrix equations. For example

$$\frac{\partial^2}{\partial x^2}u(x,y) + \frac{\partial^2}{\partial y^2}u(x,y) = x^2 u(x,y) \quad \rightarrow \quad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2\right) u(x,y) = 0$$

so we can think of the differential equation as an equation of the form $Au = 0$, a homogenous version of the general form $Au = b$.

An extra term can result in an inhomogenous equation:

$$\frac{\partial^2}{\partial x^2}u(x,y) + \frac{\partial^2}{\partial y^2}u(x,y) = x^2 u(x,y) + \cos y \quad \rightarrow \quad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2\right) u(x,y) = \cos y$$

Now we have an equation of the form $Au = b$, with $A = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2$ and $b = \cos y$.

For now, we will stick to homogenous equations, with $b = 0$.

Lec 7

Recall from last lecture:

A linear map is a function from one vector space to another that respects the structure of the vector spaces. It can also be called an **operator** (especially in Quantum Mechanics).

3.2 **Definition** ***linear map***

A ***linear map*** from $V$ to $W$ is a function $T : V \to W$ with the following properties:

**additivity**
$T(u + v) = Tu + Tv$ for all $u, v \in V$;

**homogeneity**
$T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$.

We have already seen that operations such as differentiation are linear maps. This will allow us to think about differential equations in the same way as matrix equations. For example

$$\frac{\partial^2}{\partial x^2}u(x,y) + \frac{\partial^2}{\partial y^2}u(x,y) = x^2 u(x,y) \quad \to \quad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2\right) u(x,y) = 0$$

so we can think of the differential equation as an equation of the form $Au = 0$, a homogenous version of the general form $Au = b$.

An extra term can result in an inhomogenous equation:

$$\frac{\partial^2}{\partial x^2}u(x,y) + \frac{\partial^2}{\partial y^2}u(x,y) = x^2 u(x,y) + \cos y \quad \to \quad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2\right) u(x,y) = \cos y$$

Now we have an equation of the form $Au = b$, with $A = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2$ and $b = \cos y$.

Consider a linear operator (map) $A : V \to W$. The set of vectors $k$ in $V$ such that $Ak = 0$ is called the kernel of $A$.

Theorem: the kernel of any linear map $A$ is a subspace of $V$, and therefore a vector space.

Proof: We discussed that for a set to be a subspace, all we need is closure under vector addition and scalar multiplication. Therefore, take any two vectors $k_1$ and $k_2$ in the kernel of $A$: $Ak_1 = 0$ and $Ak_2 = 0$. Then, $0 = Ak_1 + Ak_2 = A(k_1 + k_2)$, which implies that $k_1 + k_2$ is in the kernel (closure under vector addition). Further, let $\lambda$ be any number (real or complex, as is appropriate), then $0 = \lambda(Ak_1) = A(\lambda k_1)$, so that $\lambda k_1$ is in the kernel (closure under scalar multiplication). $\square$

We have been using this fact already, as the principle of linear superposition. The above proof shows you that the principle of superposition applies to any equation of the form $Ax = 0$ where $x$ is a vector and $A$ any linear map. Such equations are called linear and homogenous. Linear inhomogenous equations are $Ax = b$, and we will return to those later. However, recall than the most general solution to $Ax = b$ is of the form $x_p + x_h$ where $x_p$ is any particular solution to $Ax = b$ nad $x_h$ is the general solution to the corresponding homogeneous equation $Ax = 0$.

When we separate variables, we often end up with equations of the form

$$Au = \lambda u \ .$$

This is is not a linear equation if $\lambda$ is some unknown parameter, since the two unknows $\lambda$ and $u$ multiply each other. It is called an eigenvalue equation (or problem). We are looking for eigenvectors of the operator (linear map) $A$, ie vectors which do not change direction when $A$ acts on them. Instead, these vectors are only rescaled, by the corresponding eigenvalue $\lambda$.

Notice that separation variables just about forced us into considering these eigenvalue problems. This will turn out to be a blessing in disguise.

The boundary problems we have been solving so far had two kinds of boundary conditions: homogeneous ones (something was set to zero) and inhomogeneous ones. The homogeneous boundary conditions specify a vector subspace (of the space of all functions) that we want to solve our differential equation in. The inhomogeneous conditions should be applied at the end of the solution and can be used to determine undefined coefficients. So far, the exact form of these has been chosen so that the coefficients can be more-or-less determined by inspection, because the functions chosen were already sums of solutions to the separated equations.

For example, consider the Laplace problem $\nabla^2 u(r, \theta) = 0$ in polar coordinates, with $u(1, \theta) = \cos 2\theta$. We separated variables $u(r, \theta) = f(\theta)g(r)$ and obtained an eigenvalue equation $f''(\theta) = \lambda f(\theta)$. There is an implicit boundary condition which is that $f$ is periodic, $f(\theta) = f(\theta + 2\pi)$. This defines the particular vectors space in which we solve the eigenvalue equation $\partial_\theta^2 f(\theta) = f''(\theta) = \lambda f(\theta)$. The solutions, as we already know, are the eigenvectors $f_n = e^{in\theta}$ and the corresponding eigenvalues $\lambda_n = -n^2$, for any integer $n$. Recall further that when trying to match to boundary conditions, you wrote $u(1, \theta) = \cos(2\theta) = \frac{1}{2}e^{i\theta} + \frac{1}{2}e^{-i\theta}$, ie, you used $f_n(\theta)$ as a basis in which to write the function $\cos 2\theta$.

But what if the boundary condition was, say $u(r, \theta) = cos(cos(\theta))$? It is not immediately obvious how to write this down in terms of $f_n = e^{in\theta}$. It is not immediately obvious that

this is possible at all. Figuring this sort of thing out is where we are going.

Let's take a little detour into questions about finite dimensional (column) vectors.
**1.** Consider the following vectors, with $a, b$ some real numbers:

$$f_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \ , \qquad f_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \ , \qquad v = \begin{bmatrix} a \\ b \end{bmatrix}$$

Write $v$ in the basis $\{f_1, f_2\}$.

**2.** Redo the question above, with

$$f_1 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \ , \qquad f_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} \ , \qquad v = \begin{bmatrix} a \\ b \end{bmatrix}$$

Write $v$ in the basis $\{f_1, f_2\}$.

The key property in the above solution was orthogonality. Let's see how this would work more generaly, in $d$ dimensions.

Let $\{f_1, f_2, \ldots, f_d\}$ be $d$ orthogonal vectors in a $d$-dimensional space: $f_k \cdot f_j = 0$ for any $k \neq j$. What coefficients $c_n$ give you $\sum_{n=1}^{d} c_n f_n = v$?

Answer: Let's start with a piece of useful notation, the Kronecker delta symbol, which is a function of two integers, defined by

$$\delta_{kn} = \begin{cases} 1 & \text{when } k = n \\ 0 & \text{when } k \neq n \end{cases}$$

Then, orthogonality can be written as

$$f_k \cdot f_n = \begin{cases} f_n \cdot f_n & \text{when } k = n \\ 0 & \text{when } k \neq n \end{cases} \quad = \quad (f_n \cdot f_n) \begin{cases} 1 & \text{when } k = n \\ 0 & \text{when } k \neq n \end{cases} = \delta_{kn} \ (f_n \cdot f_n) = \delta_{kn} \ \|f_n\|^2$$

We are trying to solve

$$\sum_{n=1}^{d} c_n f_n = v$$

for the coefficients $c_n$. Project this equation onto the direction of $f_k$ by taking a dot product on both sides:

$$f_k \cdot \left( \sum_{n=1}^{d} c_n f_n \right) = f_k \cdot v$$

$$\sum_{n=1}^{d} c_n (f_k \cdot f_n) = f_k \cdot v$$

$$\sum_{n=1}^{d} c_n \|f_n\|^2 \delta_{kn} = f_k \cdot v$$

The terms is the sum are all zero, except when $n = k$, so $\sum_{n=1}^{d} c_n \|f_n\|^2 \delta_{kn} = c_k \|f_k\|^2$. Therefore,

$$c_k \|f_k\|^2 = f_k \cdot v \qquad \rightarrow \qquad c_k = \frac{f_k \cdot v}{\|f_k\|^2}$$

This is exactly what we got in the worksheet question (except there, the basis vectors were length 1, so the denominator was 1).

We see that orthogonality of a basis can be a great help with finding coefficients. It will turn out that eigenvectors (eigenfuncions) of differential operators will be orthogonal as long as we pick the right generalization of the dot product. Let's talk about how this is accomplished.

Define the inner product, which is a generalization of the dot product: a way to multiply two vectors together to get a number.

**6.3 Definition** ***inner product***

An ***inner product*** on $V$ is a function that takes each ordered pair $(u, v)$ of elements of $V$ to a number $\langle u, v \rangle \in \mathbf{F}$ and has the following properties:

**positivity**
$\langle v, v \rangle \geq 0$ for all $v \in V$;

**definiteness**
$\langle v, v \rangle = 0$ if and only if $v = 0$;

*Although most mathematicians define an inner product as above, many physicists use a definition that requires homogeneity in the second slot instead of the first slot.*

second
**additivity in ~~first~~ slot** <w,u+v> = <w,u> + <w,v>
~~$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$~~ for all $u, v, w \in V$;

second
**homogeneity in ~~first~~ slot**
~~$\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$~~ for all $\lambda \in \mathbf{F}$ and all $u, v \in V$;
<v,λu> = λ <v,u>

**conjugate symmetry**
$\langle u, v \rangle = \overline{\langle v, u \rangle}$ for all $u, v \in V$.

**6.7 Basic properties of an inner product**

(a) For each fixed $u \in V$, the function that takes $v$ to $\langle v, u \rangle$ is a linear map from $V$ to $\mathbf{F}$.

(b) $\langle 0, u \rangle = 0$ for every $u \in V$.

(c) $\langle u, 0 \rangle = 0$ for every $u \in V$.

(d) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$ for all $u, v, w \in V$.

(e) $\langle u, \lambda v \rangle = \bar{\lambda} \langle u, v \rangle$ for all $\lambda \in \mathbf{F}$ and $u, v \in V$.

Of course, we will replace:
(d) above with $\langle v + w, u \rangle = \langle v, u \rangle + \langle w, u \rangle$
and (e) above with $\langle \lambda v, u \rangle = \bar{\lambda} \langle v, u \rangle$

**6.8 Definition** ***norm,*** $\|v\|$

For $v \in V$, the ***norm*** of $v$, denoted $\|v\|$, is defined by

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

**6.10 Basic properties of the norm**

Suppose $v \in V$.

(a) $\|v\| = 0$ if and only if $v = 0$.

(b) $\|\lambda v\| = |\lambda|\,\|v\|$ for all $\lambda \in \mathbf{F}$.

**6.11 Definition** ***orthogonal***

Two vectors $u, v \in V$ are called ***orthogonal*** if $\langle u, v\rangle = 0$.

**6.13 Pythagorean Theorem**

Suppose $u$ and $v$ are orthogonal vectors in $V$. Then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

**6.15 Cauchy–Schwarz Inequality**

Suppose $u, v \in V$. Then

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|.$$

This inequality is an equality if and only if one of $u, v$ is a scalar multiple of the other.

**6.18 Triangle Inequality**

Suppose $u, v \in V$. Then

$$\|u + v\| \leq \|u\| + \|v\|.$$

This inequality is an equality if and only if one of $u, v$ is a nonnegative multiple of the other.

Before we move on, we should have at least one example of an inner product on a function space. Consider the space of continuous real functions on the interval $[0, 1]$. This is a real vector space and we define an inner product via

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx = \int_0^1 dx\ f(x)g(x)$$

and the norm is then

$$\|f\|^2 = \int_0^1 (f(x))^2 dx\ .$$

If the functions are complex, we have instead:

$$\langle f, g\rangle = \int_0^1 dx\ \overline{f(x)}g(x)dx\ .$$

The complex conjugate on $f(x)$ above is crucial for the whole thing to work, in particular in means the norm is always real and non-negative:

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 dx\ \overline{f(x)}f(x)dx = \int_0^1 dx\ |f(x)|^2$$

You can check for yourself that these definitions have the required properties.

Now comes one of the foundational arguments of all physics, which links a particular property of a linear operator to orthogonality of its eigenvectors.

Consider, in a real vector space, two eigenvectors of the same operator, with **different** eigenvalues

$$Av_1 = \lambda_1 \quad \text{and} \quad Av_2 = \lambda_2 \quad \text{with} \ \ \lambda_1 \neq \lambda_2$$

Then:

$$\langle v_2, Av_1\rangle = \langle v_2, \lambda_1 v_1\rangle = \lambda_1\langle v_2, v_1\rangle$$

and

$$\langle Av_2, v_1\rangle = \langle \lambda_2 v_2, v_1\rangle = \lambda_2\langle v_2, v_1\rangle$$

Therefore,

$$(\lambda_1 - \lambda_2)\langle v_2, v_1\rangle = \langle v_2, Av_1\rangle - \langle Av_2, v_1\rangle$$

If the RHS vanishes, then we must have $\langle v_2, v_1\rangle = 0$ (since $\lambda_1 - \lambda_2 \neq 0$). This motivates a definition and serves as a proof of a theorem:

**Definition** *symmetric* A linear operator $A$ on a real inner product space is called symmetric iff $\langle v_2, Av_1\rangle = \langle Av_2, v_1\rangle$ for any two vectors $v_1$ and $v_2$.

**Theorem** On a real vector space, eigenvectors of a symmetric operator with distinct eigenvalues are orthogonal.

**Corollary** The set of eigenvectors of a symmetric operator on a real vector space can always be made orthogonal.

As an example, consider a finite dimensional real space. The vectors can be presented as column vectors and linear operators as matrices. Taking the inner product to be the regular dot product, $\langle v, w\rangle = v\dot{w} = v^T w$. Consider a symmetric operator $A$,

$$\langle v, Aw\rangle = \langle Av, w\rangle$$

$$v^T Aw = (Av)^T w = v^T A^T w$$

where the last step comes from properties of transpose, $(AB)^T = B^T A^T$. Since $v^T Aw = v^T A^T w$ is true for all $v$ and $w$, we must have that $A = A^T$, so the matrix $A$ is symmetric.

What if the vector space is complex rather than real? An operator with $\langle v_2, Av_1\rangle = \langle Av_2, v_1\rangle$ is then called Hermitian. Immediately, if $v$ is an eigenvector of $A$ with $Av = \lambda v$, we get that

$$\langle v, Av\rangle = \langle Av, v\rangle$$

$$\langle v, \lambda v\rangle = \langle \lambda v, v\rangle$$

$$\bar{\lambda}\langle v, v\rangle = \lambda\langle v, v\rangle$$

$$\bar{\lambda} = \lambda$$

The eigenvalues are real.

Let's re-examine the proof above for a complex vector space: Consider, in a complex vector space, two eigenvectors of the same **hermitian** operator, with **different** eigenvalues

$$Av_1 = \lambda_1 \quad \text{and} \quad Av_2 = \lambda_2 \quad \text{with} \quad \lambda_1 \neq \lambda_2$$

Then:

$$\langle v_2, Av_1\rangle = \langle v_2, \lambda_1 v_1\rangle = \lambda_1\langle v_2, v_1\rangle$$

and

$$\langle Av_2, v_1\rangle = \langle \lambda_2 v_2, v_1\rangle = \bar{\lambda}_2\langle v_2, v_1\rangle = \lambda_2\langle v_2, v_1\rangle$$

Therefore,

$$(\lambda_1 - \lambda_2)\langle v_2, v_1\rangle = \langle v_2, Av_1\rangle - \langle Av_2, v_1\rangle = 0$$

for a hermitian $A$, implying that $\langle v_2, v_1\rangle = 0$ (since $\lambda_1 - \lambda_2 \neq 0$).

**Definition** *hermitian* A linear operator $A$ on a complex inner product space is called hermitian iff $\langle v_2, Av_1\rangle = \langle Av_2, v_1\rangle$ for any two vectors $v_1$ and $v_2$.

**Theorem** Eigenvalues of a hermitian operator are real.

**Theorem** On a complex vector space, eigenvectors of a hermitian operator with distinct eigenvalues are orthogonal.

**Corollary** The set of eigenvectors of a hermitian operator on a complex vector space can always be made orthogonal.

**3. (Time Permitting)** Consider $\frac{d^2}{dx^2}$ acting on real functions on the interval $[0, L]$ whose first derivative vanishes at the ends of the interval (so called free boundary conditions). Show that this operator (second derivative) is symmetric with the inner product $\langle f, g\rangle = \int_0^L dx f(x)g(x)$.

Lec 8

Recall from last lecture:

A linear map is a function from one vector space to another that respects the structure of the vector spaces. It can also be called an **operator** (especially in Quantum Mechanics).

3.2 **Definition** ***linear map***

A ***linear map*** from $V$ to $W$ is a function $T : V \to W$ with the following properties:

**additivity**
$T(u + v) = Tu + Tv$ for all $u, v \in V$;

**homogeneity**
$T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$.

We have already seen that operations such as differentiation are linear maps. Thil allow us to think about differential equations in the same way as matrix equations. For example

$$\frac{\partial^2}{\partial x^2}u(x,y) + \frac{\partial^2}{\partial y^2}u(x,y) = x^2 u(x,y) \quad \to \quad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2\right) u(x,y) = 0$$

so we can think of the differential equation as an equation of the form $Au = 0$, a homogenous version of the general form $Au = b$.

An extra term can result in an inhomogenous equation:

$$\frac{\partial^2}{\partial x^2}u(x,y) + \frac{\partial^2}{\partial y^2}u(x,y) = x^2 u(x,y) + \cos y \quad \to \quad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2\right) u(x,y) = \cos y$$

Now we have an equation of the form $Au = b$, with $A = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - x^2$ and $b = \cos y$.

Recall that when solving (finite dimentional) linear equations of the form $Af = b$, we found it very useful to know the solution to the eigenvalue equation $Af = \lambda f$. The same will be true here. Also, recall that we get equations of the form $Af = \lambda f$ from separation of variables. These two things together are the motivation for thinking about eigenvalue problems for function spaces.

The key property we want to understand is orthogonality, so we need to first generalize the concept of the inner product:

Define the inner product, which is a generalization of the dot product: a way to multiply two vectors together to get a number.

**6.3 Definition** ***inner product***

An ***inner product*** on $V$ is a function that takes each ordered pair $(u, v)$ of elements of $V$ to a number $\langle u, v\rangle \in \mathbf{F}$ and has the following properties:

**positivity**
$\langle v, v\rangle \geq 0$ for all $v \in V$;

**definiteness**
$\langle v, v\rangle = 0$ if and only if $v = 0$;

**additivity in ~~first~~ second slot** <w,u+v> = <w,u> + <w,v>
~~$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$~~ for all $u, v, w \in V$;

**homogeneity in ~~first~~ second slot**
~~$\langle \lambda u, v\rangle = \lambda \langle u, v\rangle$~~ for all $\lambda \in \mathbf{F}$ and all $u, v \in V$;
<v,λu> = λ <v,u>

**conjugate symmetry**
$\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

*Although most mathematicians define an inner product as above, many physicists use a definition that requires homogeneity in the second slot instead of the first slot.*

**6.7 Basic properties of an inner product**

(a) For each fixed $u \in V$, the function that takes $v$ to $\langle v, u\rangle$ is a linear map from $V$ to $\mathbf{F}$.

(b) $\langle 0, u\rangle = 0$ for every $u \in V$.

(c) $\langle u, 0\rangle = 0$ for every $u \in V$.

(d) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for all $u, v, w \in V$.

(e) $\langle u, \lambda v\rangle = \bar{\lambda}\langle u, v\rangle$ for all $\lambda \in \mathbf{F}$ and $u, v \in V$.

Of course, we will replace:
(d) above with $\langle v + w, u\rangle = \langle v, u\rangle + \langle w, u\rangle$
and (e) above with $\langle \lambda v, u\rangle = \bar{\lambda}\langle v, u\rangle$

**6.8 Definition** ***norm,*** $\|v\|$

For $v \in V$, the ***norm*** of $v$, denoted $\|v\|$, is defined by

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

**6.10 Basic properties of the norm**

Suppose $v \in V$.

(a) $\|v\| = 0$ if and only if $v = 0$.

(b) $\|\lambda v\| = |\lambda|\,\|v\|$ for all $\lambda \in \mathbf{F}$.

**6.11 Definition** ***orthogonal***

Two vectors $u, v \in V$ are called ***orthogonal*** if $\langle u, v\rangle = 0$.

**6.13 Pythagorean Theorem**

Suppose $u$ and $v$ are orthogonal vectors in $V$. Then

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

**6.15 Cauchy–Schwarz Inequality**

Suppose $u, v \in V$. Then

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|.$$

This inequality is an equality if and only if one of $u, v$ is a scalar multiple of the other.

**6.18 Triangle Inequality**

Suppose $u, v \in V$. Then

$$\|u+v\| \leq \|u\| + \|v\|.$$

This inequality is an equality if and only if one of $u, v$ is a nonnegative multiple of the other.

Before we move on, we should have at least one example of an inner product on a function space. Consider the space of continuous real functions on the interval $[0, 1]$. This is a real vector space and we define an inner product via

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx = \int_0^1 dx\ f(x)g(x)$$

and the norm is then

$$\|f\|^2 = \int_0^1 (f(x))^2 dx\ .$$

If the functions are complex, we have instead:

$$\langle f, g\rangle = \int_0^1 dx\ \overline{f(x)}g(x)dx\ .$$

The complex conjugate on $f(x)$ above is crucial for the whole thing to work, in particular in

means the norm is always real and non-negative:

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 dx\ \overline{f(x)} f(x) dx = \int_0^1 dx\ |f(x)|^2$$

You can check for yourself that these definitions have the required properties.

Now comes one of the foundational arguments of all physics: Consider, in a real vector space, two eigenvectors of the same operator, with **different** eigenvalues

$$Av_1 = \lambda_1 \quad \text{and} \quad Av_2 = \lambda_2 \quad \text{with} \quad \lambda_1 \neq \lambda_2$$

Then:

$$\langle v_2, Av_1\rangle = \langle v_2, \lambda_1 v_1\rangle = \lambda_1 \langle v_2, v_1\rangle$$

and

$$\langle Av_2, v_1\rangle = \langle \lambda_2 v_2, v_1\rangle = \lambda_2 \langle v_2, v_1\rangle$$

Therefore,

$$(\lambda_1 - \lambda_2)\langle v_2, v_1\rangle = \langle v_2, Av_1\rangle - \langle Av_2, v_1\rangle$$

If the RHS vanishes, then we must have $\langle v_2, v_1\rangle = 0$ (since $\lambda_1 - \lambda_2 \neq 0$). This motivates a definition and serves as a proof of a theorem:

**Definition** *symmetric* A linear operator $A$ on a real inner product space is called symmetric iff $\langle v_2, Av_1\rangle = \langle Av_2, v_1\rangle$ for any two vectors $v_1$ and $v_2$.

**Theorem** On a real vector space, eigenvectors of a symmetric operator with distinct eigenvalues are orthogonal.

**Corollary** The set of eigenvectors of a symmetric operator on a real vector space can always be made orthogonal.

As an example, consider a finite dimensional real space. The vectors can be presented as column vectors and linear operators as matrices. Taking the inner product to be the regular dot product, $\langle v, w\rangle = v\dot{w} = v^T w$. Consider a symmetric operator $A$,

$$\langle v, Aw\rangle = \langle Av, w\rangle$$

$$v^T A w = (Av)^T w = v^T A^T w$$

where the last step comes from properties of transpose, $(AB)^T = B^T A^T$. Since $v^T A w =$

$v^T A^T w$ is true for all $v$ and $w$, we must have that $A = A^T$, so the matrix $A$ is symmetric.

What if the vector space is complex rather than real? An operator with $\langle v_2, Av_1\rangle = \langle Av_2, v_1\rangle$ is then called Hermitian. Immediately, if $v$ is an eigenvector of $A$ with $Av = \lambda v$, we get that

$$\langle v, Av\rangle = \langle Av, v\rangle$$

$$\langle v, \lambda v\rangle = \langle \lambda v, v\rangle$$

$$\bar{\lambda}\langle v, v\rangle = \lambda\langle v, v\rangle$$

$$\bar{\lambda} = \lambda$$

The eigenvalues are real.

Let's re-examine the proof above for a complex vector space: Consider, in a complex vector space, two eigenvectors of the same **hermitian** operator, with **different** eigenvalues

$$Av_1 = \lambda_1 \quad \text{and} \quad Av_2 = \lambda_2 \quad \text{with} \quad \lambda_1 \neq \lambda_2$$

Then:

$$\langle v_2, Av_1\rangle = \langle v_2, \lambda_1 v_1\rangle = \lambda_1\langle v_2, v_1\rangle$$

and

$$\langle Av_2, v_1\rangle = \langle \lambda_2 v_2, v_1\rangle = \bar{\lambda}_2\langle v_2, v_1\rangle = \lambda_2\langle v_2, v_1\rangle$$

Therefore,

$$(\lambda_1 - \lambda_2)\langle v_2, v_1\rangle = \langle v_2, Av_1\rangle - \langle Av_2, v_1\rangle = 0$$

for a hermitian $A$, implying that $\langle v_2, v_1\rangle = 0$ (since $\lambda_1 - \lambda_2 \neq 0$).

**Definition** *hermitian* A linear operator $A$ on a complex inner product space is called hermitian iff $\langle v_2, Av_1\rangle = \langle Av_2, v_1\rangle$ for any two vectors $v_1$ and $v_2$.

**Theorem** Eigenvalues of a hermitian operator are real.

**Theorem** On a complex vector space, eigenvectors of a hermitian operator with distinct eigenvalues are orthogonal.

**Corollary** The set of eigenvectors of a hermitian operator on a complex vector space can always be made orthogonal.

**1.** Consider $\frac{d^2}{dx^2}$ acting on real functions on the interval $[0, L]$ whose first derivative vanishes at the ends of the interval (so called free boundary conditions). Show that this operator (second derivative) is symmetric with the inner product $\langle f, g\rangle = \int_0^L dx f(x)g(x)$.

**2.** Redo for functions defined on the interval $[0, 2\pi]$ that are periodic, ie $f(0) = f(2\pi)$, $f'(0) = f'(2\pi)$, etc...

**3.** Consider the operator from the Bessel equation,

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right)$$

Show that it is symmetric under the inner product $\langle f, g\rangle = \int_0^R rdr\ f(r)g(r)$ for functions that vanish at $r = R$ and are finite at $r = 0$.

Lec 9

We know that the solutions of $Ax = 0$ are always a vector space, we know how to think about the general solution, which, as you will recall from examples in worksheets and homework, is in a form of a sum with undertermined coefficients. Recall the concept of a **span**:

Consider a set of vectors in some vector space $V$; $S = \{v_1, v_2, v_3, ..., v_k\} \subset V$.

The (much larger!) set of vectors of the form $\sum_i^k c_i v_i$ is called the **span** of $S$. In math notation this is

$$\text{span}(S) = \left\{ \sum_i^k c_i v_i \mid c_i \in \mathbf{F} \right\}$$

span$(S)$ is a subspace of $V$. This should be pretty obvious: the sum of any two vectors of the form $\sum_i^k c_i v_i$ is also of this form, as is a scalar multiple of any such vector.

Going back to differential equations, we will proceed as follows: we want to solve the equation $Au = 0$, ie we want to find the kernel of $A$. To describe the kernel, we will find enough solutions so that the desired kernel is their span.

To solve a partial differential equation of the form $Au = 0$, we will:

- identify an appropriate vector space $V$ for $A$ to act on by imposing all relevant homogenous boundary conditions;
- separate variables to obtain a set of eigenvalue problems; identify appropriate vector space (ie, boundary conditions) for each eigevalue problem;
- recognize which of the eigenvalue problems have symmetric operators; solve those eigenvalue problems first (you will get specific values for separation constants); claim orthogonality of eigenvectors;

- multiply appropriate solutions to separated equations together to obtain a set of solutions to the equation $Au = 0$, each solution corresponding to a different value of the separation constant(s);
- write the general solution to $Au = 0$ as the span of all those solutions (summing over all possible values of the separation constant(s));
- impose nonhomogenous boundary conditions (if any) to restrict all or some of the arbitrary coefficients in the span.

Let's use this series of steps to solve a problem similar to that on Homework 2A: Laplace equation $\nabla^2 u = 0$ in two dimensions, with the Laplacian operator expressed in polar coordinates, $(r, \theta)$:

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} \ .$$

Solve this equation on the inside of the unit disk (ie, for $r < 1$) with the following boundary conditions: $u(r, \theta)$ is bounded and $u(1, \theta) = e^{e^{i\theta}} + e^{e^{-i\theta}}$ (which is a real function).

Lec 10

Consider a set of vectors in some vector space $V$; $S = \{v_1, v_2, v_3, ..., v_k\} \subset V$.

The (much larger!) set of vectors of the form $\sum_i^k c_i v_i$ is called the **span** of $S$. In math notation this is

$$\text{span}(S) = \left\{ \sum_i^k c_i v_i \mid c_i \in \mathbf{F} \right\}$$

span$(S)$ is a subspace of $V$. This should be pretty obvious: the sum of any two vectors of the form $\sum_i^k c_i v_i$ is also of this form, as is a scalar multiple of any such vector.
When span$(S)$ is all of $V$, $S$ is called a basis.
Of special interest are orthogonal bases.

Let $g_n$ be an orthogonal basis of some space $V$.

$$\langle g_n, g_m \rangle = \delta_{n,m} \|g_n\|^2$$

Then, for any vector $v$ in $V$, it must be possible to find coefficients $c_n$ so that

$$v = \sum_n c_n g_n$$

Explicitely, these coefficients are

$$c_k = \frac{\langle g_k, v \rangle}{\langle g_k, g_k \rangle} = \frac{\langle g_k, v \rangle}{\|g_k\|^2}$$

How do we know if our basis is orthogonal? A: when it is made of eigenvectors of a symmetric/hermitian operator $A$:

(i) $Ag_n = \lambda_n g_n$
(ii) $\langle g, Af \rangle = \langle Ag, f \rangle$ for all $f, g$

The trick is to use the correct inner product; once you have that, everything falls into place.
The inner product will invariably be an integral of the form

$$\langle g, f \rangle = \int_D \rho(x) dx \; \bar{g} f$$

where $D$ is the domain of the functions in $V$, and $\rho$ is some density which will often be just $\rho(x) = 1$.

We can prove the above formula for coefficients as follows: When solving an equation such as

$$\sum_{n=-\infty}^{\infty} c_n g_n = v \ ,$$

we can use the same method we used in Lecture 8. To find, say $c_0$, we take an inner product of both sides of the equation with the vector $g_0$:

$$\left\langle g_0, \sum_{n=-\infty}^{\infty} c_n g_n \right\rangle = \langle g_0, v\rangle$$

$$\sum_{n=-\infty}^{\infty} c_n \langle g_0, g_n\rangle = \langle g_0, v\rangle$$

$$\sum_{n=-\infty}^{\infty} c_n \langle g_0, g_0\rangle \delta_{0,n} = \langle g_0, v\rangle$$

$$c_0 \langle g_0, g_0\rangle = \langle g_0, v\rangle$$

$$c_0 = \frac{\langle g_0, v\rangle}{\langle g_0, g_0\rangle}$$

Similarly, to find any other coefficient $c_k$, we take an inner product with $g_k$ and obtain:

$$c_k = \frac{\langle g_k, v\rangle}{\langle g_k, g_k\rangle} = \frac{\langle g_k, v\rangle}{\|g_k\|^2}$$

Example: let's say you were solving the following problem, on a half disk in polar coordinates,

$$\nabla^2 u(r,\theta) = 0 \ , \quad \partial_\theta u(r,0) = \partial_\theta u(r,\pi) = 0 \ , \quad u(R,\theta) = \sin(\theta)$$

You are now masters at the steps that get you to a the facts that

$$u(r,\theta) = \sum_{n=0}^{\infty} c_n r^n \cos(n\theta)$$

and that

$$u(R,\theta) = \sum_{n=0}^{\infty} c_n R^n \cos(n\theta) = \sin(\theta)$$

The above equation cannot be solved by inspection for $c_n$. Instead, we go to our theory. We kno that $\partial_\theta^2 \cos(n\theta) = -n^2 \cos(n\theta)$; $\cos(n\theta)$ are eigenvectors of $\partial_\theta^2$ by definition. The relevant vector space is function with vanishing derivatives at 0 and $\pi$:

$$V = \{v : [0,\pi] \to \mathbb{R} \mid v'(0) = v'(\pi) = 0\}$$

The inner product turns out to be simple: $\langle v, w\rangle = \int_0^\pi d\theta v(\theta) w(\theta)$. We proved that the operator is orthogonal in our last worksheet $\langle \partial_\theta^2 v, w\rangle = \langle v, \partial_\theta^2 w\rangle$ (see solutions). Therefore,

go directly to the formula for coefficients:

$$R^k c_k = \frac{\int_0^\pi d\theta \sin(\theta)\cos(k\theta)}{\int_0^\pi d\theta \cos^2(k\theta)}$$

It is instructive to perform the integrals. Note that we got this one for 'free':

$$\int_0^\pi d\theta \cos(k\theta)\cos(n\theta) = 0 \quad \text{if } k \neq n \ ,$$

but we need to compute two more $\int_0^\pi d\theta \sin(\theta)\cos(k\theta)$ and $\int_0^\pi d\theta \cos^2(k\theta)$. When in doubt, convert to complex exponentials and expand:

$$\int_0^\pi d\theta \sin(\theta)\cos(k\theta) = \int_0^\pi d\theta \ \frac{e^{i\theta}-e^{-i\theta}}{2i}\frac{e^{ik\theta}+e^{-ik\theta}}{2}$$

$$\frac{1}{4i}\int_0^\pi d\theta \ \left(e^{i(1+k)\theta}+e^{i(1-k)\theta}-e^{i(-1+k)\theta}-e^{i(-1-k)\theta}\right)$$

$$\frac{1}{4i}\ \left(\frac{e^{i(1+k)\theta}}{i(1+k)}+\frac{e^{i(1-k)\theta}}{i(1-k)}-\frac{e^{i(-1+k)\theta}}{i(-1+k)}-\frac{e^{i(-1-k)}\theta}{i(-1-k)}\right)\Bigg|_0^\pi$$

$$\frac{1}{-4}\ \left(\frac{(-1)^{(1+k)}-1}{1+k}+\frac{(-1)^{(1-k)}-1}{1-k}-\frac{(-1)^{(-1+k)}-1}{-1+k}-\frac{(-1)^{(-1-k)}-1}{-1-k}\right)$$

$$\frac{1}{-4}\left(-(-1)^k-1\right)\ \left(\frac{1}{1+k}+\frac{1}{1-k}-\frac{1}{-1+k}-\frac{1}{-1-k}\right)$$

$$\frac{1}{4}\left((-1)^k+1\right)\ \left(\frac{2}{1+k}+\frac{2}{1-k}\right) = \frac{(-1)^k+1}{4}\ \frac{4}{1-k^2} = \frac{(-1)^k+1}{1-k^2}$$

What if $k=1$?? no problem, the numerator is 0. Is that the right answer, though?

$$\int_0^\pi d\theta \sin(\theta)\cos(\theta) = \int_0^\pi d\theta \ \frac{e^{i\theta}-e^{-i\theta}}{2i}\frac{e^{i\theta}+e^{-i\theta}}{2}$$

$$\frac{1}{4i}\int_0^\pi d\theta \ \left(e^{2i\theta}+1-1-e^{-2i\theta}\right) = 0$$

It is the right answer, good. The integral vanishes for all odd $k$. Next, we do

$$\int_0^\pi d\theta \cos^2(k\theta) = \frac{1}{2}\pi$$

You can: use a double angle formula $\cos^2(k\theta) = \frac{1}{2}(\cos(2k\theta)+1)$ or the complex exponentials. Best remember that $\cos^2$ and $\sin^2$ always average out to $\frac{1}{2}$ over complete quarter-periods, though. Pay attention to the special case at $k=0$, though!!!

$$\int_0^\pi d\theta \cos^2(0\theta) = \pi$$

Put this all together: for even $k > 0$,

$$R^k c_k = \frac{4}{\pi(1-k^2)}$$

and

$$c_0 = \frac{2}{\pi}$$

and therefore (let $k = 2m$)

$$u(r,\theta) = \frac{2}{\pi} \; + \; \frac{4}{\pi} \sum_{m=1}^{\infty} \frac{\cos 2m\theta}{1-4m^2} \left(\frac{r}{R}\right)^{2m}$$

This answer can now be truncated and plotted.

One important lesson: need to pay a lot of attention to special cases.

Lec 11

**1.** Consider the one-dimensional heat equation $\frac{\partial}{\partial t}T = \alpha\frac{\partial^2}{\partial x^2}T$. This describes the time evolution of a temperature profile $T(x,t)$ along, say, an insulated metal rod. We will take the initial condition to be a uniform temperature distribution, $T(x,0) = T_0$, where $T_0 > 0$ is some fixed temperature. There are two physically significant boundary conditions possible:
(1) the ends of the rod are in thermal contact with a (heat) bath at some temperature, say 0 for simplicity: $T(0,t) = T(L,t) = 0$.
(2) the end of the rod are insulated (no heat can leave through the ends): $\partial_x T(0,t) = \partial_x T(L,t) = 0$.
For each of these two cases, think about the physics and sketch qualitatively what $T(x,t)$ should look like at
(a) a very early time $t \approx 0$
(b) very late times $t \to \infty$ (ie, in thermal equilibrium)
(c) some intermediate time

**2.** For each of the two boundary conditions above, write down the general solution for $T(x,t)$.

**3.** What is $T(x,t)$ with initial condition $T(x,0) = T_0$ and under insulating boundary conditions (2)?

We will look at a problem very similar to the homework:

$$\frac{\partial}{\partial t}T(x,t) = \alpha \frac{\partial^2}{\partial x^2}T(x,t)$$

$$T(0,t) = T(L,t) = 0 \quad \text{homogeneous bc}$$

$$T(x,0) = T_0 \quad \text{inhomogeneous bc,} \quad \text{replacing } \gamma x(L-x)$$

Separation of variables $T(x,t) = f(x)g(t)$ leads to an eigenvalue problem with a symmetric operator $\partial_x^2$:

$$f''(x) = af(x) \quad \text{with} \quad f(0) = f(L) = 0$$

with solutions $f_n = \sin(n\pi x/L) \qquad a_n = -(n\pi/L)^2 \quad \text{for} \quad n \in \mathbb{Z}$

The other equation is

$$g'(t) = \alpha a g(t) = -\alpha(n\pi/L)^2 g(t)$$

with solutions proportional to

$$g_n = e^{-\alpha(n\pi/L)^2 t}$$

Together, the general solution is

$$T(x,t) = \sum_{n=1}^{\infty} c_n g_n(t) f_n(x) = \sum_{n=1}^{\infty} c_n e^{-\alpha(n\pi/L)^2 t} \sin(n\pi x/L)$$

A different way to do this is to try the following form of the most general function of two variables:

$$u(x,t) = \sum_{n=1}^{\infty} G_n(t) f_n(x) \ .$$

Substituting this into the differential equation,

$$\dot{u} = \sum_{n=1}^{\infty} \dot{G}_n(t) f_n(x) = \alpha \partial_x^2 u = \alpha \sum_{n=1}^{\infty} a_n G_n(t) f_n(x)$$

which implies that $\dot{G}_n(t) = \alpha a_n G_n(t)$, with solutions $G_n = c_n e^{\alpha a_n t} = c_n e^{-\alpha(n\pi/L)^2 t}$. The general solution is, again,

$$T(x,t) = \sum_{n=1}^{\infty} c_n g_n(t) f_n(x) = \sum_{n=1}^{\infty} c_n e^{-\alpha(n\pi/L)^2 t} \sin(n\pi x/L)$$

We need to determine the coefficients $c_n$ so that

$$u(x,0) = \sum_{n=1}^{\infty} c_n \sin(n\pi x/L) = T_0$$

Last week, we've seen that problems of the form

$$\sum_i c_i v_i = w$$

where $v_i$ are orthogonal vectors under some inner product are easy to solve. And $v_i$ are orthogonal because they are eigenvectors of a symmetric operator with distinct eigenvalues:

$$\langle v_n, v_m \rangle = \delta_{nm} \langle v_n, v_n \rangle \quad \text{if } \lambda_n \neq \lambda_m$$

We can multiply (using the inner product) both sides of the above sum by $v_j$:

$$\left\langle v_j, \sum_i c_i v_i \right\rangle = \langle v_j, w \rangle$$

Bring it inside (using the distributive law):

$$\sum_i c_i \langle v_j, v_i \rangle = \langle v_j, w \rangle$$

Realize that only $i = j$ term in the sum is nonzero, since $\langle v_j, v_i \rangle = 0$ if $i \neq j$, so that

$$c_j \langle v_j, v_j \rangle = \langle v_j, w \rangle$$

Apply this idea to our problem,

$$\sum_{n=1}^{\infty} c_n \sin(n\pi x/L) = T_0$$

to obtain

$$c_n = \frac{\langle \sin(n\pi x/L), T_0 \rangle}{\langle \sin(n\pi x/L), \sin(n\pi x/L) \rangle}$$

So now all that remains is doing the integrals.

$$\langle \sin(n\pi x/L), T_0 \rangle = \int_0^L dx \; T_0 \sin(n\pi x/L) = -\frac{T_0 L}{n\pi}(\cos(n\pi) - 1)$$

$$\langle \sin(n\pi x/L), \sin(n\pi x/L) \rangle = \int_0^L dx \; T_0 \sin^2(n\pi x/L) = \frac{L}{2}$$

The required coefficients are thus

$$c_n = -\frac{2T_0}{n\pi}(\cos(n\pi) - 1) = \begin{cases} \frac{4T_0}{n\pi} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}$$

The complete solution to the heat equation with the given initial condition is

$$T(x,t) = \frac{4T_0}{\pi} \sum_{n=1, n \text{ odd}}^{\infty} \frac{1}{n}\, e^{-n^2 \frac{\alpha\pi^2}{L^2} t}\, \sin(n\pi x/L)$$

$$= \frac{4T_0}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} e^{-(2k+1)^2 \frac{\alpha\pi^2}{L^2} t}\, \sin((2k+1)\pi x/L)\ .$$

**Is** it really possible to write any function of $x \in [0, L]$ as a sum of eigenvectors of $\partial_x^2$? Even when that function is not in the originally considered vector space?

In what sense is

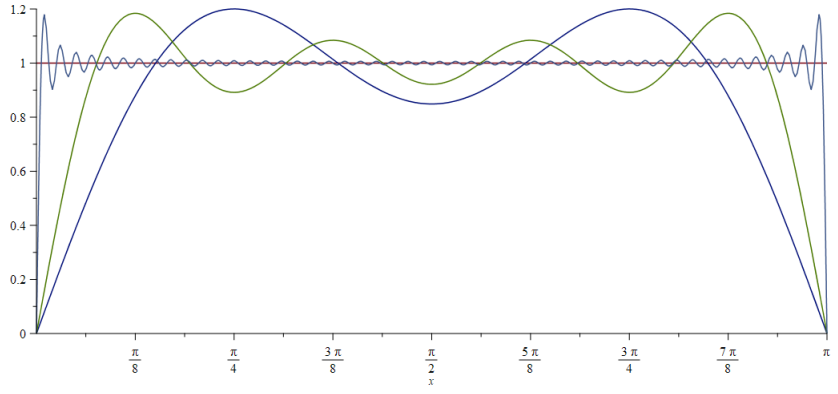
$$T_0^{-1} T(x,0) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1}\, \sin((2k+1)\pi x/L) = 1$$

true? What does it mean?

If I cut off the sum at some finite number (say, $N = 10$), can I say that

$$\frac{4}{\pi} \sum_{k=0}^{N} \frac{1}{2k+1}\, \sin((2k+1)\pi x/L) \approx 1 \quad ?$$

To see how well it actually works, consider plots for 2, 4 and 50 components:



**4.** Consider two functions $f : [0, L] \to \mathbb{R}$ and $g : [0, L] \to \mathbb{R}$. What does it mean for them to be close to each other? How could we measure how close are they?

Try to think of some ways to measure how close to each other two functions are.

The specific answer to the above question we want to use (because it's the one that gives nice theorems) is:

$$\|f - g\|^2 \quad = \quad \int_0^L dx \ (f(x) - g(x))^2$$

(Recall that $\|u\|^2 = \langle u, u \rangle$.)

With this definition, solving

$$\sum_{n=1}^{\infty} c_n \sin(n\pi x/L) = \gamma x(L - x)$$

for $c_n$ can be replaced with picking some $N$ and asking

For what $c_n$ is $\sum_{n=1}^{N} c_n \sin(n\pi x/L)$ is closest to $\gamma x(L - x)$ ?

We need to find $c_n$ that minimize

$$\left\langle \sum_{n=1}^{N} c_n \sin(n\pi x/L) - \gamma x(L - x), \sum_{n=1}^{N} c_n \sin(n\pi x/L) - \gamma x(L - x) \right\rangle$$

To find a miminum, set all derivatives w.r.t. $c_m$ to zero:

$$0 = \frac{d}{dc_m} \left\langle \sum_{n=1}^{N} c_n \sin(n\pi x/L) - \gamma x(L-x), \sum_{n=1}^{N} c_n \sin(n\pi x/L) - \gamma x(L-x) \right\rangle$$

$$= \left\langle \sin(m\pi x/L), \sum_{n=1}^{N} c_n \sin(n\pi x/L) - \gamma x(L-x) \right\rangle +$$

$$\left\langle \sum_{n=1}^{N} c_n \sin(n\pi x/L) - \gamma x(L-x), \sin(m\pi x/L) \right\rangle$$

$$= 2 \left\langle \sin(m\pi x/L), \sum_{n=1}^{N} c_n \sin(n\pi x/L) - \gamma x(L-x) \right\rangle = 0$$

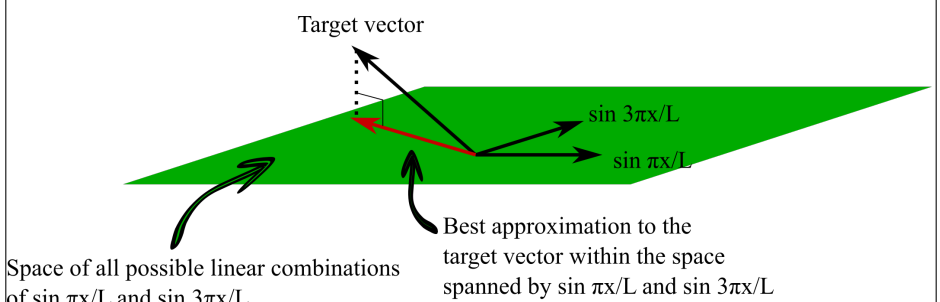
This is the same as what we had before:

$$\left\langle \sin(m\pi x/L), \sum_{n=1}^{N} c_n \sin(n\pi x/L) \right\rangle = \langle \sin(m\pi x/L), \gamma x(L-x) \rangle$$

$$c_m \langle \sin(m\pi x/L), \sin(m\pi x/L) \rangle = \langle \sin(m\pi x/L), \gamma x(L-x) \rangle$$

$$c_m = \frac{\langle \sin(m\pi x/L), \sin(m\pi x/L) \rangle}{\langle \sin(m\pi x/L), \gamma x(L-x) \rangle}$$

I find the following picture, drawn for $N = 2$, to be useful:



It's interesting to also look at how the temperature profile evolves in time, using just a finite number of components.
In the first animation, we use 50 components. Notice that the heat equation is smoothing - almost immediately, the temperature distribution becomes a smooth function that is 0 at the ends and close to constant through most of the middle.
In the second animation, we use just 2, 3 or 4 componets. At the time the first animation ended (t=0.1), the difference between the three plots is already small and the shape has converged to that seen at the end of the first animation.
The solutions look completely reasonable. And as long as we don't want to study the details of the very first instant of the process, we only need the first few components.

**5.** What is $\frac{\partial^2}{\partial x^2} f(x)$ if $f(x) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \sin((2k+1)\pi x/L)$ ?

Lesson: $\frac{\partial^2}{\partial x^2}$ seems to be a different operator depending on the boudary conditions used!

The correct procedure understanding a symmetric operator goes something like this:

Consider the space of suitably smooth functions which satisfy the boundary conditions

Find orthogonal eigenvectors of this symmetric operator in this space

Consider a span of the eigenvectors where the convergence of the infinite sum is really good; the action of the symmetric operator on a sum is determined by its eigenvalues

Now, extend the space of vectors and the action of the symmetric operator in a 'smooth' way to a more general space of functions (often, this means $L^2$). The technical terms is 'a self-adjoint extension'.

The self-adjoint operator you get depends on the boundary conditions. Its existance is assumed by physicists and (hopefully) proved by mathematicians.

A comment on our definition of inner product: Consider two functions, $f : [0, L] \to \mathbb{R}$ and $g : [0, L] \to \mathbb{R}$. We can approximate these functions by a vector of $n+1$ discrete data points, $x = iL/n$ for $i$ from 0 to $n$:

$$F = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ g(x_n) \end{bmatrix} \quad G = \begin{bmatrix} g(x_0) \\ g(x_1) \\ \vdots \\ g(x_n) \end{bmatrix}$$

Then, the dot product between these column vectors is an approximation to the inner product

$$\langle f(x), g(x) \rangle = \int_0^L f(x)g(x) \approx \sum_{i=0}^{n} f(x_i)g(x_i) = F \cdot G$$

Lec 12

**Functions of several variables**

The gradient at any point points towards the direction that the function is increasing the fastest. Another way to put this is that the linear approximation to a multivariable function is

$$f(\vec{x}) \approx f(\vec{x}_0) + \vec{\nabla} f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$

The specific example we want to talk about first is when the function represents temperature (as a function of position). When the temperature is not uniform, a non-zero temperature gradient results in a flow of heat. This flow is called heat flux, and is represented by a vector field: a functions that assigns a vector to each point in space. We will use $\vec{\Phi} = \vec{\Phi}(\vec{x}, t)$ to denote the heat flux. $\vec{\Phi}$ has units of (energy)/(area)/(time).

The heat flux is proportional to the temperature gradient,

$$\vec{\Phi} = -\kappa \vec{\nabla} T$$

where $\kappa$ is the heat conductivity.

Consider now a small, box-shaped region of space given by $x_0 < x < x_0 + \Delta x$, $y_0 < y < y_0 + \Delta y$, $z_0 < z < z_0 + \Delta z$. The rate at which heat flows into the box through the face at $x = x_0$ is given by the x-component of $\vec{\Phi}$ at $x = x_0$, $\Phi_x(x_0)$, times the area of this face $\Delta y \Delta z$: $\Phi_x(x_0)\Delta y \Delta z$. Similarly, the rate at which heat flows into the region through the face at $x = x_0 + \Delta x$ is $-\Phi_x(x_0 + \Delta x)\Delta y \Delta z$. Repeating with the other 4 faces, we get the total rate at which heat flows into the box as

$$\left[\Phi_x(x_0) - \Phi_x(x_0 + \Delta x)\right] \Delta y \Delta z + \left[\Phi_y(y_0) - \Phi_y(y_0 + \Delta y)\right] \Delta x \Delta z + \left[\Phi_z(z_0) - \Phi_z(z_0 + \Delta z)\right] \Delta x \Delta y$$

$$= -\left[\frac{\Phi_x(x_0) - \Phi_x(x_0 + \Delta x)}{\Delta x} + \frac{\Phi_y(y_0) - \Phi_y(y_0 + \Delta y)}{\Delta y} + \frac{\Phi_z(z_0) - \Phi_z(z_0 + \Delta z)}{\Delta z}\right] \Delta x \Delta y \Delta z$$

$$= -\left(\frac{\partial}{\partial x}\Phi_x + \frac{\partial}{\partial y}\Phi_y + \frac{\partial}{\partial z}\Phi_z\right) \text{Vol(box)}$$

Defining new notation:

$$\vec{\nabla} \cdot \vec{\Phi} \ := \ \frac{\partial}{\partial x}\Phi_x + \frac{\partial}{\partial y}\Phi_y + \frac{\partial}{\partial z}\Phi_z$$

for a quantity called the **divergence** of $\Phi$, from conservation of energy we can connect the rate of change of the thermal energy inside the box to the heat flux:

$$\frac{d}{dt}E = -\vec{\nabla} \cdot \vec{\Phi} \ \text{Vol(box)}$$

A more general theorem extends this result to general region shapes (notice that any region can be build up, at least approximately, from small box-shaped pieces). It is called the divergence theorem, and it states that

$$\oiint_{\partial R} ds \ \hat{n} \cdot \vec{\Phi} \ = \ \iiint_R dv \ \vec{\nabla}\vec{\Phi}$$

where $R$ is any region, $\partial R$ is the region's boundary surface, $\hat{n}$ is a vector normal to this surface pointing outwards. (Read more about this theorem in Ref. 2 (see Syllabus)).

So, for any region, it is true that

$$\frac{d}{dt}E = -\oiint_{\partial R} ds\ \hat{n}\cdot\vec{\Phi} \ = \ -\iiint_R dv\ \vec{\nabla}\vec{\Phi}$$

Returning to the infinitesimal version,

$$\frac{1}{\text{Vol(box)}}\ \frac{d}{dt}E = -\vec{\nabla}\cdot\vec{\Phi} = \kappa\vec{\nabla}\cdot(\vec{\nabla}T)$$

Thermal energy is given by $E = \text{Vol(box)}cT$ where $c$ is the heat capacity per unit volume. Then, with $\alpha = \kappa/c$, we get the **heat** equation

$$\frac{\partial}{\partial t}T = \alpha\nabla^2 T$$

where we have defined the **Laplacian** operator $\nabla^2$ by $\nabla^2 = \vec{\nabla}\cdot\vec{\nabla} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. When we separate the time variables from the space variables, we have $T = u(t)f(\vec{x})$, with

$$\dot{u} = \alpha\lambda u \quad \Rightarrow \quad u \sim e^{\alpha\lambda t}$$

and

$$\nabla^2 f(\vec{x}) = \lambda f(\vec{x})$$

There is a special solution to the heat equation, called the steady-state solution, which has no time dependence at all. It is simply the solution to $\nabla^2 T = 0$, possibly with nonhomogeneous boundary conditions to make the solution non-trivial.

To solve an initial value problem for the heat equation, with a known initial temperature distibution $T(\vec{x}, 0)$, we have to first solve the Laplacian eigenvalue problem, often called the Helmholtz equation. We want the solutions to be orthogonal, so let's prove symmetry of the Laplacian once and for all.

Consider the vector space of real functions on some arbitrary region $R$ in 3-dimensions which vanish on the boundary of $R$, $\partial R$. Prove that the Laplacian is symmetric under the inner product $\langle f, g\rangle = \iiint_R d^3x\ f(\vec{x})g(\vec{x})$. Hint: divergence theorem.

**1.** Consider the vector space of real functions on some arbitrary region $R$ in 3-dimensions which vanish on the boundary of $R$, $\partial R$. Prove that the Laplacian is symmetric under the inner product $\langle f, g\rangle = \iiint_R d^3x\ f(\vec{x})g(\vec{x})$. Hint: divergence theorem.

What other boundary conditions could we use, other than vanishing? If $\vec{n} \cdot \left(\vec{\nabla} f(\vec{x})\right)$ vanishes on the boundary for all functions $f$, then the Laplacian is also symmetric.

The physical intepretation of the vanishing boundary condition is clear: the region of interest is surrounded by something (say, water) that keeps its boundary at a constant temperature. Zero here is an arbitrary choice: you can always add a constant to the solution of the heat equation if you want the boundary temperature to be, say, 15 in some units. Also called the Dirichlet boundary condition.

The physical intepretation of $\vec{n} \cdot \left(\vec{\nabla} T(\vec{x})\right)$ is that the flux $\vec{\Phi} = -\kappa \vec{\nabla} T(\vec{x})$ is zero in the direction perpendicular to the boundary: $\vec{n} \cdot \vec{\Phi} = 0$. No energy flows across the boundary anywhere: an insulating boundary condition. Also called the Neumann boundary condition.

An in-between boundary condition (Robin) is when we assume Newton's law applies to the boundary: there is some ambient temperature $T_0$ and heat flux across the boundary is caused by the difference between that ambient temperature and the temperature of our body at its surface:

$$-\kappa \vec{n} \cdot \vec{\nabla} T(\vec{x}) = \vec{n} \cdot \vec{\Phi} = \beta \left(T(\vec{x}) - T_0\right)$$

or

$$\vec{n} \cdot \nabla T(\vec{x}) = -\frac{\beta}{\kappa}(T - T_0)$$

This condition inteporpolates between fixed/constant ($\frac{\beta}{\kappa}$ very large) and insulating ($\frac{\beta}{\kappa}$ very small).

A computation similar to that in the above question on the worksheet allows us to prove that on finite regions with vanishing boundary conditions, the Laplacian has only negative

eigenvalues. Let $\nabla^2 f = \lambda f$. Then

$$\lambda \int_R d^d x f(\vec{x}) f(\vec{x}) = \int_R d^d x f(\vec{x}) \vec{\nabla} \cdot (\vec{\nabla} f(\vec{x})) = \int_R d^d x \vec{\nabla} \cdot \Big( f(\vec{x}) (\vec{\nabla} f(\vec{x}) \Big) \; - \; \int_R d^d x \vec{\nabla} f(\vec{x}) \cdot \vec{\nabla} f(\vec{x})$$

$$\oint_{\partial R} ds \; \vec{n} \cdot \Big( f(\vec{x}) (\vec{\nabla} f(\vec{x}) \Big) \; - \; \int_R d^d x \vec{\nabla} f(\vec{x}) \cdot \vec{\nabla} f(\vec{x}) = 0 \; - \; \int_R d^d x \vec{\nabla} f(\vec{x}) \cdot \vec{\nabla} f(\vec{x}) \; \leq \; 0$$

This implies that $\lambda \geq 0$. Note that we have used that $\int_R d^d x (g(x))^2 \leq 0$.

If the boundary conditions are insulating: $\hat{n} \cdot \vec{\nabla} f(\vec{x}) = 0$ on the boundary $\partial R$, we reach the same conclusion. There is a difference, however: the constant function is a zero eigenvector of $\nabla^2$. For vanishing boundary conditions, constant nonzero functions are not in the vector space. For insulating boundary conditions, they are. In fact, constant functions are the only zero eigenvectors of $\nabla^2$ under either type of boundary condition, since $\nabla^2 f = 0$ implies that (using the above argument)

$$\int_R d^d x \vec{\nabla} f(\vec{x}) \cdot \vec{\nabla} f(\vec{x}) = - \int_R d^d x f(\vec{x}) \vec{\nabla} \cdot (\vec{\nabla} f(\vec{x})) = 0$$

which implies that $\vec{\nabla} f(\vec{x}) = \vec{0}$, which means that $f$ is constant. The conclusion is that: $\nabla^2$ has

- strictly negative eigenvalues with vanishing boundary conditions and
- non-positive eigenvalues with insulating boundary conditions, where the constant functions are the only ones with zero eigenvalue

Physically, this means that the time dependence $e^{\alpha \lambda t}$ always decays: the heat equation smooths out any initial configurations and approaches the steady state solution at late times.

Lec 13

Let's solve the heat equation in 2 dimensions (plus time):

$$\frac{\partial}{\partial t}T = \alpha\nabla^2 T \ ,$$

We want to solve for $T(x, y, t)$ with $0 < x < L$ and $0 < y < L$ with boundary conditions

$$T(0, y, t) = T(L, y, t) = T(x, 0, t) = T(x, L, t) = 0$$

In rectilinear coordinates, the Laplacian is just

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \ .$$

To illustrate some other ideas, let's use a piecewise defined initial condition ($D < L/2$):

$$T(x, y, 0) = \begin{cases} T_0 & \text{for } D < x < L - D \text{ and } D < y < L - D \\ 0 & \text{otherwise} \end{cases}$$

Let's try an approach like this: assume that we have eigenvectors of the spacial part (the Laplacian): $\nabla^2 g_s(x, y) = \lambda_s g_s(x, y)$, where $s$ is a label for the eigenvector. Then, let the solution be of the form

$$T(x, y, t) = \sum_s h_s(t) g_s(x, y)$$

. The equation for $h_s(t)$ is, as before $\dot{h}_s(t) = \alpha\lambda_s h_s(t)$

**1.** Let's start with the Laplacian eigenvalue problem:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) g = \lambda g$$

on a square with side $L$ and with vanishing boundary solutions:

$$g(0, y) = g(L, y) = g(x, 0) = g(x, L) = 0 \ .$$

Separate variables in this equation, solve each separated equation with the correct boundary conditions and explain why the (unnormalized) eigenvectors are labelled by a pair of positive integerts $m, n$ and are equal to

$$g_{n,m} = \sin(n\pi x/L)\sin(m\pi y/L)$$

Make sure to find $\lambda_{nm}$.

**2.** Write down the general solution to the heat equation,

$$\frac{\partial}{\partial t}T(x,y,t) = \alpha\nabla^2 T(x,y,t)$$

with the boundary conditions given above.

**3.** Match to the initial condition:

$$T(x,y,0) = \begin{cases} T_0 & \text{for } D < x < L-D \text{ and } D < y < L-D \\ 0 & \text{otherwise} \end{cases}$$

Lec 14

Reminder: heat equation is $\dot{T} = \alpha \nabla^2 T$. The general solution has a form

$$T = \sum_s c_s e^{\alpha \lambda_s t} f_s$$

where $f_s$ are eigenvectors of the Laplacian $\nabla^2$: $\nabla^2 f_s = \lambda_s f_s$. At very large $t$,

$$T \to \sum_{\substack{s \text{ s.t.} \\ \lambda_s = 0}} c_s f_s$$

I.e., at large time, the solution approaches a sum of zero eigenvectors of $\nabla^2$, which is the same as being a zero eigenvector of $\nabla^2$. At late times, the solutions becomes 1. time independent and 2. a solution to the Laplace equation.

The laplace equation $\nabla^2 T = 0$ can also be obtained from the heat equation by assuming that temperature is independent of time, so that $\dot{T} = 0$. This is why $\nabla^2 T = 0$ is also refered to as the time-independent heat equation. A solution to this equation is called the steady-state solution.

A simple way to find a solution to the heat equation with inhomogeneous (time independent) boundary conditions is to look for the steady-state solution and solve just the Laplace equation. This will give you the thermal equilibrium temperature distribution arising from the boundary conditions. Then, separately, you can solve the problem with homogeneous boundary conditions to get the general solution and to match to initial conditions. This is an example of the Divide and Conquer strategy. Use it for Question 2 on HW 4A.

Example: To solve $\dot{T} = \alpha \nabla^2 T$ with boundary conditions $T_{\text{boundary}} = T_0$, break it into two problems: $\nabla^2 T = 0$ with $T_{\text{boundary}} = T_0$ and $T$ independent of time (the steady state problem) and $\dot{T} = \alpha \nabla^2 T$ with boundary conditions $T_{\text{boundary}} = 0$, then add the two solutions.

Last lecture, we did proved some general statements about the Laplacian eigenvalue problem in arbitrary dimension using rectilinear coordinates. To obtain explicit solutions, however, we need to use coordinates adapted to the shape of the boundary. Today, we will talk about polar coordinates in two dimensions.

Let's solve the laplacian eigenvalue problem on a disk with radius $R$ and vanishing boundary conditions:

$$\nabla^2 u(r,\theta) = -\lambda u(r,\theta) \quad u(R,\theta) = 0 \qquad \nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}$$

As we know already, the eigenvalues of $\nabla^2$ will all be negative, so $\lambda > 0$.

Separation of variables: $u(r,\theta) = f(r)g(\theta)$. $g''(\theta) = -n^2 g(\theta)$ which leads to $g_n = e^{in\theta}$. The $r$ equation is

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r} f(r)\right) - \frac{n^2}{r^2} f(r) = -\lambda f(r)$$

This is not an equation that has a solution in terms of familiar functions. The best we can do is to show that this is the Bessel equation by changing variables to $z = \sqrt{\lambda} r$, and defining

a new function $Z$ it terms of $f(r)$ as $Z(z) = f(z/\sqrt{\lambda})$. Then, dividing both sides of the differential equation by $\lambda$:

$$\frac{1}{\sqrt{\lambda}r}\frac{d}{\sqrt{\lambda}dr}\left(\sqrt{\lambda}r\frac{d}{\sqrt{\lambda}dr}f(r)\right) - \frac{n^2}{\lambda r^2}f(r) = -f(r)$$

and then using using $\frac{d}{dz} = \frac{dr}{dz}\frac{d}{dr} = \frac{1}{\sqrt{\lambda}}\frac{d}{dr} = \frac{d}{\sqrt{\lambda}dr}$,

$$\frac{1}{z}\frac{d}{dz}\left(z\frac{d}{dz}f(z/\sqrt{\lambda})\right) - \frac{n^2}{z^2}f(z/\sqrt{\lambda}) = -f(z/\sqrt{\lambda})$$

$$\frac{1}{z}\frac{d}{dz}\left(z\frac{d}{dz}Z(z)\right) - \frac{n^2}{z^2}Z(z) = -Z(z)$$

$$\frac{d^2}{dz^2}Z(z) + \frac{1}{z}\frac{d}{dz}Z(z) + \left(1 - \frac{n^2}{z^2}\right)Z(z) = 0$$

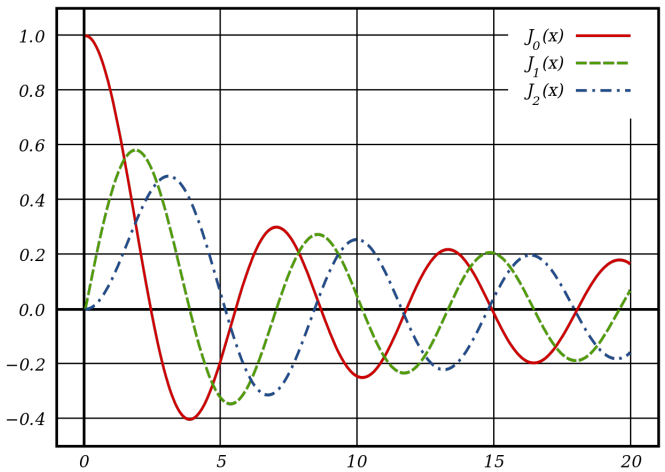
The last equation is the Bessel equation in its standard form.

Bessel functions that don't blow up at $z = 0$ are called Bessel functions of the first kind, and denoted with $J_n$. So, since $f(r) = Z(\sqrt{\lambda}r)$, our solutions are $f(r) = J_n(\sqrt{\lambda}r)$.

The boundary conditions become $Z(\sqrt{\lambda}R) = f(R) = 0$, ie $\sqrt{\lambda}$ must be a solution to the equation $J_n(\sqrt{\lambda}R) = 0$.

Let's look at some plots

From wikipedia, credit to Inductiveload, in Public Domain:




Looking at the plots, we can see that for every $n$, we get a series of values for $\lambda$ by solving the equation $J_n(\sqrt{\lambda}R) = 0$. This cannot be done in closed form. We can denote these solutions with $\lambda_{n,m}$, where $m$ denotes the number of 'half-waves' that fit from $r = 0$ to $r = R$.

Altogether, the eigenvectors of $\nabla^2$ take the form

$$u_{n,m} = J_n(\sqrt{\lambda}_{n,m}r)e^{in\theta} \ .$$

You can replace $e^{in\theta}$ with $\sin(n\theta)$ and $\cos(n\theta)$ as desired and most convenient depending on

boundary/inital conditions.

Look at some pretty pictures: https://www.chebfun.org/examples/disk/Eigenfunctions.html

We know already that these eigenvectors must be orthogonal, since we proved that $\nabla^2$ is symmetric which means that we know that

$$\langle u_{n,m}, u_{\tilde{n},\tilde{m}}\rangle \;=\; \int_0^R dr \int_0^{2\pi} r d\theta \; \overline{u_{n,m}} u_{\tilde{n},\tilde{m}} \;=\; \|u_{n,m}\|^2 \delta_{n\tilde{n}}\delta_{m\tilde{m}}$$

We have:

$$\int_0^R dr \int_0^{2\pi} r d\theta \; \overline{u_{n,m}} u_{\tilde{n},\tilde{m}} = \int_0^R dr \int_0^{2\pi} r d\theta \; J_n(\sqrt{\lambda}_{n,m} r) e^{-in\theta} J_{\tilde{n}}(\sqrt{\lambda}_{\tilde{n},\tilde{m}} r) e^{i\tilde{n}\theta} =$$

$$\int_0^{2\pi} d\theta e^{-in\theta} e^{i\tilde{n}\theta} \int_0^R r dr \; J_n(\sqrt{\lambda}_{n,m} r) J_{\tilde{n}}(\sqrt{\lambda}_{\tilde{n},\tilde{m}} r) =$$

$$2\pi\delta_{n\tilde{n}} \int_0^R r dr \; J_n(\sqrt{\lambda}_{n,m} r) J_{\tilde{n}}(\sqrt{\lambda}_{\tilde{n},\tilde{m}} r)$$

So, if we define an inner product for (real) functions of $r$ alone by $\langle f, h\rangle = \int_0^R r dr f(r) h(r)$, then we have

$$\langle J_n(\sqrt{\lambda}_{n,m} r), J_n(\sqrt{\lambda}_{n,\tilde{m}} r)\rangle = \int_0^R r dr \; J_n(\sqrt{\lambda}_{n,m} r) J_n(\sqrt{\lambda}_{n,\tilde{m}} r) = \delta_{m\tilde{m}} \|J_n(\sqrt{\lambda}_{n,m} r)\|^2$$

This is expected because $J_n(\sqrt{\lambda}_{n,m} r)$ are eigenvectors of the operator

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) - \frac{n^2}{r^2}$$

which is symmetric under the inner product $\langle f, h\rangle = \int_0^R r dr f(r) h(r)$ (see Worksheet 8, question 3).

Note: Bessel functions are orthogonal when they have the same $n$ and different $m$.

**Summmary of (cylindrical) Bessel functions**

When solving the Laplace eigenvalue problem in polar coordinates, you will have a differential equation of the form

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r} f(r)\right) \;-\; \frac{n^2}{r^2} f(r) = -\gamma f(r) \; .$$

This equation for $f$ is the Bessel equation, after we change variables to $z = \sqrt{\gamma}r$, $r(z) = z/\sqrt{\gamma}$:

$$\gamma \frac{1}{z}\frac{\partial}{\partial z}\left(z\frac{\partial}{\partial z}f(r(z))\right) - \frac{\gamma n^2}{z^2}f(r(z)) = -\gamma f(r(z))$$

$$\frac{1}{z}\frac{\partial}{\partial z}\left(z\frac{\partial}{\partial z}f(r(z))\right) + \left(1 - \frac{n^2}{z^2}\right)f(r(z)) = 0$$

The solutions to the Bessel equation are denoted by $J_n(z)$ and $Y_n(z)$. $Y_n(z)$ blow up at $z = 0$, but this does not bother us here since $r = 0$ is not part of the region where we are solving the equation. We need to keep both to match boundary conditions, so we get a general solution of the above equation as $f(r(z)) = AJ_n(z) + BY_n(z)$, or $f(r) = AJ_n(\sqrt{\gamma}r) + BY_n(\sqrt{\gamma}r)$.

Lec 15

The heat equation is not the only important equation in physics that is solved by separation of variables. Other such equations include the Schrodinger equation and the wave equation. These equations all have the Laplacian in them, but the time derivatives enter differently.

Let $\nabla^2 f_s = -\lambda_s f_s$, so $f_s$ are eigenvectors of $\nabla^2$.

HEAT:

$$\dot{T} = \alpha \ \nabla^2 T \qquad\qquad T = \sum_s c_s \ e^{-\alpha \lambda_s t} \ f_s$$

SCHRODINGER (free particle):

$$i\hbar \frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2m}\nabla^2 \Psi \qquad\qquad \Psi = \sum_s c_s \ e^{-\frac{i}{\hbar}\frac{\hbar^2 \lambda_s}{2m} t} \ f_s$$

WAVE:

$$\ddot{u} = v^2 \ \nabla^2 u \qquad\qquad u = \sum_s \quad ??? \quad f_s$$

The Schrodinger equation is basically the heat equation with an imaginary $\alpha$ (or you can think of it as the heat equation with imaginary time). $\frac{\hbar^2 \lambda_s}{2m}$ is the energy eigenvalue. For more, see Phys 304.

In this lecture, we will focus on the wave equation. We can derive it (in 1 dimension) by considering a system of masses connected by springs (like atoms in a 1d solid). This will be the same argument we saw in the first Jupiter worksheet, but presented backwards.

At equilibrium, the atoms are spaced by spacing $a$. The spring constant of each spring is $k$ and the atoms have mass $m$. If each atom moves away from the equilibirium spacing by $y_n$, then the $n^{\text{th}}$ atom will experience a force given by

$$F = k(y_{n-1} - y_n) + k(y_{n+1} - y_n) = m\ddot{y}_n$$

This is a system of differential equations for $y_n$, which we can write as

$$m\frac{d^2}{dt^2}\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{n+1} \end{bmatrix} = K \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{n+1} \end{bmatrix}$$

$$K = \begin{bmatrix} k & -2k & k & & & \\ & k & -2k & k & & \\ & & \ddots & \ddots & \ddots & \\ & & & k & -2k & k \\ & & & & & \end{bmatrix}$$

We introduced the extra atom at each end to be able to impose a fixed boundary condition, $y_0 = y_{n+1} = 0$. We can now get rid of those:

$$m\frac{d^2}{dt^2}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = K \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} \qquad \ddot{y} = \frac{1}{m}Ky$$

$$K = \begin{bmatrix} -2k & k & & & & \\ k & -2k & k & & & \\ & k & -2k & k & & \\ & & \ddots & \ddots & \ddots & \\ & & & k & -2k & k \\ & & & & k & -2k \end{bmatrix}$$

**1.** Let $n = 1$. Write an equation for $y_1$ and solve.

**2.** Let $n = 2$. Find the eigenvectors $v_1$ and $v_2$ of the $2 \times 2$ matrix $K$.

As we saw above, $\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = c_1(t)v_1 + c_2(t)v_2$. What are $c_1$ and $c_2$?

When $n$ is very large, it's interesting to zoom out of the atomic picture and consider a continuum limit. At a point $x = ka$ along the chain, the displacement of an atom is

$y(x) = y(ka) = y_k$. The equation for $y(x)$ is

$$m\ddot{y}(x) = k(y(x-a) - y(x)) + k(y(x+a) - y(x)) = ka^2 \frac{\frac{y(x+a)-y(x)}{a} - \frac{y(x)-y(x-a)}{a}}{a}$$

Taking $a$ to be very small, this becomes:

$$m\ddot{y}(x) = ka^2 \frac{y'(x) - y'(x-a)}{a} = ka^2 y''(x)$$

We have supressed the $t$ dependence of $y$ for clarity, replacing it we have

$$\ddot{y}(x,t) = \frac{ak}{m/a} \frac{\partial^2}{\partial^2 x} y(x,t)$$

$m/a$ is simply the (linear) density of the chain of atoms (mass per length). $ak$ has an intepretation as an elastic modulus. The easiest case to understand is when if the relaxed length of each spring is zero. Then, $ak$ is the tension in the chain of springs. More generaly, if $a$ is the equilibrium length of a string, then we have a rigid body such as a rod, then $ak$ is the elastic modulus, which is defined as the ratio of stress to strain:
elastic modulus = stress/strain = force / $(\Delta L/L)$ = $L$(force/$\Delta L$) = $L$(k/n)=($L$/n)k = $ak$
In either case, $ak/(m/a)$ has a macroscopic intepretation in terms of the collection of masses and springs as a whole.

Let's recall something from first year physics: a periodic moving wave is described by

$$\sin\left(2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right) + \phi\right)$$

This is a wave moving to the right; when moving to the left, we have instead

$$\sin\left(2\pi\left(\frac{x}{\lambda} + \frac{t}{T}\right) + \phi\right)$$

Substituting this into the wave equation, we see that it is a solution as long as

$$\frac{ak}{m/a} = \frac{\text{tension}}{\text{density}} = \frac{\text{elastic modulus}}{\text{density}} = \frac{\lambda^2}{T^2} = v^2$$

We will write the wave equation then as

$$\ddot{y}(x,t) = v^2 \frac{\partial^2}{\partial^2 x} y(x,t)$$

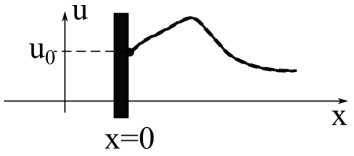
The higher dimensional version is

$$\ddot{y} = v^2 \nabla^2 y$$

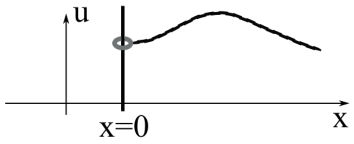
and these equations apply equally well to transverse oscillations as to longitudinal oscillations.

Boundary conditions for the 1D wave equation: $\ddot{u} = v^2 \frac{\partial^2}{\partial x^2} u$

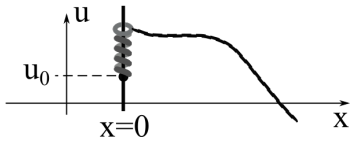
Fixed bc (Dirichlet)
$u(0,t) = u_0$




Free bc (Neumann)
$\partial_x u(0,t) = 0$




In-between bc (Robin)
$\partial_x u(0,t) = -\mu(u(0,t) - u_0)$




**3.** Let $\ddot{u} = v^2 \nabla^2 u$. Then, $u = \sum_s$ ??? $f_s(\vec{x})$. What does ??? represent?
Comment: $f_s(\vec{x})$ are called the normal modes.
**4.** What are the normal modes of oscillation for a circular drum membrane with radius $R$ and wave speed $v$? Find the spacial dependence $f_s(\vec{x})$ and the corresponding frequency.

Lec 16

**1.** (Repeated from last lecture's worksheet, since we did not get to it.)
What are the normal modes of oscillation for a circular drum membrane with radius $R$ and wave speed $v$? Find the spacial dependence $g_s(\vec{x})$ and the corresponding frequency. Recall that in polar coordinates,

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}$$

Since, at a fixed $n$, $J_n(k_{nm}r)$ all eigenvectors of $D_n := \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) - \frac{n^2}{r^2}$ with distinct eigenvalues, they must be orthogonal if we can find an inner product for which this operator is symmetric. We saw such an inner product in a previous Worksheet:

$$\langle a, b\rangle_r := \int_0^R rdr\ a(r)b(r)$$

(subscript $r$ will help us keep track of different inner products here). We have

$$\langle h_{nm}, h_{nm'}\rangle_r = \|h_{nm}\|^2\delta_{mm'}$$

Note: same $n$!

We also have, for (complex) functions of the angle $\theta$,

$$\langle c, d\rangle_\theta = \int_0^{2\pi} d\theta\ \bar{c}(\theta)d(\theta)$$

and the two-dimensional inner product for function on a disk with a surface element $ds^2 = r\ dr\ d\theta$:

$$\langle u, v\rangle_{\text{disk}} = \int_{r\leq R} ds^2\bar{u}(r,\theta)v(r,\theta) = \int_0^R rdr\ \int_0^{2\pi} d\theta\ \bar{u}(r,\theta)v(r,\theta)$$

These are all compatible with each other, namely, for $u = ac$ and $v = bd$,

$$\langle ac, bc\rangle_{\text{disk}} = \int_0^R rdr \int_0^{2\pi} d\theta\ a(r)\bar{c}(\theta)b(r)d(\theta) =$$

$$\int_0^R rdr\ a(r)b(r) \int_0^{2\pi} d\theta\ \bar{c}(\theta)d(\theta) = \langle a, b\rangle_r \langle c, d\rangle_\theta$$

The inner product that made the $r$-part of the two dimensional Laplacian, $D_n = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) - \frac{n^2}{r^2}$, symmetric comes from the Jacobian for spherical coordinates.

Orthogonality of $g_{nm}$ works out as well (as it must, since they are eigenvectors of a Laplacian): $g_{nm} = e^{in\theta}h_{nm}$,

$$\langle g_{nm}, g_{jk}\rangle_{\text{disk}} = \langle h_{nm}, h_{jk}\rangle_r \langle e^{in\theta}, e^{ij\theta}\rangle_\theta = \langle h_{nm}, h_{jk}\rangle_r\ (2\pi)\ \delta_{nj} =$$

$$2\pi\langle h_{nm}, h_{nk}\rangle_r \delta_{nj} = 2\pi\|h_{nm}\|^2\delta_{mk}\delta_{nj}$$

Therefore, for a general function on the disk, $u(r, \theta)$, we can write

$$u(r,\theta) = \sum_{n=-\infty, m=1}^{\infty} c_{nm}h_{nm}(r)e^{in\theta}$$

with

$$c_{nm} = \frac{\langle h_{nm}e^{in\theta}, u\rangle_{\text{disk}}}{2\pi\|h_{nm}\|^2}$$

Phew!

## Divide and Conquer

When you are comfortable with thinking of functions as vectors and differentiation etc... as linear operators, you can view a linear partial differential equation as 'built' out of linear operators. Boundary conditions can also be described in terms of linear operators: for example, consider the space $V$ of functions of two variables, $r$ and $\theta$. Then define the linear operator (map) $L : V \to W$ where $W$ is a function of one variable, which assigns to each such function $f$ a function of one variable, $\theta$ given by $\partial_r f(1, \theta)$. In symbols $g = L(f)$ when $g(\theta) = \partial_r f(1, \theta)$, or even $L(f)(\theta) = \partial_r f(1, \theta)$. The homogeneous boundary condition $f(1, \theta) = 0$ is then given by the equation linear homogeneous equation $Lf = 0$.

A completely homogeneous problem would be given by a set such as

$$L_1 f = 0$$

$$L_2 f = 0$$

$$L_3 f = 0$$

where $L_1$ could represent differential operators such as $\nabla^2$ or $\nabla^2 - \alpha\partial_t$, etc... and others could represent boundary conditions, or even other differential equations (especially if there are multiple unknown functions).

Example of a linear, homogenous system: let's consider two unknown functions, say $u(r,\theta,t)$ and $w(\theta,t)$ with equations such as

$$\begin{aligned}
\nabla^2 u - c^2\partial_t^2 u &= 0 \\
u(R,\theta,t) - w(\theta,t) &= 0 \\
\rho\partial_t^2 w - \tau\partial_\theta^2 w - \partial_r u(R,\theta,t) &= 0 \\
u(r,\theta,t) - u(r,\theta+2\pi,t) &= 0 \\
w(\theta,t) - w(\theta+2\pi,t) &= 0
\end{aligned}$$

This could model a drum with a rim that has some flex to it.

If we add enough homogeneous equation, we might end up in the situation where the only solution is the trivial one where everything is zero. Here, we could add zero initial conditions

$$\begin{aligned}
u(r,\theta,0) &= 0 \\
\partial_t u(r,\theta,0) &= 0
\end{aligned}$$

Without these extra equations, we might have a need for a solution we will call the 'solution to the homogeneous problem', or the 'homogeneous solution' (it's sometimes also called the complementary solution).

We can start complicating this problem by adding inhomogeneous terms to the right hand sides of the equations. The trick is to find at least one solution to the inhomogeneous problem, and then add the general homogenous solution to it.

However, we cannot approach a problem with a whole bunch of inhomogeneous terms at once!

Recall that, given sufficient boundary conditions, the differential linear operators are symmetric under some inner product. Boundary condition operators and differential operators serve different functions: the homogeneous boundary conditions define the vector space on which the differential linear operators are symmetric. Working in the basis of eigenvectors of

the symmetric operator(s), you can reduce a partial differential equation to ordinary differential equations for the coefficients of these eigenvectors. Inhomogeneous boundary and/or initial conditions can be written in the eigenvector basis as well (using the fact that the eigenvectors are orthogonal), and result in boundary conditions for these ordinary differential equation.

Ig we have too many inhomogeneous boundary conditions, there might not seem to be enough homogeneous ones present. The 'divide and conquer' strategy is to split the problem into several 'elementary' ones, each containing only one inhomogeneous boundary condition.

As a simple example, consider the Laplacian problem on a disk:

$$\nabla^2 u = g(r,\theta) \qquad \partial_r u(1,\theta) = h(\theta)$$

The corresponding homogeneous problem is, of course,

$$\nabla^2 u_H = 0 \qquad \partial_r u_H(1,\theta) = 0$$

We know that eigenvalues of the laplacian are generally non-positive, and that the constant function is the only zero eigenvector. Therefore, the only solution here is $u_H(r,\theta) = c_0$, a constant. You can, if you don't believe the theorem, confirm this by separating variables and solving the usual way.

Now, we have two different homogeneous problems, one of which we know how to solve:

$$\nabla^2 u_1 = 0 \qquad \partial_r u(1,\theta) = h(\theta)$$

$$\nabla^2 u_2 = g(r,\theta) \qquad \partial_r u_2(1,\theta) = 0$$

The first we know how to do. The second involves inverting a symmetric operator, ie understanding the meaning of $u_2 = (\nabla^2)^{-1} g(r,\theta)$. Should we find some solutions (any solutions, don't have to be the most general ones), we can write the most general solution as $u = u_1 + u_2 + u_H$.

Notice that this technique parallels the one you learned for solving systems of linear equations.

Another simple example, similar to the homework: the heat equation $\nabla^2 T = \alpha \dot{T}$ with a boundary condition of the form $\partial_r T(R,\theta,t) = \kappa(T(R,\theta,t) - T_0)$, which is not homogeneous, and an inhomogeneous initial condition, $T(r,\theta,0) = T_i$. We could consider the following inhomogeneous subproblems:

$$\nabla^2 T = \alpha \dot{T} \qquad \partial_r T(R,\theta,t) = \kappa(T(R,\theta,t) - T_0) \qquad T(r,\theta,0) = 0 \qquad (1)$$

$$\nabla^2 T = \alpha \dot{T} \qquad \partial_r T(R,\theta,t) = \kappa T(R,\theta,t) \qquad T(r,\theta,0) = T_i \qquad (2)$$

but this does not help us, since the first problem still does not have a good symmetric operator. Instead, let's think about the physics. There should be a steady state solution, independent of time and possibly initial conditions. Removing time dependence, we have a simple problem $\nabla^2 T = 0$, $\partial_r T = \kappa(T - T_0)$ which is solved by $T = T_0$. To make this a

solution of the first problem above, we have to modify it like this:

$$\nabla^2 T = \alpha \dot{T} \qquad \partial_r T(R,\theta,t) = \kappa(T(R,\theta,t) - T_0) \qquad T(r,\theta,0) = T_0 \tag{3}$$

$$\nabla^2 T = \alpha \dot{T} \qquad \partial_r T(R,\theta,t) = \kappa T(R,\theta,t) \qquad T(r,\theta,0) = T_i - T_0 \tag{4}$$

We compensate for this modification in the second problem so that $T_1 + T_2$ solves the original problem still.

A lesson: It pays to be clever in how we divide up the problem into pieces to make your life easier. Use symmetry, physical intuition, check for constant solutions and generally try to identify the simplest set of inhomogeneous problems you can find.

Our top-level linear algebra view allows us to see the whole last problem:

$$\nabla^2 T = \alpha \dot{T} \qquad \partial_r T = \kappa(T - T_0) \qquad T(r,\theta,0) = T_i$$

as being of the form $LT = b$ where the linear operator $L$ is a list of operators:

$$L = (\nabla^2 - \alpha \dot{T}, \quad (\text{substitute } r = R)(\partial_r - \kappa), \quad \text{substitute } t = 0)$$

and where

$$b = (0, -\kappa T_0, T_i)$$

To divide and conquer, we consider the following three problems:

$$\begin{aligned} LT_H &= \underline{0} := (0,0,0) \\ LT_1 &= b_1 := (0, -\kappa T_0, T_0) \\ LT_2 &= b_2 := (0, 0, T_i - T_0) \end{aligned}$$

Once we solved all three, we add the solutions to obtain a solution to the original problem $T_H + T_1 + T_2$. This works because $b_1 + b_2 = b$, and therefore $LT = L(T_H + T_1 + T_2) = LT_H + LT_1 + LT_2 = \underline{0} + b_1 + b_2 = b$.

In the homogeneous case, we need to find the most general $T_H$. In the inhomogeneous cases, any particular solution will suffice. This is because a difference between any two solutions $T_A$ and $T_B$ of an inhomogeneous equation $LT_A = b$ and $LT_B = b$ is a solution to the homogeneous equation $L(T_A - T_B) = LT_A - LT_B = \underline{0}$.

In this case, the solution to the homogeneous problem is just $T_H = 0$, and the solution to the entire problem is therefore unique (something we have been glossing over up to this point).

As explained above, chosing $b_1 = (0, -\kappa T_0, T_0)$ and $b_2 := (0, 0, T_i - T_0)$ makes your life easy because then, by physical reasoning and/or inspection $T_1 = T_0$. Always consider simple solutions to problems similar to the one asked to see if you can make one or more of the inhomogeneous sub-problems easy (or at least easier).

**2.** Solve the heat equation $\dot{T} = \alpha \nabla^2 T$ in two dimensions in polar coordinates $r, \theta$, for a bounded temperature distribution $T(r, \theta, t)$, $R_1 \leq r \leq R_2$, $t \geq 0$, with the following boundary and initial conditions:

$$T(R_1, \theta, t) = T_1$$

$$T(R_2, \theta, t) = T_2$$

$$T(r, \theta, 0) = T_i$$

**(a)** Divide and conquer: try to divide this problem into more basic ones that are similar to what we have seen before.

**(b)** What Laplacian eigenvalue problem do you think you will need to solve?

**(c)** Write down a formal solution to the Laplacian eigenvalue problem.

**(d)** Solve your sub-problems from part (a).

Lec 17

Test 4 + a clarification of last lecture

Lec 18

Below are three problems; the second is a repetition from two lectures ago. For each, if necessary, figure out a 'divide and conquer' strategy, reducing the problem to two or more basic ones you know how to solve. Then, state what symmetric operator's eigenvectors will give you a good basis to work in. Don't necessarily work out the details of the solution, rather, focus on coming up with a strategy.

**1.** Solve the Laplace equation $\nabla^2 u = 0$ for $u$ as a function of $x, y, z$, on a unit cube: $0 < x < 1$, $0 < y < 1$, $0 < z < 1$, with boundary conditions

$$\begin{array}{lll} u(x,y,0) = x(1-x) & u(x,0,z) = \sin(z/\pi) & \partial_x u(0,y,z) = 0 \\ u(x,y,1) = y(1-y) & u(x,1,z) = \sin(x/\pi)\sin(z/\pi) & \partial_x u(1,y,z) = 0 \end{array}$$

**2.** Solve the heat equation $\dot{T} = \alpha \nabla^2 T$ in two dimensions in polar coordinates $r, \theta$, for a bounded temperature distribution $T(r,\theta,t)$, $R_1 \leq r \leq R_2$, $t \geq 0$, with the following boundary and initial conditions:

$$T(R_1,\theta,t) = T_1$$
$$T(R_2,\theta,t) = T_2$$
$$T(r,\theta,0) = T_i$$

**3.** Solve the Schrodinger equation

$$i\hbar \frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi + K\Psi$$

for the wavefunction of a spin 1/2 particle in one dimension, which has the form

$$\Psi(x,t) = \begin{bmatrix} \psi_+(x,t) \\ \psi_-(x,t) \end{bmatrix}$$

where $K$ is an interaction term between the two spin states $+$ and $-$:

$$K = \begin{bmatrix} 0 & r \\ r & 0 \end{bmatrix}$$

$r$ is (real) constant. Let the particle be confined to a box with length $L$:

$$\psi_\pm(0,t) = \psi_\pm(0,t) = 0$$

and take the initial wavefunction to be

$$\Psi(x,0) = \begin{bmatrix} \sin^2(\pi x/L) \\ 0 \end{bmatrix} .$$

Lec 19

Today, we are going to learn to solve inhomogeneous PDEs. The over-all methodology is very similar to solving inhomogeneous linear equations, familiar from matrix algebra. Let's review

The equations we are interested in are of the form $Av = b$, where $A$ is a matrix (ie, a linear operator), $v$ is a column vector and $b$ is a column vector. We will focus on the case where $v$ and $b$ are the same size (ie, live in the same vector space $V$ and $A : V \to V$).

We know that we can write $v = A^{-1}b$, **if** $A^{-1}$ exists. But, there are equations where $A^{-1}$ does not exist ($A$ has a nontrivial kernel, ie some zero eigenvectors). There are also equations where there are no solutions, because $b$ is not in the image of $A$. We learned that if we know the eigenvectors of $A$ with their eigenvalues, we can solve for $v$ it terms of those eigenvectors. Let $Au_s = \lambda_s u_s$, then if

$$b = \sum_s c_s u_s$$

for some coefficients $c_s$ (which we know how to compute), we can write

$$v = \sum_s \frac{c_s}{\lambda_s} u_s \ .$$

The tricky part is when one (or more) of $\lambda_s$ are zero. Let's say that there exists an eigenvectors with zero eigenvalue, denoted with $u_0$: $Au_0 = 0$. Then, there are two possibilities: if $c_0 = 0$, there exists a solution, just omit the zero eigenvector from the sum and add the kernel to the particular solution:

$$v = \sum_{s, s\neq 0} \frac{c_s}{\lambda_s} u_s \ + \ d_0 v_0$$

If $c_0 \neq 0$, then there is no solution!

Let's now see how this plays out with differential equations.

Example 1: Solve $\nabla^2 f(x,y) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(x,y) = 1$ with boundary conditions $f(0,y) = f(\pi,y) = f(x,0) = f(x,\pi) = 0$. Recall that the Laplacian with zero boundary conditions has no zero eigenvectors. (we proved in lecture 11 that the laplacian with vanishing boundary conditions has strictly negative eigenvalues). This means that the solution here will be unique.

Example 2: Solve $\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}\right) f(x,y) = 1$ with boundary conditions $f(0,y) = f(\pi,y) = f(x,0) = f(x,\pi) = 0$. This is more interesting, because the operator $A := \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$ does have a nontrivial kernel (vector space of zero eigenvectors) on the space of functions with these boundary conditions. We know (or can easily obtain by our usual methods) what the

most general vector in this kernel looks like (solution to $Af = 0$)

$$\sum_{n=1}^{\infty} c_n \sin(nx)\sin(nx)$$

In general, eigenvectors of $A$ are given by $\sin(nx)\sin(mx)$: $A\sin(nx)\sin(mx) = -(n^2 + m^2)\sin(nx)\sin(mx)$. If we just had a single (particular) solution to the inhomogeneous equation with the correct boundary conditions... call it $f_P$, the general solution would have the form

$$f(x,y) = f_P(x,y) + \sum_{n=1}^{\infty} c_n \sin(nx)\sin(nx)$$

We can obtain $f_P$ in the same way as for the finite dimensional case before, writing the RHS as a sum over eigenvectors of $A$ and dividing each term by its eigenvalue.

A: Assuming that $A$ is a symmetric operator, eigenvectors come to the rescue. Let $v_q$ be eigenvectors of $A$ with nonzero eigenvalues: $Av_q = \lambda_q v_q$ with $\lambda_q \neq 0$. If you can write $b$ as a linear combination of those $v_q$: $b = \sum_q c_q v_q$, then you can check that

$$A\sum_q \frac{1}{\lambda_q} c_q v_q = \sum_q \frac{1}{\lambda_q} c_q A v_q = \sum_q \frac{1}{\cancel{\lambda_q}} c_q \cancel{\lambda_q} v_q = \sum_q c_q v_q = b$$

so $v_P = \sum_q \frac{1}{\lambda_q} c_q v_q$ is a particular solution to $Av = b$.

However, if $A$ has any zero eigenvectors, $Av_0 = 0$, the equation might have no solution. One way to check is to compute $\langle b, v_0\rangle$. If this is nonzero, $Af = b$ has no solutions.

Let's put this theory to some practice. Example 1 in question 1. Example 2 in question 2.

**1.** (a) Find any particular solution to the equation $\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(x,y) = 1$ with boundary conditions $f(0,y) = f(\pi,y) = f(x,0) = f(x,\pi) = 0$.
(b) Write the general solution to the above problem and finish solving Example 1.

**2.** (a) Find a particular solution to the equation $\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}\right) f(x,y) = 1$ with boundary conditions $f(0,y) = f(\pi,y) = f(x,0) = f(x,\pi) = 0$.

(b) Find the general solution to $\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}\right) f(x,y) = \sin(3x)\sin y$ with the same boundary conditions as in part (a).

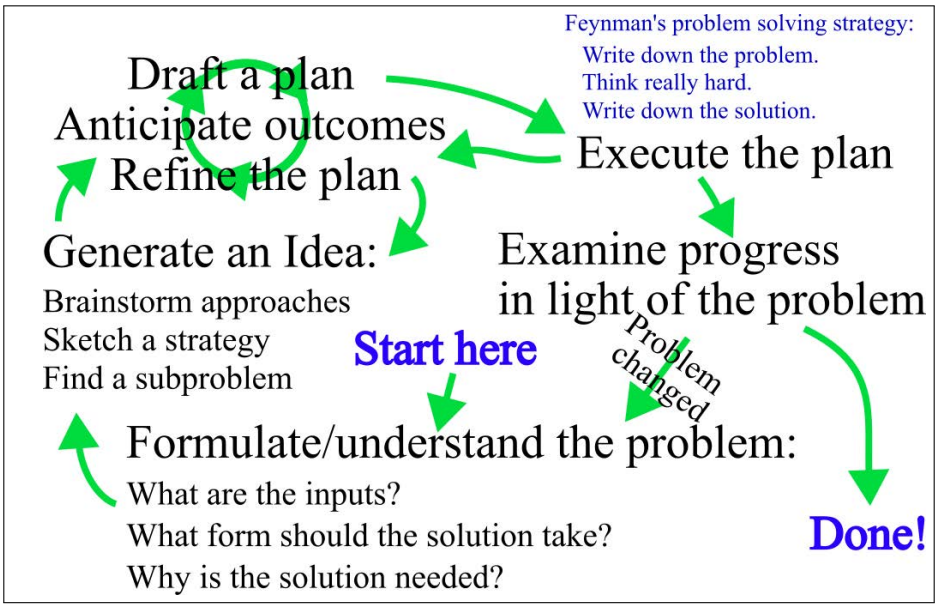
**3.** Solve the heat equation in 1d with a source: $\dot{T} = \alpha\partial_x^2 T + S(x)$ for $T(x,t)$ where

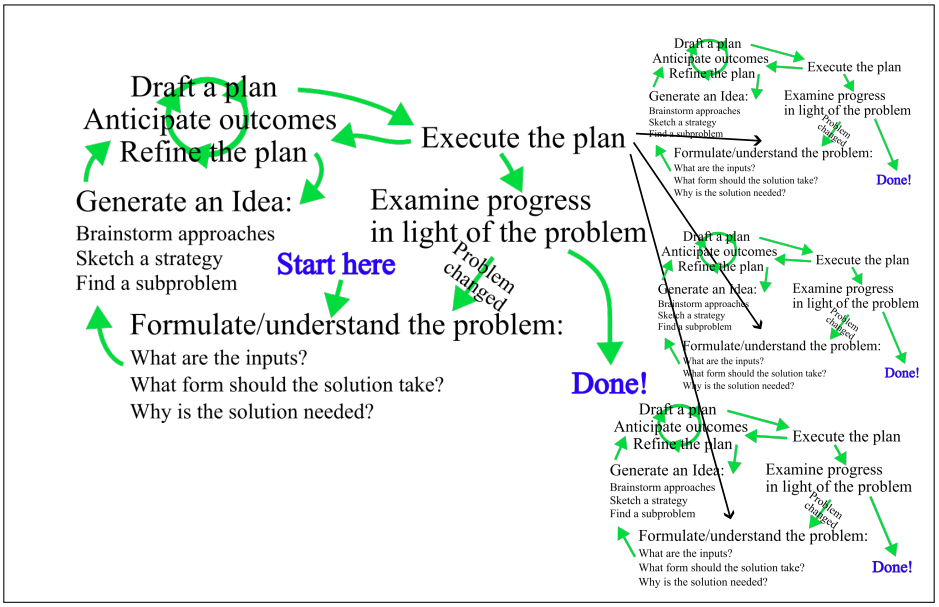
$$S(x) = \begin{cases} 0 & \text{for } x < L/3 \text{ or } x > 2L/3 \\ \Phi & \text{for } L/3 < x < 2L/3 \end{cases}$$

using insulating boundary conditions: $\partial_x T(0,t) = \partial_x T(L,t) = 0$ and a zero initial condition

$T(x,0) = 0$. The physical solution has a continuous, differentiable temperature with a continuous derivative in the $x$ direction (the derivative is proportional to the rate of heat flow along the $x$ direction).

Lec 20


Feynman's problem solving strategy:
Write down the problem.
Think really hard.
Write down the solution.
Draft a plan
Anticipate outcomes
Refine the plan
Execute the plan
Generate an Idea:
Brainstorm approaches
Sketch a strategy
Find a subproblem
Start here
Examine progress
in light of the problem
Problem changed
Formulate/understand the problem:
What are the inputs?
What form should the solution take?
Why is the solution needed?
Done!


Draft a plan
Anticipate outcomes
Refine the plan
Execute the plan
Generate an Idea:
Brainstorm approaches
Sketch a strategy
Find a subproblem
Start here
Examine progress
in light of the problem
Problem changed
Formulate/understand the problem:
What are the inputs?
What form should the solution take?
Why is the solution needed?
Done!

Solve the heat equation in 1d with a source:

$$\dot{T}(x,t) = \alpha \partial_x^2 T(x,t) + S(x)$$

$$S(x) = \begin{cases} 0 & \text{for } x < L/3 \text{ or } x > 2L/3 \\ \Phi & \text{for } L/3 < x < 2L/3 \end{cases}$$

Use insulating boundary conditions: $\partial_x T(0,t) = \partial_x T(L,t) = 0$ and a zero initial condition $T(x,0) = 0$.

The physical solution has a continuous, differentiable temperature with a continuous derivative in the $x$ direction (the derivative is proportional to the rate of heat flow along the $x$ direction).

Solve the heat equation in 1d with a source: $\dot{T} = \alpha \partial_x^2 T + S(x)$ for $T(x,t)$ where

$$S(x) = \begin{cases} 0 & \text{for } x < L/3 \text{ or } x > 2L/3 \\ \Phi & \text{for } L/3 < x < 2L/3 \end{cases}$$

using insulating boundary conditions: $\partial_x T(0,t) = \partial_x T(L,t) = 0$ and a zero initial condition $T(x,0) = 0$. The physical solution has a continuous, differentiable temperature with a continuous derivative in the $x$ direction (the derivative is proportional to the rate of heat flow along the $x$ direction).

Lec 21

In the problems we have solved so far in the course, we always had a symmetric operator with a complete basis of orthogonal eigenfunctions to base our solutions on.

There are, however, other (still physical) problems where the domain of the problem is infinite, there are no boundaries and the derivative operators are not symmetric in the sense in which we have understood it so far.

Consider the following pairs of examples:

(A) Solve the Laplace equation

$$\nabla^2 u = 0$$

where

A1: In polar coordinates $r, \theta$, let $u(R, \theta) = f(\theta)$. Solve for $u(r, \theta)$, $r < R$.

A2: In rectilinear coordinates $x, y$, $v(x, 0) = g(x)$, $v \to 0$ for $y \to \infty$. Solve for $v(x, y)$, $y > 0$.

(B) Solve the equation

$$\partial_x^2 u(x) = S(x)$$

where

B1: $u(x)$ and $S(x)$ are defined on an interval from 0 to $L$ with $u(0) = u(L) = 0$.

B2: $u(x)$ and $S(x)$ are defined on the entire real axis, and $\lim_{x \to \pm\infty} S(x) = 0$.[a]

[a] The real convergence condition will turn out to be somewhat different, but let's leave it like this for now.

(C) Solve the wave equation

$$\ddot{u} - u'' = F(t)$$

where

$$u(0,t) = u(L,t) = 0$$

$$u(x,0) = 0 \qquad \partial_t u(x,0) = 0$$

and where
C1: $F(t) = \cos\omega t$
C2: $F(t)$ is an arbitrary function with $F(t) = 0$ for $t < 0$.

(D) Solve the wave equation

$$\ddot{u} - u'' = 0$$

where

$$u(0,t) = 0$$

$$u(L,t) = F(t)$$

$$u(x,0) = 0 \qquad \partial_t u(x,0) = 0$$

and where
D1: $F(t) = \cos\omega t$
D2: $F(t)$ is an arbitrary function with $F(t) = 0$ for $t < 0$.

For A1, we would consider the eigenvectors of $\partial_\theta^2$ with periodic boundary conditions, $e^{im\theta}$ to expand $f(\theta)$ in. How to expand $g(x)$ in eigenvectors of $\partial_x^2$ the entire real line? Formally, $e^{ikx}$ is an eigenvector of this: $\partial_x^2 e^{ikx} = -k^2 e^{ikx}$, but what do we do with the continuous label $k$?

For B1, we can expand the inhomogeneous term $S(x)$ as a linear combination of $\sin(n\pi x/L)$. The coefficients can be found using inner products. Can we do the same in B2, write $S(x)$ as a linear combination of $e^{ikx}$ somehow? how do we find the coefficients?

For C1, we can be clever and try a solution of the form $u(x,t) = f(x)\cos\omega t$. You saw that this works in HW5B, question 2. But for C2, an arbitrary function is not an eigenvector of $\partial_t^2$ the way $\cos\omega t$ is, so that won't work. Can we write $F(t)$ as a linear combination of $e^{i\omega t}$ for various $\omega$s? How, again, do we find the relevant coefficients?

The same considerations as in (C) apply to (D): you can try $f(x)\cos(\omega t)$ in D1, but need something more sophisticated in D2.

Definition: Fourier transform $\hat{f}$ of a function $f$ is given by

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)\ e^{-ikx}$$

The Fourier inversion formula is

$$f(y) = \int_{-\infty}^{\infty} dk\ \hat{f}(k)\ e^{iky}$$

You can think of this as writing the function $f$ in the basis of $e^{ikx}$, which are eigenvectors of $-i\partial_x$ with real eigenvalues in the absense of any boundary conditions. The Fourier transform allows you to compute the coefficient $\hat{f}(k)$ of the eigenvector of $i\partial_x$ with eigenvalue $k$, ie $e^{ikx}$. The Fourier inversion formula says that when the Fourier transform is used as coefficients is the linear combination of these eigenvectors, the original function $f(x)$ is recovered. This is the Fourier inversion theorem, which holds when $f(x)$ the relevant integrals are sufficiently convergent (more about this later).

These should be compared as follows with an expansion in an orthogonal set of eigenvectors $g_n$

$$f(y) = \int_{-\infty}^{\infty} dk\ \hat{f}(k)\ e^{iky} \qquad g(y) = \sum_n c_n g_n(y)$$

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)\ e^{-ikx} \qquad c_n = \frac{\langle g_n, g\rangle}{\langle g_n, g_n\rangle}$$

The similarity is even stronger when we take $g_n(x) = e^{inx}$, ie $x$ is periodic with period $2\pi$:

$$f(y) = \int_{-\infty}^{\infty} dk\ \hat{f}(k)\ e^{iky} \qquad g(y) = \sum_n c_n e^{iny}$$

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)\ e^{-ikx} \qquad c_n = \frac{\langle e^{inx}, g(x)\rangle}{\langle e^{inx}, e^{inx}\rangle} = \frac{1}{2\pi}\int_0^{2\pi} dx\ g(x)e^{-inx}$$

We will now do the example (A) above, A1 to guide us and A2 to see how the Fourier transform works.

In this question, we will be solving the Laplace equation $\nabla^2 u = 0$ or $\nabla^2 v = 0$ in two dimensions, with nontrivial boundary conditions. $f$ and $g$ are some given functions. We will be solving the following two problems:

A1 In polar coordinates $r, \theta$, let $u(R, \theta) = f(\theta)$. Solve for $u(r, \theta)$, $r < R$.

A2 In rectilinear coordinates $x, y$, $v(x, 0) = g(x)$, $v \to 0$ for $y \to \infty$. Solve for $v(x, y)$, $y > 0$.

Start by solving the first problem below. Then, copy the key steps of your solution to the next page, and use them as a guide to solving the second problem.

Lec 22

Consider the wave equation with a time dependent boundary condition: $\ddot{u} = c^2 u''$, $u = u(x,t)$, $u(0,t) = 0$, $u(L,t) = \sin(\omega t)$. We will be looking for a particular solution.
(a) Use an ansatz $u(x,t) = f(x)\sin(\omega t)$. What is $f(x)$?

(b) Change the boundary condition at $x = L$ to $u(L,t) = A\sin(\omega t)$. What is the corresponding solution?

(c) Change the boundary condition at $x = L$ to $u(L,t) = e^{i\omega t}$. What is the corresponding solution?

(d) Change the boundary condition at $x = L$ to $u(L,t) = \frac{1}{1+e^{t-T}+e^{-t-T}} e^{i\omega t}$, where $T$ is a positive constant. What is the corresponding solution?

Consider a function $f(x)$ for $x$ from 0 to $L$. We can write

$$f(x) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{L} \ .$$

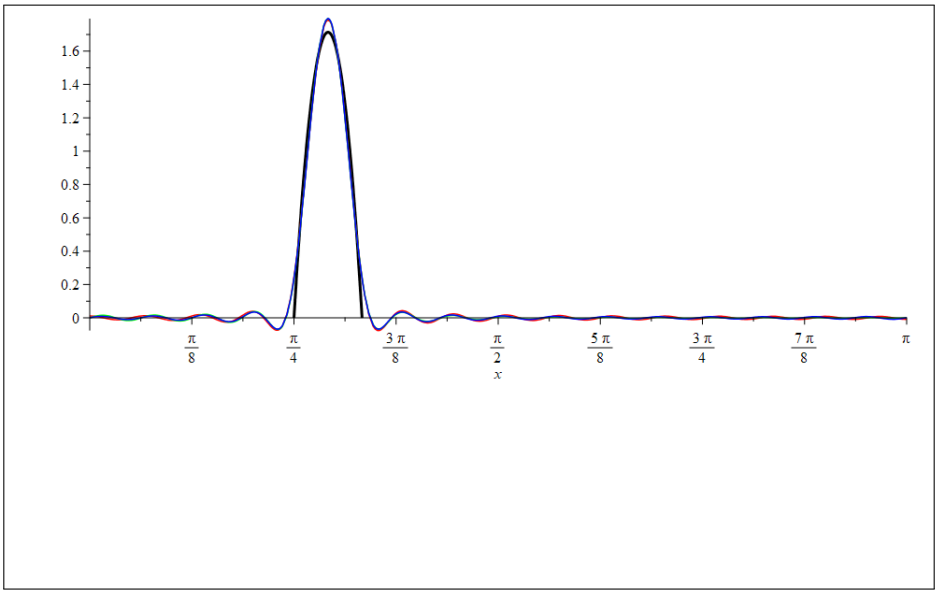
Consider a function $f(x)$ for $x$ from 0 to $L$. We can write

$$f(x) = d_0 + \sum_{n=1}^{\infty} d_n \cos \frac{n\pi x}{L} \ .$$

Consider a function $f(x)$ for $x$ from 0 to $L$. We can write

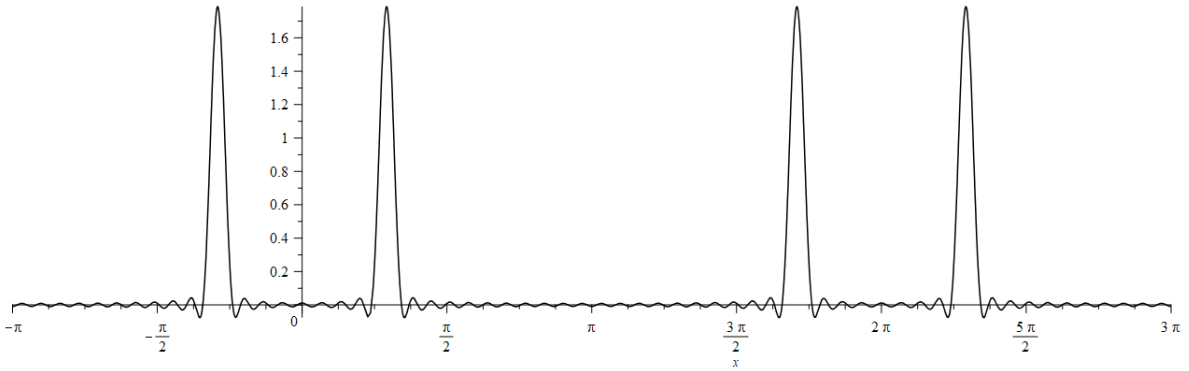
$$f(x) = \sum_{n=-\infty}^{\infty} e_n e^{2\pi i \frac{nx}{L}}$$

As an example, I have computed the first 30 coefficients in these three series (sin, cos, periodic), for a function on the interval 0 to $L = \pi$ which is $(x - \pi/4)(\pi/3 - x)$ between $\pi/4$ and $\pi/3$ and 0 otherwise. All three series do pretty well approximating this function between 0 and $\pi$:

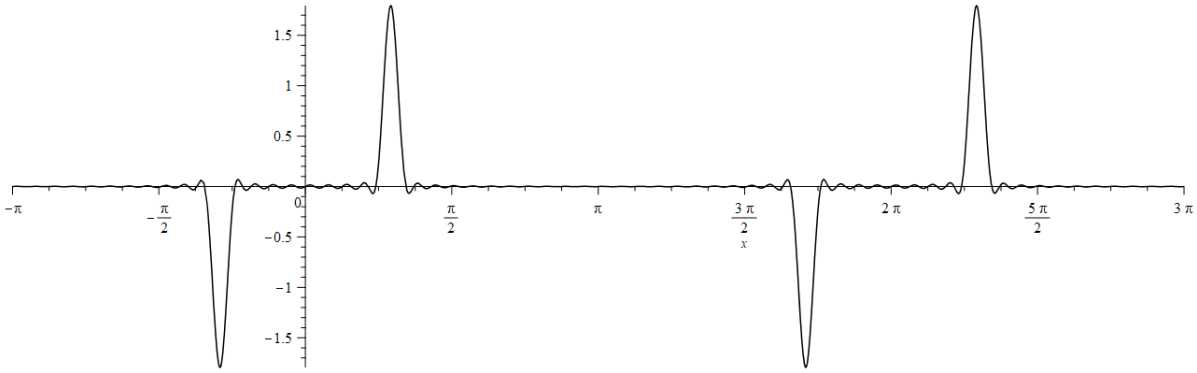



When we extend the plots of the functions outside of this interval, what we get is determined by the symmetries of each series:

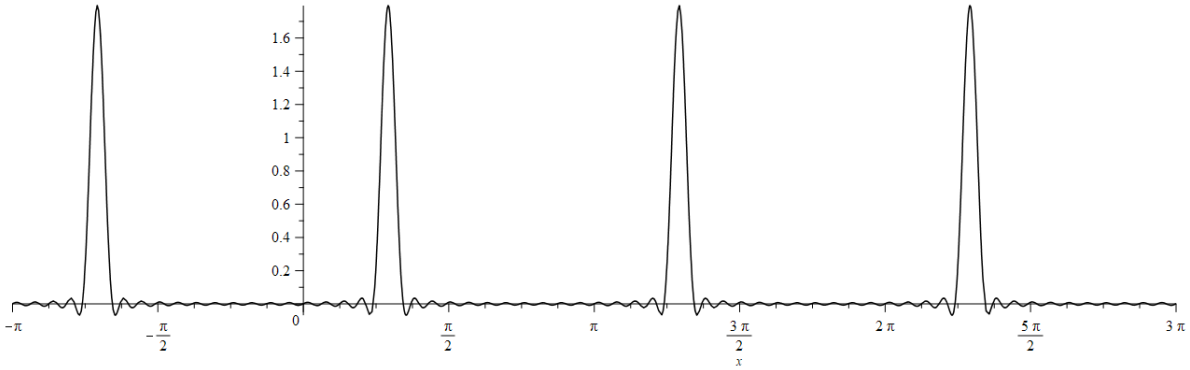
- The cosine series is symmetric under reflection $x \to -x$ and periodic under $x \to x + 2L = x + 2\pi$. The extended graph is:



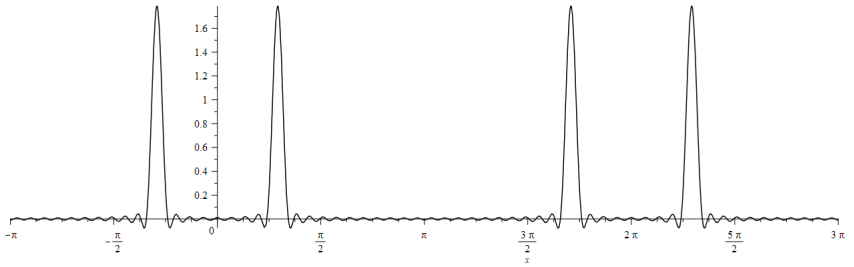

- The sine series is anti-symmetric under reflection $x \to -x$ and periodic under $x \to x + 2L = x + 2\pi$, so half of the peaks are flipped over. The extended graph is:



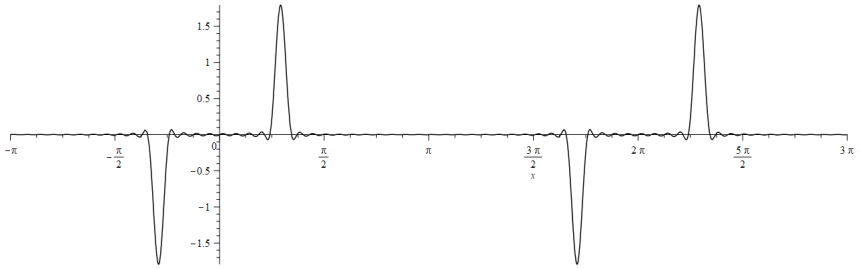

- Finally, the periodic series is simply periodic under $x \to x + L = x + \pi$ with no further symmetries, therefore the extended graph is:



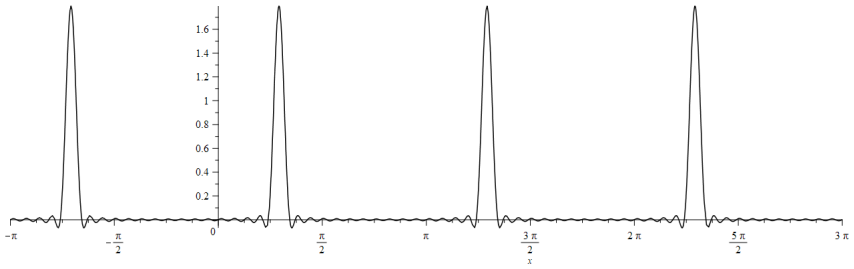

If we want to avoid having repeated peaks all together and simply have the function be zero outside the 'primary' peak, we need to have no periodic symmetry present at all. The only way to do that is to use a continuous set of wavevectors, ie use a Fourier transform instead of one of the periodic series. Last lecture, we saw that the fourier transform and its inverse are analogs of the period series expansions:




The cosine series is symmetric under reflection $x \to -x$ and periodic under $x \to x + 2L = x + 2\pi$.




The sine series is anti-symmetric under reflection $x \to -x$ and periodic under $x \to x + 2L = x + 2\pi$.




The periodic series is simply periodic under $x \to x + L = x + \pi$ with no further symmetries.

These should be compared as follows with an expansion in an orthogonal set of eigenvectors $g_n$

$$f(y) = \int_{-\infty}^{\infty} dk\ \hat{f}(k)\ e^{iky} \qquad g(y) = \sum_n c_n g_n(y)$$

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)\ e^{-ikx} \qquad c_n = \frac{\langle g_n, g\rangle}{\langle g_n, g_n\rangle}$$

The similarity is even stronger when we take $g_n(x) = e^{inx}$, ie $x$ is periodic with period $2\pi$:

$$f(y) = \int_{-\infty}^{\infty} dk\ \hat{f}(k)\ e^{iky} \qquad g(y) = \sum_n c_n e^{iny}$$

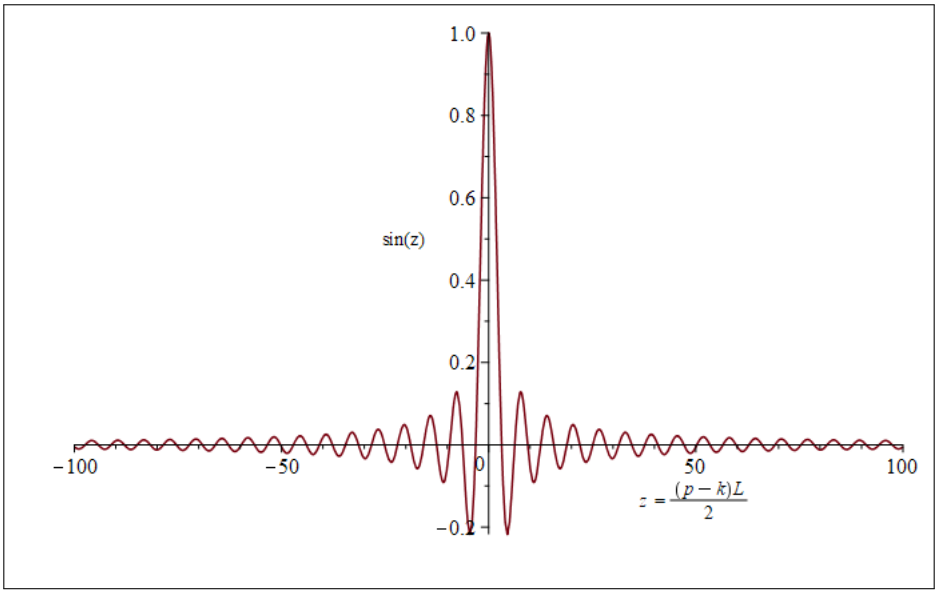
$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)\ e^{-ikx} \qquad c_n = \frac{\langle e^{inx}, g(x)\rangle}{\langle e^{inx}, e^{inx}\rangle} =$$

$$= \frac{1}{2\pi}\int_0^{2\pi} dx\ g(x)e^{-inx}$$

The fourier transform and its inverse allow us to 'expand' functions of the entire real axis in much the same way that we expanded functions on an interval using, say, the sine expansion. There is, however, a missing element in the correspondence: the fourier series we have seen (sine, cosine, periodic) are all based on orthogonal sets of (eigen)vectors. Is there a sense in which the functions $f_k(x) = e^{ikx}$ with different values of $k$ are orthogonal to each other?

Consider an inner product $\langle f, g\rangle = \int_{-L/2}^{L/2} dx\ \overline{f(x)}g(x)$. Then,

$$\langle e^{ikx}, e^{ipx}\rangle = \int_{-L/2}^{L/2} dx\ e^{-ikx}e^{ipx} = \frac{e^{i(p-k)L/2} - e^{-i(p-k)L/2}}{i(p-k)} = L\frac{\sin\left(\frac{(p-k)L}{2}\right)}{\frac{(p-k)L}{2}}$$




If you squint hard enough, you might believe that at the large $L$ limit, the result is zero unless $p = k$. Of course, at $L = \infty$, the result is infinite. We get a weird sort of 'function'. This is the continuous analog of the Kronecker delta $\delta_{ij}$, called the (Dirac) delta function and denoted with $\delta(p-k)$ (up to an overall factor of $2\pi$ that we will figure out in a moment).

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

$$\delta(p-k) = \begin{cases} \infty & \text{if} \quad p = k \\ 0 & \text{if} \quad p \neq q \end{cases}$$

Unfortunately, this is not really a definition we can work with. To be more precise, recall that the main property of the Kronecker delta was that

$$\sum_i a_i \delta_{ij} = a_j$$

for any sequence $a_i$. The analogous property for the delta function is that, for any 'test' function $\hat{f}(p)$,

$$\int dp \ \hat{f}(p)\delta(p-k) = \hat{f}(k)$$

In words, the function $\delta(p-k)$ is so sharply peaked around $p = k$ that only the value of $\hat{f}$ at $p = k$ contributes to the integral. The above formula is the definition of the delta function (actually, it's a distribution, which is a generalization of a function). The main takeaway should be that a delta function only makes sense multiplied by a smooth function under an integral.

Let's see how this is compatible with our formulas for fourier transform:

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \ f(x)e^{-ikx} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \ \left(\int_{-\infty}^{\infty} dp \ \hat{f}(p) \ e^{ipx}\right) \ e^{-ikx} =$$

$$= \int_{-\infty}^{\infty} dp\ \hat{f}(p) \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} dx\ e^{-ikx} e^{ipx} \right)$$

So, by definition, we extract that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dx\ e^{-ikx} e^{ipx} = \delta(p-k)$$

or

$$\int_{-\infty}^{\infty} dx\ e^{-ikx} e^{ipx} = 2\pi\delta(p-k)$$

By changing the variables around we also have

$$\int_{-\infty}^{\infty} dp\ e^{-ipy} e^{ipx} = 2\pi\delta(x-y)$$

The first statement is orthogonality of the vectors $e^{ikx}$. The second can be thought of as the statement that the fourier transform of $e^{ipx}$ is $2\pi\delta(x-y)$. A less rigorous way to derive these statements is to consider that, from the definition of the delta function:

$$e^{ikx} = \int_{-\infty}^{\infty} dp\ \delta(p-k) e^{ipx}$$

If we think about this equation as the inversion formula with $f(x) = e^{ikx}$ and $\hat{f}(p) = \delta(p-k)$, then we have that

$$\delta(p-k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx\ e^{ipx} e^{-ikx}$$

As an example using the delta function, let's consider a follow up question to the earlier worksheet problem, and put the delta function on in the boundary condition. The resulting solutions are sometimes called the impulse response.

Consider the question from Worksheet 22: we solved $\nabla^2 v = 0$ for $v(x,y)$ in rectilinear coordinates $x, y$, with $y > 0$ and $v(x,0) = g(x)$. We were able to perform a sum and obtain an answer in the form:

$$v(x,y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dz\ g(z)\ \frac{2y}{(x-z)^2 + y^2}$$

Now, let $g(x) = \delta(x - x_0)$ for some $x_0$. Find an explicit formula for $v(x,y)$ (it should depend on $x_0$).

Check that $v(x,y)$ you obtain satisfies the differential equation and the boundary condition.

Lec 23

Fourier transform is like a fourier series on an infinite interval. Compare with the periodic series, where $g_n(x) = e^{inx}$, $x$ is periodic with period $2\pi$.

$$f(y) = \int_{-\infty}^{\infty} dk\ \hat{f}(k)\ e^{iky} \qquad g(y) = \sum_n c_n e^{iny}$$

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)\ e^{-ikx} \qquad c_n = \frac{\langle e^{inx}, g(x)\rangle}{\langle e^{inx}, e^{inx}\rangle} =$$

$$= \frac{1}{2\pi}\int_0^{2\pi} dx\ g(x) e^{-inx}$$

Let's return to the example from lecture 21.

In this question, we were solving the Laplace equation $\nabla^2 u = 0$ or $\nabla^2 v = 0$ in two dimensions, with nontrivial boundary conditions. $f$ and $g$ are some given functions. We were solving the following two problems:

A1 In polar coordinates $r, \theta$, let $u(R, \theta) = f(\theta)$. Solve for $u(r, \theta)$, $r < R$.

A2 In rectilinear coordinates $x, y$, $v(x, 0) = g(x)$, $v \to 0$ for $y \to \infty$. Solve for $v(x, y)$, $y > 0$.

The equation in A2 is $\partial_x^2 v + \partial_y^2 v = 0$ and the boundary condition is $v(x, 0) = g(x)$. Take a fourier transform of both of these. Equation first.

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ \left(\partial_x^2 v(x, y) + \partial_y^2 v(x, y)\right)\ e^{-ikx} = 0$$

Integration by parts moves the $x$ partial derivative to the exponential:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ \left(v(x, y)\partial_x^2 e^{-ikx} + \partial_y^2 v(x, y) e^{-ikx}\right)\ = 0$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ \left(-k^2 v(x, y) + \partial_y^2 v(x, y)\right) e^{-ikx}\ = 0$$

$$\left(-k^2 + \partial_y^2\right) frac12\pi \int_{-\infty}^{\infty} dx\ v(x, y) e^{-ikx}\ = 0$$

Introduce notation for the fourier transforms:

$$(-k^2 + \partial_y^2)\hat{v}(k,y) = 0$$

Now do the boundary condition

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx \ \ (v(x,0)) \ \ e^{-ikx} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \ \ (g(x)) \ \ e^{-ikx}$$

sot hat

$$\hat{v}(k,0) = \hat{g}(k)$$

$k$ can now be treated like a fixed parameter, and we have an ODE in y. Easy to solve, the solution is

$$\hat{v}(x,y) = \hat{g}(k)e^{-|k|y}$$

where we took care to keep only the solution that decays at $y = \infty$, as instructed.

Before we finish, it's instructive to consider the same path in question A1. The equation is

$$\frac{1}{r}\partial_r(r\partial_r u(r,\theta)) + \frac{1}{r^2}\partial_\theta^2 u(r,\theta) = 0$$

and we can take inner products with a complete basis of periodic functions: $e^{in\theta}$:

$$\left\langle e^{in\theta}, \frac{1}{r}\partial_r(r\partial_r u(r,\theta)) + \frac{1}{r^2}\partial_\theta^2 u(r,\theta) \right\rangle = 0$$

$$\int_0^{2\pi} d\theta e^{in\theta}\left(\frac{1}{r}\partial_r(r\partial_r u(r,\theta)) + \frac{1}{r^2}\partial_\theta^2 u(r,\theta)\right) = 0$$

Integrate by parts in the $\theta$ direction, assuming that $u$ is periodic which kills the boundary terms:

$$\int_0^{2\pi} d\theta \left(\frac{1}{r}\partial_r(r\partial_r u(r,\theta))e^{-in\theta} + \frac{1}{r^2}u(r,\theta)\partial_\theta^2 e^{-in\theta}\right) = 0$$

$$\int_0^{2\pi} d\theta \left(\frac{1}{r}\partial_r(r\partial_r u(r,\theta)) + \frac{-n^2}{r^2}u(r,\theta)\right)e^{-in\theta} = 0$$

$$\left(\frac{1}{r}\partial_r(r\partial_r - n^2\right)\int_0^{2\pi} d\theta u(r,\theta)e^{-in\theta} = 0$$

Introduce notation for the integral (ie, coefficients):

$$\left(\frac{1}{r}\partial_r(r\partial_r - n^2\right)\hat{u}_n(r) = 0$$

Also take inner product in the boundary condition

$$\left\langle e^{in\theta}, u(R,\theta)\right\rangle = \left\langle e^{in\theta}, f(\theta)\right\rangle$$

$$\hat{u}_n(R) = \hat{f}_n$$

Now, we can solve the ODE for $\hat{u}_n(r)$, with the result

$$u_n(r) = \hat{f}_n (r/R)^n$$

It remains to sum this back up, as follows:

$$u(r,\theta) = \sum_{-\infty}^{\infty} \frac{u_n(r)}{\|e^{in\theta}\|^2} e^{in\theta} = \frac{1}{2\pi} \sum_{-\infty}^{\infty} \hat{f}_n \ (r/R)^n$$

What is the equivalent step in question A2? it's the fourier inversion thm (which we cannot prove): If $\hat{v}(k,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \ v(x,y) e^{-ikx}$, then

$$v(x,y) = \int_{-\infty}^{\infty} dk \ \hat{v}(k,y) e^{ikx}$$

Putting it all together,

$$v(x,y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp \ \left( \int_{-\infty}^{\infty} dz \ g(z) e^{-ipz} \right) e^{ipx} e^{-|p|y}$$

We can reverse the order of integration and do the $p$ integral (equivalent to doing the sum in the other question):

$$v(x,y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dz \ g(z) \ \left( \int_{-\infty}^{\infty} dp \ e^{-ipz} e^{ipx} e^{-|p|y} \right)$$

We have

$$\int_{-\infty}^{\infty} dp \, e^{-ipz} e^{ipx} e^{-|p|y} = \int_0^{\infty} dp \, e^{-ipz} e^{ipx} e^{-|p|y} + \int_0^{\infty} dp \, e^{ipz} e^{-ipx} e^{-|p|y} = 2\mathrm{Re} \int_0^{\infty} dp \, e^{-ipz} e^{ipx} e^{-py}$$

$$= 2\mathrm{Re} \frac{1}{i(z-x) - y} = \frac{2y}{(x-z)^2 + y^2}$$

Therefore

$$v(x,y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dz \ g(z) \ \frac{2y}{(x-z)^2 + y^2}$$

Why is the fourier inversion theorem true, though? one way to view it is to accept the existance of a delta 'function':

$$\int dp \ \hat{f}(p) \delta(p-k) = \hat{f}(k)$$

In words, the function $\delta(p-k)$ is so sharply peaked around $p = k$ that only the value of $\hat{f}$ at $p = k$ contributes to the integral. The above formula is the definition of the delta function (actually, it's a distribution, which is a generalization of a function). The main takeaway should be that a delta function only makes sense multiplied by a smooth function under an

integral.

Let's see how this is compatible with our formulas for fourier transform:

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)e^{-ikx} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ \left(\int_{-\infty}^{\infty} dp\ \hat{f}(p)\ e^{ipx}\right)\ e^{-ikx} =$$

$$= \int_{-\infty}^{\infty} dp\ \hat{f}(p)\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ e^{-ikx}e^{ipx}\right)$$

So, by definition, we extract that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ e^{-ikx}e^{ipx} = \delta(p-k)$$

or

$$\int_{-\infty}^{\infty} dx\ e^{-ikx}e^{ipx} = 2\pi\delta(p-k)$$

By changing the variables around we also have

$$\int_{-\infty}^{\infty} dp\ e^{-ipy}e^{ipx} = 2\pi\delta(x-y)$$

The first statement is orthogonality of the vectors $e^{ikx}$. The second can be thought of as the statement that the fourier transform of $e^{ipx}$ is $2\pi\delta(x-y)$. A less rigorous way to derive these statements is to consider that, from the definition of the delta function:

$$e^{ikx} = \int_{-\infty}^{\infty} dp\ \delta(p-k)e^{ipx}$$

If we think about this equation as the inversion formula with $f(x) = e^{ikx}$ and $\hat{f}(p) = \delta(p-k)$, then we have that

$$\delta(p-k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ e^{ipx}e^{-ikx}$$

We can also use the delta function to derive the inversion, as follows. Start with the definition:

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)e^{-ikx}$$

Insert delta:

$$\int_{-\infty}^{\infty} dq\delta(q-k)\hat{f}(q) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)e^{-ikx}$$

CHange to inner prdocut

$$\int_{-\infty}^{\infty} dq\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} dx e^{ipx}e^{-ikx}\right)\hat{f}(q) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)e^{-ikx}$$

Rearrange:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\left(\int_{-\infty}^{\infty} dq e^{iqx}\hat{f}(q)\right)e^{-ikx} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\ f(x)e^{-ikx}$$

Conclude:

$$\int_{-\infty}^{\infty} dq e^{iqx} \hat{f}(q) = f(x)$$

And there it is!

As an example using the delta function, let's consider a follow up question to the earlier worksheet problem, and put the delta function on in the boundary condition. The resulting solutions are sometimes called the impulse response.

Consider the question from Worksheet 22: we solved $\nabla^2 v = 0$ for $v(x, y)$ in rectilinear coordinates $x, y$, with $y > 0$ and $v(x, 0) = g(x)$. We were able to perform a sum and obtain an answer in the form:

$$v(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dz \; g(z) \; \frac{2y}{(x - z)^2 + y^2}$$

Now, let $g(x) = \delta(x - x_0)$ for some $x_0$. Find an explicit formula for $v(x, y)$ (it should depend on $x_0$).

Check that $v(x, y)$ you obtain satisfies the differential equation and the boundary condition.