MATH 4310: PARTIAL DIFFERENTIAL EQUATIONS

Daily Log for Lectures and Readings

Timothy E. Faver

December 2, 2024

CONTENTS

Day 1: Monday, August 12.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Section 1.2 has a broad overview of the subject and some important terms (like linear PDE and superposition). You definitely don't have to understand everything in here, but it gives a good vision of the subject and some important examples. We will revisit some of this material throughout the term.

Broadly, we care about PDE (I use this as both a singular and plural noun, depending on the context) because many interesting quantities in life depend on more than one variable. Your ODE course treated functions of a single variable, often time ($y = y(t)$), and now we will have multiple variables, often time and at least one spatial dimension ($u = u(x,t)$). Notation is always a nightmare, and I will say things like u_t and $\partial_t[u]$ to mean the partial derivative of u with respect to t. So if $u(x,t) = \cos(xt)$, then $u_t(x,t) = -\sin(xt)x$. I probably won't write

$$\frac{\partial u}{\partial t} \quad \text{or} \quad \frac{\partial}{\partial t}[u(x,t)].$$

You have no reason to care about these PDE right now, but here are some things that we will study:

$$\begin{array}{ll} u_t + u_x = 0 & \text{Transport equation} \\ u_t - u_{xx} = 0 & \text{Heat equation} \\ u_{tt} - u_{xx} = 0 & \text{Wave equation} \\ u_{tt} + u_{xx} = 0 & \text{Laplace's equation.} \end{array}$$

It turns out that we can represent all solutions to the transport equation very explicitly and compactly, and so that PDE will be a great "lab rat" as we develop new techniques—we can always see how something new compares to what we know about transport. We will not be so comprehensive with the other examples above, and in particular "boundary conditions"—whether x lives in a bounded interval or on the whole real line (maybe with some limit conditions on u as $x \to \pm\infty$)—will play a larger role there. Also, the algebraic structure of these PDE *sometimes* has a profound effect; we will see that if you know how to solve transport or heat as given, then you effectively get solutions to $u_t - u_x = 0$ and $u_t + u_{xx} = 0$. However, the $\pm$ distinction between wave and Laplace is important.

All four PDE are **LINEAR AND HOMOGENEOUS** in the sense that if u and v are solutions and c_1, $c_2 \in \mathbb{R}$, then $c_1u + c_2v$ is also a solution.

1.1 Problem. Prove that.

This phenomenon is sometimes gussied up with the term **SUPERPOSITION**, which fails for nonlinear problems. Here are two nonlinear equations that we will eventually study:

$$\begin{array}{ll} u_t + uu_x = 0 & \text{Burgers's equation} \\ u_t + u_{xxx} + uu_x = 0 & \text{Korteweg–de Vries (KdV) equation.} \end{array}$$

1.2 Problem. If u and v solve Burgers's equation, what goes wrong if you try to show that $c_1u + c_2v$ is also a solution for c_1, $c_2 \in \mathbb{R}$?

A major goal of our course, perhaps achievable only in hindsight, will be to understand how the algebraic structure of these PDE contributes to the existence and properties of solutions, and how the existence results and the exact properties differ from equation to equation. Simple, seemingly banal changes in the algebraic structure (the arrangement of $+$ and $-$, linearity or nonlinearity) and the analytic structure (what derivatives appear, how many, where) can lead to profound changes in the behavior of solutions to PDE.

Lawrence C.Evans, in his magisterial graduate-level text *Partial Differential Equations*, captures the challenge and the orientation of PDE study quite evocatively:

> *"There is no general theory concerning the solvability of all partial differential equations. Such a theory is extremely unlikely to exist, given the rich variety of physical, geometric, and probabilistic phenomena which can be modeled by PDE. Instead, research focuses on various particular partial differential equations that are important for applications within and outside of mathematics, with the hope that insight from the origins of these PDE can give clues as to their solutions."*

Peter Olver's book *Introduction to Partial Differential Equations* gives the following as a mission statement for a first undergraduate course in PDE, and I agree with it fully:

> *"[T]he primary purpose of a course in partial differential equations is to learn the principal solution techniques and to understand the underlying mathematical analysis."*

We will focus rather less on deriving PDE from models and physical principles and rather more on the solution techniques and mathematical analysis—and connections to other classes. No course is an island, and we will see here why you might want to study things in real and complex analysis and topology. However, when we do have explicit solutions for a PDE, we will comment on how their behavior reflects physical reality, or not.

Two of our major tools in this course will be integrals (definite and improper) and fundamental results from ODE. We will start by reviewing essential properties of the definite integral and then applying them to redevelop familiar results from ODE at a more abstract level (and more rapid pace). Throughout the course, we'll see that integrals fundamentally *measure and/or extract useful data about functions* (and all the cool kids want to be data scientists these days) and also *represent functions in convenient and/or meaningful ways*. You have already seen this in your calculus classes: the number

$$\frac{1}{b-a}\int_a^b f(x)\ dx$$

gives a good measure of the "average value" that the function f takes on the interval $[a, b]$, while the function

$$F(x) := \int_a^x f(t)\ dt$$

is an antiderivative of f in the sense that $F'(x) = f(x)$. Eventually we will see that integrals like

$$\int_a^b |f(x)|\ dx \quad \text{and} \quad \left(\int_a^b |f(x)|^2\ dx\right)^{1/2}$$

are good measures of "size" for f (that is, they are integral **NORMS**). We will also find representing functions via (inverse) Fourier transforms, which are defined via improper integrals, particularly convenient.

But to get anywhere, we need to be comfortable with how integrals work. I claim that you only need four properties of integrals in order to get the fundamental theorem of calculus (FTC), and all of those properties have geometric motivations (there are other motivations, too, but geometry/area is probably the most universally accessible). For simplicity (and to annoy the Calc II instructors), I'll write $\int_a^b f$ most of the time, and we'll agree that the dummy variable of integration doesn't matter:

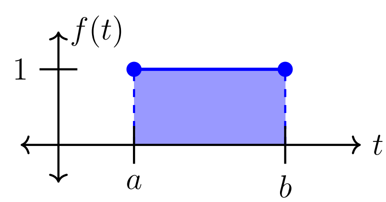
$$\int_a^b f = \int_a^b f(x)\ dx = \int_a^b f(u)\ du = \int_a^b f(s)\ ds = \int_a^b f(t)\ dt = \int_a^b f(\tau)\ d\tau.$$

That last dummy variable τ is the Greek letter "tau."

Here are those properties (taken from my complex analysis lecture notes).

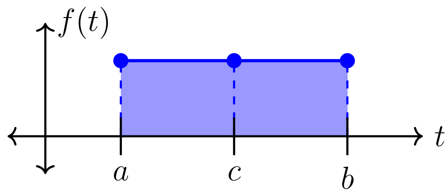
($\int$1) First, the integral of a function $f\colon [a,b] \subseteq \mathbb{R} \to \mathbb{R}$ should somehow measure the net area of the region between the graph of f and the interval $[a,b]$. Since the most fundamental area is the area of a rectangle, we should expect

$$\int_a^b 1\ dt = b - a.$$



($\int$2) If we divide the region between the graph of f and the interval $[a,b]$ into multiple components, measure the net area of those components, and add those net areas together, we should get the total net area of the region between the graph of f and the interval $[a,b]$. There are many such ways to accomplish this division, but perhaps one of the most straightforward is to split $[a,b]$ up into two or more subintervals and consider the net areas of the regions between the graph of f and those subintervals. So, we expect that if $a \le c \le b$, then

$$\int_a^c f(t)\ dt + \int_c^b f(t)\ dt = \int_a^b f(t)\ dt.$$

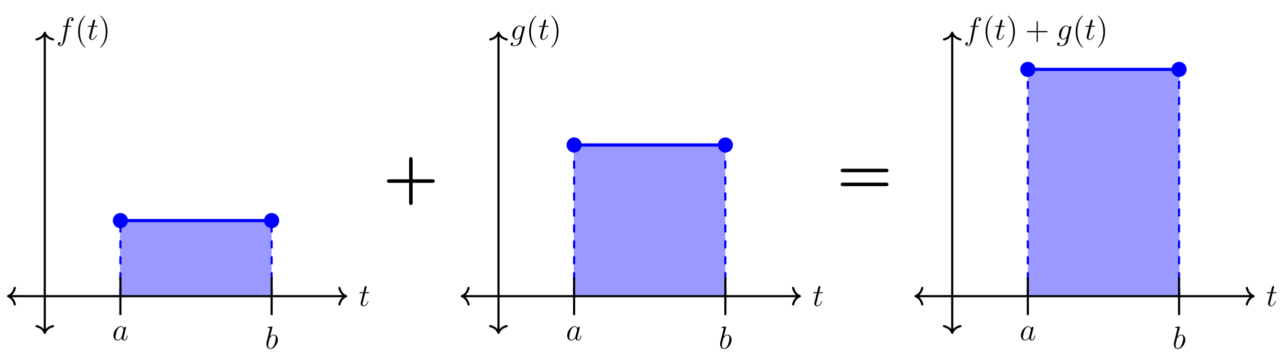


($\int$3) If f is nonnegative, the net area of the region between the graph of f and the interval $[a, b]$ should be the genuine area of the region between the graph of f and the interval $[a, b]$, and this should be a positive quantity. So, we expect that if $0 \leq f(t)$ on $[a, b]$, then

$$0 \leq \int_a^b f(t)\ dt.$$

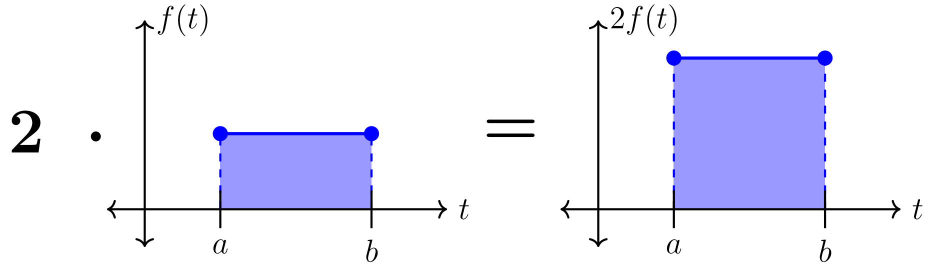
($\int$4) Adding two functions $f, g\colon [a, b] \subseteq \mathbb{R} \to \mathbb{R}$ should "stack" the graphs of f and g on top of each other. Then the region between the graph of f and the interval $[a, b]$ gets "stacked" on top of region between the graph of g and the interval $[a, b]$. Consequently, the net area of the region between the graph of $f + g$ and the interval $[a, b]$ should just be the sum of these two areas:

$$\int_a^b f(t)\ dt + \int_a^b g(t)\ dt = \int_a^b \big[f(t) + g(t)\big]\ dt.$$



Next, multiplying a function $f\colon [a, b] \subseteq \mathbb{R} \to \mathbb{R}$ by a constant $\alpha \in \mathbb{R}$ should somehow "scale" the net area of the region between the graph of f and the interval $[a, b]$ by that factor α. For example, the area under the graph of $2f$ over $[a, b]$ should be double the area under the graph. Consequently, the net area of the region between the graph of αf and the interval $[a, b]$ should be the product

$$\int_a^b \alpha f(t)\ dt = \alpha \int_a^b f(t)\ dt.$$



Day 2: Wednesday, August 14.

Here is a more formal and less geometric approach to the integral. Let $I \subseteq \mathbb{R}$ be an interval (for the rest of today, I is *always* an interval). Denote by $\mathcal{C}(I)$ the set of all continuous real-valued functions on I. We should be able to integrate every $f \in \mathcal{C}(I)$, and we can.

2.1 Theorem. *Let $I \subseteq \mathbb{R}$ be an interval and denote by $\mathcal{C}(I)$ the set of all continuous functions from I to $\mathbb{R}$. There exists a map*

$$\int : \{(f,a,b) \mid f \in \mathcal{C}(I),\ a,b \in I\} \to \mathbb{R} \colon (f,a,b) \mapsto \int_a^b f$$

with the following properties.

($\int$1) [Constants] *If a, $b \in I$, then*

$$\int_a^b 1 = b - a.$$

($\int$2) [Additivity of the domain] *If $f \in \mathcal{C}(I)$ and a, b, $c \in I$, then*

$$\int_a^c f + \int_c^b f = \int_a^b f.$$

($\int$3) [Monotonicity] *If $f \in \mathcal{C}(I)$ and a, $b \in I$ with $a \le b$ and $0 \le f(t)$ for all $t \in [a,b]$, then*

$$0 \le \int_a^b f.$$

If in particular $0 < f(t)$ for all $t \in [a,b]$ and if $a < b$, then

$$0 < \int_a^b f.$$

($\int$4) [Linearity in the integrand] *If f, $g \in \mathcal{C}(I)$, a, $b \in I$, and $\alpha \in \mathbb{R}$, then*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g \quad \text{and} \quad \int_a^b \alpha f = \alpha \int_a^b f.$$

The number $\int_a^b f$ is the DEFINITE INTEGRAL OF f FROM a TO b.

Properties ($\int 4$) encodes the linearity of the integral as an operator on the integrand with the limits of integration fixed, while property ($\int 2$) is its **ADDITIVITY** over subintervals with the integrand fixed. Property ($\int 3$) encodes the idea that a nonnegative function should have a nonnegative integral, while property ($\int 1$) defines the one value of the integral that it most certainly should have from the point of view of area.

Specifically, we can express the definite integral as a limit of Riemann sums—among them, the right-endpoint sums:

$$\int_a^b f = \lim_{n\to\infty} \frac{b-a}{n} \sum_{k=1}^n f\left(a + \frac{k(b-a)}{n}\right). \tag{2.1}$$

That this limit exists is a fundamental result about continuous functions, which we will not prove. From (2.1) we can prove properties ($\int 1$), ($\int 3$), and ($\int 4$) quite easily. Property ($\int 2$) is not so obvious from (2.1), and in fact this property hinges on expressing $\int_a^b f$ as a "limit" of several kinds of Riemann sums, not just the right-endpoint sum. And then there is still the challenge of ensuring that limits of all sorts of "well-behaved" Riemann sums for f (including, but not limited to, left and right endpoint and midpoint sums) all converge to the same number. Moreover, it is plausible that one might want to integrate functions that are not continuous. (We will eventually have to handle this.)

2.2 Problem ($\star$). Let $I \subseteq \mathbb{R}$ be an interval and f, $g \colon I \to \mathbb{R}$ be continuous. Let a, b, $c \in I$ and $\alpha \in \mathbb{R}$. Using only Theorem 2.1, prove the following. *You should not use the Riemann sum formula* (2.1) *at all. The goal is to see how other properties of the integral follow directly from the essential features of Theorem 2.1.*

(i) [Generalization of ($\int 1$)] $\displaystyle\int_a^b \alpha = \alpha(b-a)$

(ii) $\displaystyle\int_a^a f = 0$

(iii) $\displaystyle\int_a^b f = -\int_b^a f$

2.3 Problem (Wholly optional, only if you know induction). Use induction to generalize additivity as follows. Let $I \subseteq \mathbb{R}$ be an interval and $f \colon I \to \mathbb{C}$ be continuous. If $t_0, \ldots, t_n \in I$, then

$$\int_{t_0}^{t_n} f = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} f.$$

2.4 Problem. Let $I \subseteq \mathbb{R}$ be an interval.

(i) Suppose that f, $g \colon I \to \mathbb{R}$ are continuous and a, $b \in \mathbb{R}$ with $a \leq b$. If $f(t) \leq g(t)$ for all $t \in [a, b]$, show that

$$\int_a^b f \le \int_a^b g. \tag{2.2}$$

(ii) Continue to assume a, $b \in I$ with $a \le b$. Prove the **TRIANGLE INEQUALITY**

$$\left|\int_a^b f\right| \le \int_a^b |f|.$$

[Hint: *recall that if x, $r \in \mathbb{R}$ with $r \ge 0$, then $-|x| \le x \le |x|$ and $|x| \le r$ if and only if $-r \le x \le r$. Use this to estimate $f(t)$ in terms of $\pm|f(t)|$ and then apply part (i).*]

(iii) Continue to assume a, $b \in I$ with $a \le b$. Suppose that $f \colon I \to \mathbb{R}$ is continuous and there are m, $M \in \mathbb{R}$ such that $m \le f(t) \le M$ for all $t \in [a, b]$. Show that

$$m(b-a) \le \int_a^b f \le M(b-a). \tag{2.3}$$

(iv) Show that if we remove the hypothesis $a \le b$, then the triangle inequality becomes

$$\left|\int_a^b f\right| \le \left|\int_a^b |f|\right|.$$

Why is the extra absolute value on the right necessary here?

We now have only a handful of results about the definite integral, and yet they are enough to prove the fundamental theorem of calculus. (Conversely, by themselves, they do not help us evaluate integrals more complicated than $\int_a^b \alpha$ for $\alpha \in \mathbb{C}$!) This is our first rigorous verification that an integral gives a meaningful representation of a function. Specifically, integrals represent antiderivatives.

2.5 Theorem (FTC1). *Let $f \colon I \to \mathbb{C}$ be continuous and fix $a \in I$. Define*

$$F \colon I \to \mathbb{C} \colon t \mapsto \int_a^t f$$

Then F is an antiderivative of f on I.

Proof. Fix $t \in I$. We need to show that F is differentiable at t with $F'(t) = f(t)$. That is, we want

$$\lim_{h \to 0} \frac{F(t+h) - F(t)}{h} = f(t),$$

equivalently,

$$\lim_{h \to 0} \frac{F(t+h) - F(t) - hf(t)}{h} = 0.$$

We first compute

$$\begin{aligned} F(t+h) - F(t) &= \int_a^{t+h} f(\tau)\, d\tau - \int_a^t f(\tau)\, d\tau \\ &= \int_a^{t+h} f(\tau)\, d\tau + \int_t^a f(\tau)\, d\tau \\ &= \int_t^{t+h} f(\tau)\, d\tau. \end{aligned}$$

Next,

$$hf(t) = f(t)[(t+h) - t] = f(t) \int_t^{t+h} 1\, d\tau = \int_t^{t+h} f(t)\, d\tau.$$

We then have

$$F(t+h) - F(t) - hf(t) = \int_t^{t+h} f(\tau)\, d\tau - \int_t^{t+h} f(t)\, d\tau = \int_t^{t+h} \big[f(\tau) - f(t)\big]\, d\tau.$$

Note that this is one instance in which using the variable of integration τ clarifies the fact that t is constant here. It therefore suffices to show that

$$\lim_{h \to 0} \frac{1}{h} \int_t^{t+h} \big[f(\tau) - f(t)\big]\, d\tau = 0, \tag{2.4}$$

and we do that in the following lemma. ■

2.6 Problem. Reread, and maybe rewrite, the preceding proof. Identify explicitly each property of or result about integrals that was used without reference.

This is a specific instance of a more general phenomenon in manipulating difference quotients and doing "derivatives by definition." The difference quotient has h in the denominator, and we are sending $h \to 0$, so the denominator is small. A quotient of the form $1/h$ with $h \approx 0$ is large, and large numbers are problematic in analysis. The limit as $h \to 0$ of the difference quotient exists because the numerator is sufficiently small compared to the denominator for the numerator to "cancel out" the effects of that h. In particular, to show

$$\lim_{h \to 0} \frac{F(t+h) - F(t) - hf(t)}{h} = 0,$$

we want the numerator $F(t+h) - F(t) - hf(t)$ to be even smaller than the denominator. *The answer to small denominators is smaller numerators.*

Below is a proof of (2.5) for completeness; I will not require you to know it.

2.7 Lemma. *Let $I \subseteq \mathbb{R}$ be an interval and let $f \colon I \to \mathbb{C}$ be continuous. Then*

$$\lim_{h \to 0} \frac{1}{h} \int_t^{t+h} \big[f(\tau) - f(t)\big] \, d\tau = 0$$

for any $t \in I$.

Proof. We use the squeeze theorem. The triangle inequality implies

$$\left|\frac{1}{h} \int_t^{t+h} \big[f(\tau) - f(t)\big] \, d\tau\right| \le \frac{1}{|h|}|t+h-h| \max_{0 \le s \le 1} |f((1-s)t+s(t+h))-f(t)| = \max_{0 \le s \le 1} |f(t+sh)-f(t)|.$$

We now need to show that

$$\lim_{h \to 0} \max_{0 \le s \le 1} |f(t+sh) - f(t)| = 0.$$

This will involve the definition of continuity.

Let $\epsilon > 0$, so our goal is to find $\delta > 0$ such that if $0 < |h| < \delta$, then

$$\max_{0 \le s \le 1} |f(t+sh) - f(t)| < \epsilon. \tag{2.5}$$

Since f is continuous at t, there is $\delta > 0$ such that if $|t - \tau| < \delta$, then $|f(\tau) - f(t)| < \epsilon$. Suppose $0 < |h| < \delta$. Then if $0 \le s \le 1$, we have

$$|(t+sh) - t| = |sh| \le |h| < \delta,$$

thus (2.5) holds. ■

2.8 Problem. Prove that the left limit in (2.5) holds. What specific changes are needed when $h < 0$?

With FTC1, we can prove a second version that facilitates the computation of integrals via antiderivatives, but first we need to review the mean value theorem.

2.9 Theorem (Mean value). *Let a, $b \in \mathbb{R}$ with $a < b$ and let $f \colon [a,b] \to \mathbb{R}$ be continuous with f differentiable on (a,b). Then there is $c \in (a,b)$ such that*

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

2.10 Problem. **(i)** Let $I \subseteq \mathbb{R}$ be an interval. Suppose that $f \colon I \to \mathbb{R}$ is differentiable with $f'(t) = 0$ for all $t \in I$. Show that f is constant on I. [Hint: *fix $t_0 \in I$ and let $t \in I \setminus \{t_0\}$. Assuming that $t > t_0$, use the mean value theorem to express the difference quotient $(f(t) - f(t_0))/(t - t_0)$ as a derivative, which must be 0. What happens if $t < t_0$?*]

(ii) Give an example of a function f defined on the set $[-1,1] \setminus \{0\}$ that is differentiable with $f'(t) = 0$ for all t but f is not constant. [Hint: *go piecewise.*]

2.11 Corollary (FTC2). *Let $I \subseteq \mathbb{R}$ be an interval and let $f: I \to \mathbb{C}$ be continuous. If F is any antiderivative of f on I, then*

$$\int_a^b f = F(b) - F(a)$$

for all $a, b \in I$.

Proof. Let $G(t) := \int_a^t f$, so G is an antiderivative of f by FTC1. Put $H = G - F$, so $h' = 0$ on I. Since I is an interval, the mean value theorem mplies that H is constant. The most important inputs here are a and b, so we note that $H(a) = H(b)$, and so

$$G(a) - F(a) = G(b) - F(b).$$

But $G(a) = \int_a^a f = 0$, so this rearranges to

$$G(b) = F(b) - F(a),$$

and $G(b) = \int_a^b f$. ■

The fundamental theorems of calculus are, of course, the keys to both substitution and integration by parts, two of the most general techniques for evaluating integrals in terms of simpler functions. Recall that substitution involves turning the more complicated integral $\int_a^b f(\varphi(t))\varphi'(t)\,dt$ into the simpler integral $\int_{\varphi(a)}^{\varphi(b)} f(u)\,du$. For this to make sense, the function φ should be defined and continuous on an interval containing a and b, and f should be defined and continuous on an interval containing $\varphi(a)$ and $\varphi(b)$. But the first integrand should be continuous on I, and that requires φ' to be continuous on I.

2.12 Definition. *Let $I \subseteq \mathbb{R}$ be an interval. A function $\varphi: I \to \mathbb{R}$ is* CONTINUOUSLY DIFFERENTIABLE *if φ is differentiable on I (and thus continuous itself on I) and if also φ' is continuous on I. We denote the set of all continuously differentiable functions on I by $\mathcal{C}^1(I)$.*

2.13 Theorem (Substitution). *Let $I, J \subseteq \mathbb{R}$ be intervals with $a, b \in I$. Let $\varphi \in \mathcal{C}^1(I)$ and $f \in \mathcal{C}(J)$ with $\varphi(t) \in J$ for all $t \in I$. Then*

$$\int_a^b (f \circ \varphi)\varphi' = \int_{\varphi(a)}^{\varphi(b)} f.$$

Proof. We use FTC2. Let F be any antiderivative of f on J (say $F(\tau) = \int_{\varphi(a)}^{\tau} f$). The chain rule implies that $F \circ \varphi$ is an antiderivative of $(f \circ \varphi)\varphi'$; indeed,

$$(F \circ \varphi)' = (F' \circ \varphi)\varphi' = (f \circ \varphi)\varphi'.$$

Then FTC2 implies both

$$\int_{\varphi(a)}^{\varphi(b)} f = F(\varphi(b)) - F(\varphi(a)) \quad \text{and} \quad \int_a^b (f \circ \varphi)\varphi' = (F \circ \varphi)(a) - (F \circ \varphi)(b).$$

These differences are equal. ■

Day 3: Friday, August 16.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 4–6 review first-order linear ODE via integrating factors. This is not the method that we used in class, and I don't think it will be very helpful when we want to apply these ODE techniques to PDE. You might try redoing the textbook's examples with variation of parameters.

A recurring theme of our subsequent applications of integrals will be that we are trying to estimate or control some kind of difference (this is roughly 90% of analysis), and it turns out to be possible to rewrite that difference in a tractable way by introducing an integral. It may be possible to manipulate further terms under consideration by rewriting them as integrals, too. The fundamental identity that we will use in the future is (3.1) below.

3.1 Example. FTC2 allows us to rewrite a functional difference as an integral. When we incorporate substitution, we can get a very simple formula for that difference. Suppose that $I \subseteq \mathbb{R}$ is an interval, $f \in \mathcal{C}^1(I)$, and a, $b \in I$. Then

$$f(b) - f(a) = \int_a^b f'.$$

We will reverse engineer substitution and make the limits of integration simpler and the integrand more complicated. (This turns out to be a good idea.)

Define

$$\varphi \colon [0,1] \to \mathbb{R} \colon t \mapsto (1-t)a + tb = a + (b-a)t.$$

Then $\varphi(0) = a$, $\varphi(1) = b$, and $a \le \varphi(t) \le b$ for all t if $a \le b$, and otherwise $b \le \varphi(t) \le a$ for all t if $b \le a$. (Here is a proof of the first case, assuming $a \le b$. Then $b - a \ge 0$ and $t \ge 0$, so $(b-a)t \ge 0$, thus $a \le a + (b-a)t$. But also $(1-t)a \le (1-t)b$ since $1 - t \ge 0$ and $a \le b$, thus $(1-t)a + tb \le (1-t)b + tb = b$.) In other words, we think of φ as "parametrizing" the line segment between the points a and b on the real line.

Substitution implies

$$\int_a^b f' = \int_0^1 f(\varphi(t))\varphi'(t)\ dt,$$

and we calculate $\varphi'(t) = b - a$. Thus

$$\int_a^b f' = (b-a) \int_0^1 f'(a + (b-a)t)\ dt.$$

In conclusion, if $f \in \mathcal{C}^1(I)$ and a, $b \in I$, then

$$f(b) - f(a) = (b-a)\int_0^1 f'(a+(b-a)t)\ dt. \tag{3.1}$$

This represents explicitly how $f(b)-f(a)$ depends on the quantity $b-a$; if we know how to control f' (maybe f' is bounded on an interval containing a and b), then we have an estimate for the size of $f(b)-f(a)$ in terms of $b-a$. While the mean value theorem would allow us to rewrite $(f(b)-f(a))/(b-a)$ in terms of f', that result is existential and not nearly as explicit as (3.1).

3.2 Problem. Prove the following variant of Example 3.1: if $I \subseteq \mathbb{R}$ is an interval, $f \in \mathcal{C}^1(I)$, and t, $t+h \in I$, then

$$f(t+h) - f(t) = h\int_0^1 f'(t+\tau h)\ d\tau.$$

3.3 Problem. Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is continuous and p-**PERIODIC** for some $p \in \mathbb{R}$, in the sense that $f(t+p) = f(t)$ for all $t \in \mathbb{R}$. Then the integral of f over any interval of length p is the same:

$$\int_a^{a+p} f = \int_0^p f$$

for all $a \in \mathbb{R}$. Give two proofs of this identity as follows.

(i) Define

$$F\colon \mathbb{R} \to \mathbb{R}\colon a \mapsto \int_a^{a+p} f$$

and use FTC1 and the p-periodicity of f to show that $F'(a) = 0$ for all a. Since F is also defined on an interval (the interval here is $\mathbb{R}$), F must be constant.

(ii) First explain why

$$\int_a^{a+p} f = \int_0^p f + \left(\int_p^{a+p} f - \int_0^a f\right).$$

Then substitute $u = t - p$ to show

$$\int_p^{a+p} f = \int_0^a f(t-p)\ dt$$

and use the p-periodicity of f.

3.4 Problem (Integration by parts). Let $I \subseteq \mathbb{R}$ be an interval and f, $g \in \mathcal{C}^1(I)$ and a, $b \in I$. Prove that

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g. \tag{3.2}$$

[Hint: *this is equivalent to an identity for* $\int_a^b (f'g + fg')$ *and that integrand is a perfect derivative by the product rule.*]

3.5 Problem. Suppose that f, f', and f'' are continuous on $\mathbb{R}$; we might say $f \in \mathcal{C}^2(\mathbb{R})$. Suppose also that $f'(0) = 0$ and there is $M > 0$ such that

$$|f''(t)| \le M \text{ for all } t \in \mathbb{R}.$$

Show that

$$|f(x) - f(y)| \le M\big(|x| + |y|\big)|x - y|.$$

By considering the special case $f(x) = x^2$, explain why we might call this a "difference of squares" estimate. [Hint: *use Example 3.1 to rewrite the difference* $f(x)-f(y)$ *as an integral involving* f' *and expose the factor* $x-y$. *That is,* $f(x)-f(y) = (x-y)\mathcal{I}(x,y)$, *where* $\mathcal{I}(x,y)$ *represents this integral. Since* $f'(0) = 0$, *we have* $\mathcal{I}(x,y) = \mathcal{I}(x,y) - (x-y)\int_0^1 f'(0)\ dt$ *Rewrite this difference as an integral from* 0 *to* 1 *of some integrand (which involves* f'*) and apply Example 3.1 again to that integrand so that, in the end,* $\mathcal{I}(x,y)$ *is a double integral involving* f''.]

We have now built enough machinery to study elementary ODE, all of which will reappear in our study of genuine PDE. It will We proceed through three kinds of first-order problems—specifically, all are initial value problems (IVP).

The first is the **DIRECT INTEGRATION** problem

$$\begin{cases} y' = f(t) \\ y(0) = y_0. \end{cases} \tag{3.3}$$

Here $f \in \mathcal{C}(I)$ is a given function and $y_0 \in \mathbb{R}$. Also, $I \subseteq \mathbb{R}$ is an interval with $0 \in I$. The goal is to find a differentiable function y on I such that $y'(t) = f(t)$ for all $t \in I$. (In general, when solving an ODE, one wants a differentiable function y defined on an interval that "makes the ODE true" when values from that interval are substituted in. Also, the domain of a solution should be an interval to reflect the physical ideal that time should be "unbroken"—and because it makes things nice mathematically. In particular, the interval should contain 0 so that we can evaluate $y(0)$ and find $y(0) = y_0$. Last, the derivative should be continuous to reflect the physical ideal that the rates of change do not vary too much—and because it makes things nice mathematically.)

We work backwards. Assume that the problem has a solution y, so $y'(t) = f(t)$ for all $t \in I$. For $t \in I$ fixed, integrate both sides of this equality from 0 to t to find

$$\int_0^t y'(\tau)\ d\tau = \int_0^t f(\tau)\ d\tau.$$

Be very careful to change the variable of integration from t to τ (or anything other than t), since t is now in the limit of integration. We cannot do anything more for the integral on the right, but on the left FTC2 gives

$$\int_0^t y'(\tau)\ d\tau = y(t) - y(0) = y(t) - y_0.$$

That is,

$$y(t) - y_0 = \int_0^t f(\tau)\ d\tau,$$

and so

$$y(t) = y_0 + \int_0^t f(\tau)\ d\tau.$$

Thus if y solves the IVP (3.3), then y has the form above. This is a *uniqueness* result: the only possible solution is this one. But is this really a solution? We check

$$y'(t) = \frac{d}{dt}\left[y_0 + \int_0^t f(\tau)\ d\tau\right] = f(t)$$

by FTC1 and

$$y(0) = y_0 + \int_0^0 f(\tau)\ d\tau = y_0 + 0 = y_0.$$

Yes.

We write this up formally.

3.6 Theorem. *Let $I \subseteq \mathbb{R}$ be an interval with $0 \in I$, let $f \in \mathcal{C}(I)$, and let $y_0 \in \mathbb{R}$. The only solution to*

$$\begin{cases} y' = f(t) \\ y(0) = y_0 \end{cases}$$

is

$$y(t) = y_0 + \int_0^t f(\tau)\ d\tau.$$

3.7 Example. To solve

$$\begin{cases} y' = e^{-t^2} \\ y(0) = 0, \end{cases}$$

we integrate:

$$y(t) = 0 + \int_0^t e^{-\tau^2}\ d\tau = \int_0^t e^{-\tau^2}\ d\tau.$$

We stop here, because we cannot evaluate this integral in terms of "elementary functions." (Long ago with times tables, working with t^2 was hard; then that got easier, but we got older and wiser and sadder and took trig, and working with $\sin(t)$ was hard. Now we are even older, and by the end of the course, working with $\int_0^t f(\tau)\ d\tau$ should feel just as natural as working with any function defined in more "elementary" terms.)

Now we make the ODE more complicated and introduce y-dependence on the left side: we study the **LINEAR FIRST-ORDER** problem

$$\begin{cases} y' = ay + f(t) \\ y(0) = y_0. \end{cases}$$

Again, $f \in \mathcal{C}(I)$ for some interval $I \subseteq \mathbb{R}$ with $0 \in I$, and a, $y_0 \in \mathbb{R}$. The function f is sometimes called the **FORCING** or **DRIVING** term. And, again, the expression $y' = ay + f(t)$ means that we want y to satisfy $y'(t) = ay(t) + f(t)$ for all t in the domain of y (which hopefully will turn out to be I). If $a = 0$, this reduces to a direct integration problem, and it would be nice if our final solution formula will respect that.

To motivate our solution approach, we first suppose $f = 0$ and consider the exponential growth problem

$$y' = ay.$$

Calculus intuition suggests that all solutions have the form $y(t) = Ce^{at}$, where necessarily $C = y(0) = y_0$. We will make that intuition rigorous shortly. The valuable, if surprising, idea that has come down to us through the generations is to replace the constant C with an unknown function u and guess that $y(t) = u(t)e^{at}$ solves the more general problem $y' = ay + f(t)$. This is the first appearance of an **ANSATZ** in this course—that is, we have made a *guess* that a solution has a particular form.

Now the goal is to solve for u. Under the ansatz $y(t) = u(t)e^{at}$, we compute, with the product rule,

$$y'(t) = u'(t)e^{at} + u(t)ae^{at},$$

and we substitute that into our ODE $y' = ay + f(t)$. Then we need

$$u'(t)e^{at} + u(t)ae^{at} = au(t)e^{at} + f(t).$$

The same term $u(t)ae^{at}$ appears on both sides (this is a hint that we made the right ansatz), and we subtract it, leaving

$$u'(t)e^{at} = f(t).$$

We solve for things by getting them by themselves, so divide to find

$$u'(t) = e^{-at}f(t).$$

This is an ODE for u, but it would be nice if it had an initial condition. We know $y(t) = u(t)e^{at}$ and $y(0) = y_0$, so

$$y_0 = y(0) = u(0)e^{a0} = u(0).$$

That is, u must solve the direct integration problem

$$\begin{cases} u' = e^{-at}f(t) \\ u(0) = y_0, \end{cases}$$

and so, from our previous work,

$$u(t) = y_0 + \int_0^t e^{-a\tau} f(\tau)\, d\tau.$$

Returning to the ansatz $y(t) = u(t)e^{at}$, we have

$$y(t) = e^{at}\left(y_0 + \int_0^t e^{-a\tau} f(\tau)\, d\tau\right),$$

and so we have proved another theorem.

3.8 Theorem. *Let $f \in \mathcal{C}(I)$ for some interval $I \subseteq \mathbb{R}$ with $0 \in I$ and a, $y_0 \in \mathbb{R}$. Then the only solution to*

$$\begin{cases} y' = ay + f(t) \\ y(0) = y_0 \end{cases} \tag{3.4}$$

is

$$y(t) = e^{at}\left(y_0 + \int_0^t e^{-a\tau} f(\tau)\, d\tau\right). \tag{3.5}$$

Is it?

3.9 Problem. **(i)** Check that the function y in (3.5) actually solves (3.4). (Does y satisfy $y'(t) = ay(t) + f(t)$ for all t in some interval containing 0? Do we have $y(0) = y_0$? Is y' continuous?)

(ii) Check that we recover the direct integration result of Theorem 3.6 from Theorem 3.8 when $a = 0$.

By the way, the ODE $y' = ay + f(t)$ is sometimes more precisely called a **FIRST-ORDER CONSTANT COEFFICIENT LINEAR ODE**. It is constant-coefficient because the coefficient a on y is a constant real number. This ODE is **HOMOGENEOUS** if $f(t) = 0$ for all t and otherwise **NONHOMOGENEOUS**. The uniqueness part of Theorem 3.8 proves that all solutions to $y' = ay$ have the form $y(t) = y(0)e^{at}$. Sometimes this is established with separation of variables, which we will consider shortly.

3.10 Example. We study

$$\begin{cases} y' = 2y + 3e^{-4t} \\ y(0) = 1, \end{cases}$$

and rather than just use the formula from (3.5), we repeat the "variation of parameters" argument with the concrete data at hand. The corresponding homogeneous problem is $y' = 2y$, which has the solutions $y(t) = Ce^{2t}$, and so we guess that our nonhomogeneous problem has the solution $y(t) = u(t)e^{2t}$. Substituting this into both sides of the ODE, we

want
$$u'(t)e^{2t} + u(t)(2e^{2t}) = 2u(t)e^{2t} + 3e^{-4t},$$
thus
$$u'(t)e^{2t} = 3e^{-4t},$$
and so
$$u'(t) = 3e^{-6t}.$$
With the initial condition $u(0) = y(0) = 1$, this is the direct integration problem
$$\begin{cases} u' = 3e^{-6t} \\ u(0) = 1, \end{cases}$$
and the solution to that is
$$u(t) = 1 + \int_0^t 3e^{-6\tau}\, d\tau = 1 + \frac{3e^{-6\tau}}{-6}\bigg|_{\tau=0}^{\tau=t} = 1 + \frac{3e^{-6t} - 3}{-6} = \frac{3}{2} + \frac{e^{-6t}}{2}.$$
Thus the solution to the original IVP is
$$y(t) = e^{2t}\left(\frac{3}{2} + \frac{e^{-6t}}{2}\right).$$

Day 4: Monday, August 19.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas
Pages 2–3 review separation of variables for ODE.

Our experience with ODE in general, and our concrete work with the linear problem, tell us that initial conditions should determine solutions uniquely. But sometimes in both ODE and PDE, one is less concerned with the initial state of the solution and more with its behavior at a "boundary." For example, what is the long-time asymptotic behavior of a solution? Does it have a limit at infinity, or does it settle down into some coherent shape? Here is one toy problem of how boundary behavior determines the solution.

4.1 Example. Let $f \in \mathcal{C}(\mathbb{R})$ and $a \in \mathbb{R}$. We know that all solutions to
$$y' = ay + f(t)$$
have the form
$$y(t) = e^{at}\left(y(0) + \int_0^t e^{-a\tau} f(\tau)\, d\tau\right). \tag{4.1}$$
What, if any, choices for the initial condition $y(0)$ guarantee
$$\lim_{t \to \infty} y(t) = 0?$$

We consider some cases on a, starting with the easiest $a = 0$. Here

$$y(t) = y(0) + \int_0^t f(\tau)\ d\tau,$$

so we want

$$\lim_{t\to\infty} \left(y(0) + \int_0^t f(\tau)\ d\tau \right) = 0.$$

Since

$$\lim_{t\to\infty} \int_0^t f(\tau)\ d\tau = \int_0^\infty f(\tau)\ d\tau,$$

this suggests

$$y(0) = -\int_0^\infty f(\tau)\ d\tau.$$

This both specifies the value of $y(0)$ and adds an additional constraint into our problem: f must be improperly integrable on $[0, \infty)$. With this choice of $y(0)$, we have

$$y(t) = -\int_0^\infty f(\tau)\ d\tau + \int_0^t f(\tau)\ d\tau = -\int_t^\infty f(\tau)\ d\tau,$$

and we expect from calculus that

$$\lim_{t\to\infty} \int_t^\infty f(\tau)\ d\tau = 0.$$

Now we consider the case $a > 0$. From (4.1), we note that our solution is the product of two functions, one of which blows up as $t \to \infty$ (since $\lim_{t\to\infty} e^{at} = \infty$ for $a > 0$). We probably want the other factor in the product to tend to 0 as $t \to \infty$; if that factor limited, say, to a nonzero constant, then the whole limit would be ∞ times that constant, which would definitely not be 0. Indeed, we can see this using the definition of limit: if we assume $\lim_{t\to\infty} y(t) = 0$, then there is $M > 0$ such that if $t \geq M$, then $|y(t)| \leq 1$. With y given by (4.1), we find

$$\left| y(0) + \int_0^t e^{-a\tau} f(\tau)\ d\tau \right| \leq e^{-at}.$$

Since $a > 0$, this inequality and the squeeze theorem imply

$$\lim_{t\to\infty} \left(y(0) + \int_0^t e^{-a\tau} f(\tau)\ d\tau \right) = 0,$$

and thus

$$y(0) = -\int_0^\infty e^{-a\tau} f(\tau)\ d\tau \quad \text{and} \quad y(t) = -e^{at} \int_t^\infty e^{-a\tau} f(\tau)\ d\tau.$$

This directly generalizes the case of $a = 0$. In fact, we get a little more freedom here, in that for $a > 0$, it is easier for $\int_0^\infty e^{-a\tau} f(\tau)\ d\tau$ to exist (see below).

We leave the case $a < 0$ as a (possibly surprising) exercise.

4.2 Problem. Suppose that y solves $y' = ay + f(t)$ with $a < 0$ and $\lim_{t\to\infty} y(t) = 0$. As in the previous example, there is $M > 0$ such that for all $t \ge M$, we have

$$\left| y(0) + \int_0^t e^{-a\tau} f(\tau)\, d\tau \right| \le e^{-at}.$$

However, since $-a > 0$, this does not imply any convergence of the integral term to $y(0)$.

Consider the concrete problem

$$y' = -2y + 3e^{-t}.$$

Show that every solution to this problem satisfies $\lim_{t\to\infty} y(t) = 0$, and thus the boundary condition is of no help in specifying the initial condition.

4.3 Problem. We clarify a remark from the previous example about improper integrals. In the following, let $a > 0$.

(i) Suppose that $f \in \mathcal{C}(\mathbb{R})$ is absolutely integrable on $[0, \infty)$; that is,

$$\int_0^\infty |f| := \lim_{b\to\infty} \int_0^b |f|$$

converges. Show that $\int_0^\infty e^{-a\tau} f(\tau)\, d\tau$ converges as well.

(ii) Suppose that $f \in \mathcal{C}(\mathbb{R})$ is bounded on $[0, \infty)$; that is, there is $M > 0$ such that

$$|f(t)| \le M$$

for all $t \ge 0$. Show that $\int_0^\infty e^{-a\tau} f(\tau)\, d\tau$ still converges. Give an example to show that f need not be absolutely integrable on $[0, \infty)$.

Now we move to **separable** ODE. Before defining and solving this kind of ODE in general, we do a pedestrian, but illustrative, example.

4.4 Example. We study

$$\begin{cases} y' = y^2 \\ y(0) = y_0. \end{cases}$$

If $y_0 = 0$, then we can take $y(t) = 0$ for all t to get a solution; indeed, $y'(t) = 0$ and $(y(t))^2 = 0^2 = 0$ for all t.

Otherwise, suppose $y_0 \neq 0$. We expect that $y \neq 0$, so we can divide to find

$$y' \frac{1}{y^2} = 1.$$

This is a time when Leibniz notation is more evocative. Write

$$\frac{dy}{dt}\frac{1}{y^2} = 1,$$

pretend that dy and dt are parts of a fraction, and separate variables:

$$\frac{dy}{y^2} = dt.$$

Slap (...indefinite...) integrals on both sides to get

$$\int \frac{dy}{y^2} = \int dt$$

and antidifferentiate to find

$$-\frac{1}{y} = t + C.$$

Solve for y:

$$y(t) = -\frac{1}{t+C}.$$

Last, solve for C: we want $y(0) = y_0$, so

$$-\frac{1}{0+C} = y_0,$$

and then

$$C = -\frac{1}{y_0}.$$

Since $y_0 \neq 0$, we have no qualms about division here. All together,

$$y(t) = -\frac{1}{t - \dfrac{1}{y_0}} = \frac{1}{y_0^{-1} - t}.$$

Recalling that a formula alone is not sufficient to describe a function, we also establish the domain of this solution. As a formula alone, y above is defined on $\mathbb{R} \setminus \{y_0^{-1}\}$, but that is not an interval. Remember that we want the domain of the solution to this IVP to be an interval containing 0. The largest intervals in $\mathbb{R} \setminus \{y_0^{-1}\}$ (go big or go home) are $(-\infty, y_0^{-1})$ and (y_0^{-1}, ∞). Which interval we use depends on whether $y_0 < 0$ or $y_0 > 0$; if $y_0 < 0$, then $y_0^{-1} < 0$, too, so $0 \notin (-\infty, y_0^{-1})$ but $0 \in (y_0^{-1}, \infty)$. The reverse holds when $y_0 > 0$, and so there we take the domain to be $(-\infty, y_0^{-1})$.

Both situations illustrate a "blow-up in finite time." If we send t to the boundary of the domain, then the solution explodes to $\pm\infty$. For example, when $y_0 > 0$, the solution is defined on $(-\infty, y_0^{-1})$, and we have

$$\lim_{t \to (y_0^{-1})^-} y(t) = \lim_{t \to (y_0^{-1})^-} \frac{1}{y_0^{-1} - t} = \infty.$$

Note that here we are only using the limit from the left.

Now we generalize this work substantially. Let f and g be continuous functions (quite possibly on different subintervals of $\mathbb{R}$), and consider the IVP

$$\begin{cases} y' = f(t)g(t) \\ y(0) = y_0. \end{cases}$$

If $g(y_0) = 0$, then we claim that $y(t) = y_0$ is a solution to this IVP, which we call an EQUILIBRIUM SOLUTION.

4.5 Problem. Prove that.

Suppose that $g(y_0) \neq 0$. Since g is continuous, for y "close to" y_0, we have $g(y) \neq 0$. In fact, $g(y)$ is either positive for all y close to y_0 or negative for all y close to y_0.

Now we work backward. Assume that y solves this IVP with $g(y_0) \neq 0$. Since y is continuous and $y(0) = y_0$, for t close to 0, we have $y(t)$ close to y_0, and thus $g(y(t)) \neq 0$. We can then divide to find that for t close to 0, y must also satisfy

$$\frac{y'(t)}{g(y(t))} = f(t).$$

This is the heart of separation of variables: division. And division is only possible when the denominator is nonzero. We integrate both sides from 0 to t, still keeping t close to 0:

$$\int_0^t \frac{y'(\tau)}{g(y(\tau))} \, d\tau = \int_0^t f(\tau) \, d\tau. \tag{4.2}$$

There is not much more that we can say about the integral on the right, but on the left we take the composition $g \circ y$ as a hint to substitute $u = y(t)$. This yields

$$\int_0^t \frac{y'(\tau)}{g(y(\tau))} \, d\tau = \int_{y(0)}^{y(t)} \frac{du}{g(u)} = \int_{y_0}^{y(t)} \frac{du}{g(u)}. \tag{4.3}$$

Combining (4.2) and (4.3), we conclude that if y solves the separable IVP with $y_0 \neq 0$, then for t sufficiently close to 0, we have

$$\int_{y_0}^{y(t)} \frac{du}{g(u)} = \int_0^t f(\tau) \, d\tau.$$

We rewrite this one more time. Put

$$H(y,t) := \int_{y_0}^{y} \frac{du}{g(u)} - \int_0^t f(\tau) \, d\tau.$$

Here the domain of H is all y such that $g(u) \neq 0$ for u between y_0 and y and all t such that f is defined between 0 and t. Thus if y solves the separable IVP with $y_0 \neq 0$, then for t sufficiently close to 0, we have

$$H(y(t), t) = 0.$$

This is an IMPLICIT EQUATION for y.

It would be nice if we could reverse our logic and conclude that if $H(y(t), t) = 0$, then y solves the separable IVP. More generally, why should we be able to solve $H(y,t) = 0$? That we can is the content of the following problem.

4.6 Problem. The IMPLICT FUNCTION THEOREM says the following. Let a, b, $c \in \mathbb{R}$ with $a < b$ and $c > 0$. Let H be defined on $\mathcal{D} := \{(y,t) \in \mathbb{R}^2 \mid a < y < b,\ |t| < c\}$, and suppose that the partial derivatives H_y and H_t exist and are continuous on $\mathcal{D}$. Suppose that $H(y_0, 0) = 0$ for some $y_0 \in (a,b)$ with $H_y(y_0,0) \neq 0$. Then there exist δ, $\epsilon > 0$ and a continuously differentiable function $Y\colon (-\delta,\delta) \to (y_0-\epsilon, y_0+\epsilon)$ such that $H(y,t) = 0$ for $|t| < \delta$ and $|y - y_0| < \epsilon$ if and only if $y = Y(t)$. In particular, $Y(0) = y_0$.

We use the implicit function theorem to prove the existence of solutions to separable IVP.

(i) For practice, consider $H(y,t) := y^2 + t^2 - 1$. Check that $H(1,0) = 0$ and $H_y(1,0) \neq 0$ and conclude that $H(Y(t),t) = 0$ for some function Y defined on a subinterval $(-\delta,\delta)$. Then do algebra and find an explicit formula for Y.

(ii) In this part and the following, consider the separable problem

$$\begin{cases} y' = f(t)g(y) \\ y(0) = y_0, \end{cases}$$

where g is continuous on (a,b), f is continuous on $(-c,c)$, and $y_0 \in (a,b)$ with $g(y_0) \neq 0$. Without loss of generality, we will assume $g(y) > 0$ for $y \in (a,b)$. Our goal is to solve the implicit equation

$$H(y,t) := \int_{y_0}^{y} \frac{du}{g(u)} - \int_0^t f(\tau)\, d\tau = 0$$

First check that $H(y_0,0) = 0$ and $H_y(y_0,0) \neq 0$, and obtain the existence of a function Y meeting the conclusions of the implicit function theorem with $Y(0) = 1$. (In particular, we get $Y(0) = y_0$.)

(iii) Now we show that Y solves the original ODE. Differentiate the identity $H(Y(t),t) = 0$ with respect to t, use the multivariable chain rule and FTC1, and conclude that $Y' = f(t)g(Y)$.

(iv) It turns out that just from $H(Y(0),0) = 0$ we can obtain $Y(0) = y_0$, even without the implicit function theorem. To see this, use properties of integrals to show that $H(Y(0),0) = 0$ implies

$$\int_{y_0}^{Y(0)} \frac{du}{g(u)} = 0.$$

Suppose that $Y(0) \neq y_0$ and use the monotonicity of the integral and the fact that $g(u) > 0$ for u between y_0 and $Y(0)$ to obtain a contradiction.

Day 5: Wednesday, August 21.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

There are many examples of second-order constant-coefficient linear ODE on pp. 6–13. Example 8, while worth reading, is probably more complicated than any problem that we will encounter at this level for some time. (If we need variation of parameters for second-order problems, we will review it later.)

We do one more separable ODE to illustrate the utility of definite integrals.

5.1 Example. We know full well that the solution to the exponential growth problem

$$\begin{cases} y' = ry \\ y(0) = y_0 \end{cases}$$

is $y(t) = y_0 e^{rt}$. (Here $r \in \mathbb{R}$ is a fixed parameter.) Suppose we did not know this from calculus. How would we see it at the level of separation of variables?

If $y_0 = 0$, then $y(t) = 0$ is the equilibrium solution, so assume $y_0 \neq 0$. Working backwards, if y solves this IVP, then since $y(t) = y_0 \neq 0$, we have $y(t) \neq 0$ for $t \approx 0$. More precisely—and this will be important later—either $y(t) > 0$ for $t \approx 0$ or $y(t) < 0$ for $t \approx 0$.

We divide to find

$$\frac{y'(t)}{y(t)} = r$$

and integrate, for $t \approx 0$, to find

$$\int_0^t \frac{y'(\tau)}{y(\tau)}\, d\tau = \int_0^t r\, d\tau = rt.$$

We substitute $u = y(\tau)$ on the left:

$$\int_0^t \frac{y'(\tau)}{y(\tau)}\, d\tau = \int_{y(0)}^{y(t)} \frac{du}{u} = \int_{y_0}^{y(t)} \frac{du}{u} = \ln(|y(t)|) - \ln(|y_0|) = \ln\left(\left|\frac{y(t)}{y_0}\right|\right).$$

Here we are using the notion that the natural logarithm is the integral

$$\ln(t) := \int_1^t \frac{d\tau}{\tau}.$$

We obtain

$$\ln\left(\left|\frac{y(t)}{y_0}\right|\right) = rt$$

and exponentiate to find

$$\left|\frac{y(t)}{y_0}\right| = e^{rt},$$

so

$$|y(t)| = |y_0|e^{rt}. \tag{5.1}$$

At this point our algebraic (and ODE) experience would tell us $y(t) = \pm y_0 e^{rt}$, and then we would handwave the $\pm$ away. However, we can be more precise.

Recall that we are assuming $y_0 \neq 0$, so either $y_0 > 0$ or $y_0 < 0$. If $y_0 > 0$, when $t \approx 0$, we have $y(t) > 0$, too. Then $|y(t)| = y(t)$ and $|y_0| = y_0$, so (5.1) becomes $y(t) = y_0 e^{rt}$. If $y_0 < 0$, when $t \approx 0$, we have $y(t) < 0$, too. Then $|y(t)| = -y(t)$ and $y_0 = -y_0$, so (5.1) becomes $-y(t) = -y_0 e^{rt}$, thus $y(t) = y_0 e^{rt}$.

The moral of the story is that using definite integrals and initial values cuts down on much of the nonsense of the constant of integration and the absolute value manipulations that appear in a first ODE course when separating variables for this problem.

The final kind of ODE that we need to review for this course is the second-order constant-coefficient linear problem, which reads

$$ay'' + by' + cy = f(t),$$

with a, b, $c \in \mathbb{R}$, $a \neq 0$ (so that the problem is genuinely second-order), and f continuous on some interval containing 0. We largely focus here on the homogeneous case of $f = 0$. Then one studies the **CHARACTERISTIC EQUATION**

$$a\lambda^2 + b\lambda + c = 0$$

and develops solution patterns based on the root structure. They are the following.

Root structure	**Solution pattern**
Two distinct real roots $\lambda_1 \neq \lambda_2$	$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$
One repeated real root λ_0	$y(t) = c_1 e^{\lambda_0 t} + c_2 t e^{\lambda_0 t}$
Two complex conjugate roots $\alpha \pm i\beta$ ($\beta \neq 0$)	$y(t) = e^{\alpha t}\big(c_1 \cos(\beta t) + c_2 \sin(\beta t)\big)$

That any of these solution patterns actually works can be checked by directly substituting it into the ODE and using the structure of a, b, and c that results from the root pattern. (For example, in the repeated real root case one has $b^2 - 4ac = 0$, thus $c = b^2/4a$, and also $\lambda_0 = -b/2$.) The proof that *all* solutions fall into these patterns is more involved and we do not pursue it, outside of some special cases that illustrate "energy" methods that will subsequently be useful. Instead, we do some canonical examples.

5.2 Example. **(i)** The characteristic equation of $y'' - y = 0$ is $\lambda^2 - 1 = 0$. Factoring the difference of perfect squares, we have $\lambda = \pm 1$. These are distinct real roots, so all solutions are $y(t) = c_1 e^t + c_2 e^{-t}$.

(ii) The characteristic equation of $y'' = 0$ is $\lambda^2 = 0$, so $\lambda = 0$. This is a repeated real root, so all solutions are $y(t) = c_1 e^{0t} + c_2 t e^{0t} = c_1 + c_2 t$. (Of course, we could directly integrate twice to get the same result.)

(iii) The characteristic equation of $y'' + y = 0$ is $\lambda^2 + 1 = 0$, so $\lambda^2 = -1$ and thus $\lambda = \pm i$. These are complex conjugate roots with $\alpha = 0$ (which is certainly allowed) and $\beta = 1$. All solutions are $y(t) = e^{0t}\big(c_1\cos(t) + c_2\sin(t)\big) = c_1\cos(t) + c_2\sin(t)$.

Now we show uniqueness of solutions to an IVP in a special form. The result is unsurprising; the key trick will be useful. Suppose that u and v both solve the IVP

$$\begin{cases} y'' + \lambda^2 y = f(t) \\ y(0) = y_0 \\ y'(0) = y_1 \end{cases}$$

for $\lambda \neq 0$, y_0, $y_1 \in \mathbb{R}$, and some function f. (That is, taking $y = u$ and $y = v$ produces equalities throughout the IVP.) Set $z(t) := u(t) - v(t)$. We will show that $z(t) = 0$ for all t, which forces $u = v$.

5.3 Problem. Use the linearity of the IVP to show that z satisfies

$$\begin{cases} z'' + \lambda^2 z = 0 \\ z(0) = 0 \\ z'(0) = 0. \end{cases}$$

Now here is the trick: multiply both sides of $z'' + \lambda^2 z = 0$ by z' to get

$$z''z' + \lambda^2 zz' = 0. \tag{5.2}$$

A moment's calculation gives

$$\partial_t[z^2] = 2zz'$$

by the chain rule, thus

$$\partial_t\left[\frac{z^2}{2}\right] = zz'. \tag{5.3}$$

Likewise,

$$\partial_t[(z')^2] = 2z'z'',$$

so

$$z''z' = \partial_t\left[\frac{(z')^2}{2}\right]. \tag{5.4}$$

We combine (5.2), (5.3), and (5.4) to find

$$\partial_t[(z')^2 + \lambda^2 z^2] = 0,$$

so the function $(z')^2 + \lambda^2 z^2$ is constant. We know its value at precisely one point: $t = 0$. Thus

$$(z'(t))^2 + \lambda^2(z(t))^2 = (z'(0))^2 + \lambda^2(z(0))^2 = 0$$

for all t, since $z'(0) = z(0) = 0$ by Problem 5.3.

Now here is another trick: if a, $b \in \mathbb{R}$, then

$$0 \le a^2 \le a^2 + b^2.$$

So if $a^2 + b^2 = 0$, then

$$0 \le a^2 \le 0,$$

and so $a^2 = 0$, thus $a = 0$. With $a = \lambda z(t)$ and $b = z'(t)$, we find $\lambda z(t) = 0$ for all t, and since $\lambda \ne 0$, this gives $z(t) = 0$ for all t, as desired.

5.4 Problem. Generalize the preceding work as follows. Let $\mathcal{V} \in \mathcal{C}^1(\mathbb{R})$ with $\mathcal{V}(r) > 0$ for all $r \ne 0$, $\mathcal{V}(0) = 0$, and $\mathcal{V}'(0) = 0$. Show that the only solution to the IVP

$$\begin{cases} y'' + \mathcal{V}'(y) = 0 \\ y(0) = 0 \\ y'(0) = 0 \end{cases}$$

is $y = 0$. [Hint: *for existence, be sure to explain why $y = 0$ is actually a solution. For uniqueness, suppose that y solves the IVP, multiply by y', and obtain that $(y')^2/2 + \mathcal{V}(y)$ is constant. What is its value? What does that tell you about $\mathcal{V}(y)$?*]

5.5 Problem. Let $\lambda \ne 0$. This problem connects solutions to $y'' + \lambda^2 y = 0$ and $y'' - \lambda^2 y = 0$ via hyperbolic trig functions.

(i) Show that the only solution to the IVP

$$\begin{cases} y'' + \lambda^2 y = 0 \\ y(0) = y_0 \\ y'(0) = y_1 \end{cases}$$

is

$$y(t) = y_0 \cos(\lambda t) + \frac{y_1}{\lambda} \sin(\lambda t).$$

For t fixed, what is $\lim_{\lambda \to 0} y(t)$? [Hint: *L'Hospital's rule will be helpful.*] Compare this to part (ii) of Example 5.2.

(ii) The HYPERBOLIC SINE AND COSINE, respectively, are

$$\sinh(t) := \frac{e^t - e^{-t}}{2} \quad \text{and} \quad \cosh(t) := \frac{e^t + e^{-t}}{2}.$$

Show that the only solution to the IVP

$$\begin{cases} y'' - \lambda^2 y = 0 \\ y(0) = y_0 \\ y'(0) = y_1 \end{cases}$$

is

$$y(t) = y_0 \cosh(\lambda t) + \frac{y_1}{\lambda} \sinh(\lambda t).$$

[Hint: *suppose that u and v solve this IVP and define $w(t) := u(t/\lambda) - v(t/\lambda)$. Show that $w'' - w = 0$ with $w(0) = w'(0) = 0$. Show next that $(w'+w)' = w'+w$, so $w'(t)+w(t) = Ce^t$ for some constant C. Take $t = 0$ to conclude $C = 0$, so $w' = -w$. Thus $w(t) = Ke^{-t}$ for some constant K; take $t = 0$ again to conclude $K = 0$.*] For t fixed, what is $\lim_{\lambda \to 0} y(t)$? Again, compare this to part (ii) of Example 5.2.

Day 6: Friday, August 23.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Section 1.2 contains a variety of PDE that are, more or less, really ODE (or that can be solved with ODE ideas and no fancy new PDE ones). Examples 1 through 6 are worth reading and attempting; pay no attention to the "general" vs. "generic" distinction for solutions. Pages 48–50 focus specifically on PDE that are ODE. A version of the transport equation is derived on pp. 85–86 under "An application to gas flow."

We are finally ready to study some PDE, although the first few will be artificial PDE that are really ODE. We begin with a convention.

6.1 Undefinition. *A function u is a* SOLUTION *to a PDE if u solves that PDE at each point in its domain and if every partial derivative of u up to the highest-order derivative in the PDE exists and is continuous.*

We will solve the PDE

$$u_t = 0$$

in a moment, with $u = u(x,t)$. The partial derivatives u_t and u_x must exist and be continuous; even though u_x does not appear in the PDE, we still require its existence and continuity. Similarly, a solution $u = u(x,t)$ to the heat equation

$$u_t = u_{xx}$$

must have continuous partial derivatives u_t, u_x, u_{tt}, u_{xx}, u_{xt}, and u_{tx}, even though only two of these actually appear in the PDE. (Recall that since the mixed partials are continuous, they are equal: $u_{xt} = u_{tx}$.)

Here are some examples in which we use ODE techniques. The major change is that initial data will now be functions, and we will have to consider the regularity of those functions.

6.2 Example. If $u = u(x,t)$ satisfies $u_t = 0$, then we think that u is constant in t, so u should be a function of x alone, maybe $u(x,t) = f(x)$. Indeed, $\partial_t[f(x)] = 0$. We might see this at the level of integrals: since $u_t(x,t) = 0$ for all t, we have

$$0 = \int_0^t u_t(x,\tau)\ d\tau = u(x,t) - u(x,0),$$

thus

$$u(x,t) = u(x,0),$$

and u "ignores" the contribution of t. This, however, neglects the domain of u; the integral calculation is valid if u is defined on the interval containing 0 and t, but what if it is not?

6.3 Problem. Draw the set

$$\mathcal{D} := \{(x,t) \in \mathbb{R}^2 \mid x^2 + t^2 < 1 \text{ or } (x-3)^2 + t^2 < 1\}$$

and construct a function u on $\mathcal{D}$ that solves $u_t = 0$ with u not constant in t.

6.4 Example. Cautioned by that domain problem, we solve

$$\begin{cases} u_t = u, & -\infty < x,\ t < \infty \\ u(x,0) = f(x), & -\infty < x < \infty. \end{cases}$$

This is really a "family" of ODE "indexed" by x; for each x, we want to solve

$$\begin{cases} u_t(x,t) = u(x,t), & -\infty < t < \infty \\ u(x,0) = f(x). \end{cases}$$

Of course this is the same as

$$\begin{cases} y' = y \\ y(0) = y_0, \end{cases}$$

and so our solution to the PDE is

$$u(x,t) = f(x)e^t.$$

Since u_x must exist and be continuous, we want $f \in \mathcal{C}^1(\mathbb{R})$. Thus we need to be more careful and restrictive with the initial data for a PDE than we were for an ODE.

6.5 Example. We solve

$$\begin{cases} u_{tt} + x^2 u = 0 & -\infty < x,\ t < \infty \\ u(x,0) = f(x), & -\infty < x < \infty \\ u_t(x,0) = g(x), & -\infty < x < \infty. \end{cases}$$

When $x = 0$, this is effectively the ODE $y'' = 0$, and we can solve that via direct integration.

That is, for $x = 0$, we have

$$u(0, t) = g(0)t + f(0).$$

For $x \neq 0$, we can use Problem 5.5 (with x playing the role of λ) to write

$$u(x, t) = f(x)\cos(xt) + g(x)\frac{\sin(xt)}{x}.$$

Of course, we want

$$\lim_{x \to 0} u(x, t) = u(0, t),$$

and Problem 5.5 assures us that that is true. Finally, we need f, $g \in \mathcal{C}^2(\mathbb{R})$ because we want u_{xx} to exist and be continuous.

These PDE were really ODE because derivatives with respect to only one variable appeared in them. Now we derive from (nebulous) physical principles our first genuine PDE.

Consider a substance that moves or flows along an infinite path parallel to a horizontal line—maybe a pollutant moving through a stream, maybe cars along a road, maybe gas through a pipe. We think of the path as the real line $\mathbb{R} = (-\infty, \infty)$. The substance enters the path from "far away" on the left and flows to the right; once on the path, the substance does not leave the path, and there are no other sources for the substance along the path. (If the path is a road and the substance is cars, there are no on or off ramps.)

Suppose that we measure position along this path by the variable x, and let $u(x, t)$ be the density of the substance at position x and time t. Usually density = mass/volume, but this may feel strange—how can there be volume at a single point in space? We will adopt the one-dimensional point of view that u measures density via the approximation

$$u(x, t) \approx \frac{\text{the amount of the substance between points } x - h \text{ and } x + h \text{ on the path at time } t}{2h}$$

when $h > 0$ is small.

Let $a < b$. A Riemann sum argument suggests that the amount of the substance between position a and position b on the path is

$$\int_a^b u(x, t) \, dx,$$

and we will take this as the definition of "amount."

6.6 Remark. *Here is that argument. Divide the interval $[a, b]$ into the n subintervals $[x_k, x_{k+1}]$ for $k = 0, \ldots, n - 1$ with*

$$x_k := a + \left(\frac{b - a}{n}\right) k.$$

For $x_k \leq x \leq x_{k+1}$, we have $u(x, t) \approx u(x_k, t)$ if n is large and the subinterval is small (and if u is continuous).

$$\frac{u(x_k, t) \text{ amount of substance}}{\text{unit length}} \times (x_{k+1} - x_k) \text{ length} = u(x_k, t)(x_{k+1} - x_k) \text{ amount.}$$

So, over all of $[a, b]$*, there is approximately*

$$\sum_{k=0}^{n} u(x_k, t)(x_{k+1} - x_k) \text{ amount},$$

and this is a Riemann sum for the integral $\int_a^b u(x, t)\ dx$*.*

Thus the rate of change of the amount of the substance between positions a and b at time t is

$$\partial_t \left[\int_a^b u(x, t)\ dx \right].$$

Without knowing u, this is not a very helpful quantity, but the following is true. For a "sufficiently nice" function u, we have

$$\partial_t \left[\int_a^b u(x, t)\ dx \right] = \int_a^b u_t(x, t)\ dx. \tag{6.1}$$

This equality is called "differentiating under the integral," and it will be a hugely useful technique for us in the future. We will revisit it and discuss it at length. For now, just assume that it is valid.

A partial derivative has entered the stage, and we should be happy. But we have nothing to compare this partial derivative to, no equality, and so we do not yet have a PDE. We therefore introduce something new: let $q(x, t)$ be the rate of change of the amount of this substance at position x and time t. We call q the **FLUX** of this substance. This, too, is a little change: is the substance zero-dimensional so that it can exist at a single point in space? We adopt another one-dimensional point of view: q measures this rate of change if

$$q(x, t) \approx \frac{\text{the amount of substance that passes through point } x \text{ between times } t - k \text{ and } t + k}{2k}$$

for $k > 0$.

Consider any "interval" $[a, b]$ on the path. The substance enters the interval at position a with rate $q(a, t)$ and leaves the interval at position b with rate $q(b, t)$. Remember that the substance is not added to or removed from the path at all, so entering from the left and leaving from the right is the only way that the amount of the substance in $[a, b]$ can change. Thus the rate of change of the amount of the substance in $[a, b]$ is "rate in minus rate out" (a good paradigm for population models in ODE!), and so that rate is

$$q(a, t) - q(b, t) = -\int_a^b q_x(x, t)\ dx.$$

Here we have rewritten the difference as an integral (a good trick!) to make things consonant with our previous calculation of the rate of change in (6.1). That is,

$$\int_a^b u_t(x, t)\ dx = -\int_a^b q_x(x, t)\ dx,$$

and so

$$\int_a^b \left[u_t(x,t) + q_x(x,t)\right]\, dx = 0. \tag{6.2}$$

Now here is a marvelous fact about integrals.

6.7 Problem. Let $I \subseteq \mathbb{R}$ be an interval and let $g \in \mathcal{C}(I)$ such that

$$\int_a^b g = 0$$

for all a, $b \in I$ with $a \le b$. Prove that $g(x) = 0$ for all $x \in I$. [Hint: *fix $a \in I$ and let $G(x) := \int_a^x g$. What do you know about G'? Calculate it in two ways.*]

We combine this result and the fact that a and b were arbitrary to conclude from (6.2) that

$$u_t(x,t) + q_x(x,t) = 0$$

for all x and t. This is good, because it is an equation, and a PDE at that, but not so good in that we have two quantities (density and flux) and only one equation—not usually a recipe for success. One way of proceeding is to assume that flux is somehow related to density, which is not unreasonable—surely the density should somehow affect the rate of change of the amount of the substance. Perhaps the simplest relation is linear: assume

$$q(x,t) = cu(x,t)$$

for some constant c. Then u must satisfy

$$u_t + cu_x = 0.$$

This is (one version of) the **TRANSPORT EQUATION**, and we will study it in detail.

Day 7: Monday, August 26.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 58–67 treat the somewhat broader problem $a_u x + bu_t + cu = f(x,t)$. We will work our way up to this full problem. The book also has a slightly different approach via the early introduction of characteristic curves (which we will meet later when we allow the coefficients a, b, and c to depend on space and/or time). Reading pp. 58–61 (stopping with Example 1) and comparing it to our approach below is a worthwhile exercise.

We solve the transport equation with $c = 1$ and claim that from this solution we can obtain all solutions to the more general problem with $c \neq 1$. We defer the study of this claim until later.

So, consider the problem

$$\begin{cases} u_t + u_x = 0, \ -\infty < x, \ t < \infty \\ u(x,0) = f(x). \end{cases}$$

To avoid irrelevant strangeness with the domain, we are looking for solutions defined on all of $\mathbb{R}^2$. The key to success here is to recognize the presence of some hidden coefficients:

$$u_t + u_x = (1 \cdot u_t) + (1 \cdot u_x).$$

This is really a dot product:

$$(1 \cdot u_t) + (1 \cdot u_x) = \begin{pmatrix} u_x \\ u_t \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The first vector is the gradient of u, $\nabla u = (u_x, u_t)$, and so we have

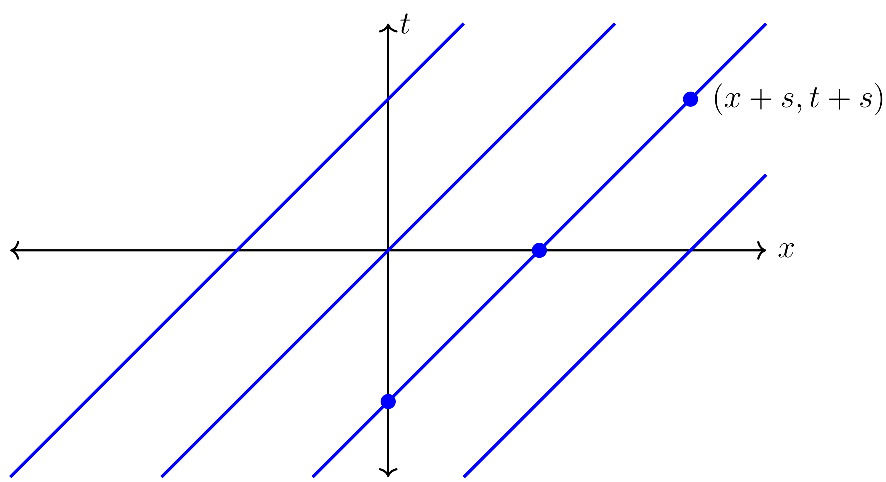
$$\nabla u \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0.$$

This dot product is the **DIRECTIONAL DERIVATIVE**: it measures how fast u is changing in the direction of the vector $(1,1)$, and the equality above says that u is really *constant* in that direction.

What does this mean? Fix $(x,t) \in \mathbb{R}^2$. "Moving" through (x,t) in the direction of the vector $(1,1)$ means moving along the line parametrized by

$$\begin{pmatrix} x \\ t \end{pmatrix} + s \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{7.1}$$

And u should be constant on any such line (as drawn in blue below).



That is, we expect that $u(x,t)$ equals the value of u at any point above in (7.1), for any choice of $s \in \mathbb{R}$. Perhaps we can choose s cleverly and maybe bring in the initial condition.

We make this more precise. With $(x,t) \in \mathbb{R}^2$ still fixed, we put

$$v(s) := u(x+s, t+s)$$

and compute, via the multivariable chain rule, that

$$v'(s) = u_x(x+s, t+s) + u_t(x+s, t+s) = 0$$

for all s. Thus v is constant. In particular,

$$u(x,t) = v(0) = v(s)$$

for any s. We can make the initial condition show up by taking $t+s=0$, thus $s=-t$. That is,

$$u(x,t) = v(-t) = u(x-t, t-t) = u(x-t, 0) = f(x-t).$$

We have proved a theorem.

7.1 Theorem. *Let $f \in \mathcal{C}^1(\mathbb{R})$ and suppose that u solves*

$$\begin{cases} u_t + u_x = 0, & -\infty < x,\ t < \infty \\ u(x,0) = f(x), & -\infty < x < \infty. \end{cases} \tag{7.2}$$

Then

$$u(x,t) = f(x-t).$$

This is a uniqueness result: the only possible solution to the IVP (7.2) is the one above. But is it really a solution?

7.2 Problem. Check that.

More generally, we claim that the only solution to

$$\begin{cases} u_t + cu_x = 0 \\ u(x,0) = f(x) \end{cases}$$

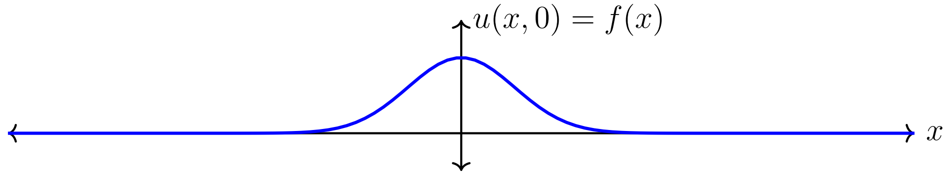
is

$$u(x,t) = f(x-ct).$$

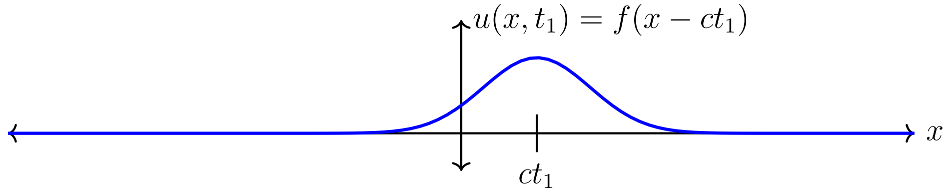
That this u is a solution can be checked as in Problem 7.2. (Do that.) That this u is the *only* solution still needs proof, which we will provide later.

So, here is our first reason for adoring the transport equation: it is a genuine PDE (that is not an ODE) and we know all of its solutions. The second reason is that these solutions respect our physical intuition: it turns out that the initial data f just gets "propagated"—dare we say, "transported"—along the x-axis with "speed" c. This is best seen through some pictures.

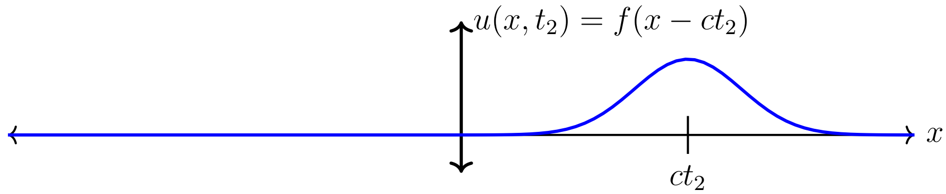
Here is a graph for the initial data f.



Say that $c > 0$ and we consider the solution u at time $t_1 > 0$. Then $u(x, t_1) = f(x - ct_1)$, and this graph is just the graph of f "shifted" by ct_1 units to the right.



We let time evolve more and the graph gets shifted more.



What we are really seeing here is the structure of a "traveling wave"—a fixed profile steadily translated in the same direction. We will explore the traveling wave structure of solutions to PDE much more in the future.

7.3 Example. The only solution to

$$\begin{cases} u_t + 3u_x = 0, \ -\infty < x, \ t < \infty \\ u(x, 0) = \sin(x), \ -\infty < x < \infty \end{cases}$$

is

$$u(x, t) = \sin(x - 3t).$$

We can take advantage of the diverse, flexible geometry of our domain $\mathbb{R}^2$ to specify the behavior of a solution not via an initial condition (i.e., via its behavior on the x-axis) but via a "side condition" in which we prescribe the solution's behavior on a one-dimensional curve in $\mathbb{R}^2$.

7.4 Example. We consider the problem

$$\begin{cases} u_t + 3u_x = 0, & -\infty < x,\ t < \infty \\ u(s,s) = \sin(s), & -\infty < s < \infty. \end{cases}$$

This prescribes the behavior of u on the line $x = t$. We know that if $u_t + 3u_x = 0$, then $u(x,t) = f(x - 3t)$, where $f(x) = u(x,0)$. Working backward, if we have a solution with the side condition, then

$$\sin(s) = u(s,s) = f(s - 3s) = f(-2s).$$

If we can express f explicitly, then we will know u. Here is where some algebraic trickery helps: put $\sigma = -2s$, so $s = -\sigma/2$. Then

$$f(-2s) = f(\sigma) = \sin\left(-\frac{\sigma}{2}\right) = -\sin\left(\frac{\sigma}{2}\right).$$

Then we expect that the solution is

$$u(x,t) = -\sin\left(\frac{x-t}{2}\right),$$

and we could always check that explicitly.

7.5 Problem. In what sense is any initial condition a side condition?

7.6 Example. Here is a situation in which we will not be as transparently successful in managing the side condition. Consider

$$\begin{cases} u_t + 3u_x = 0, & -\infty < x,\ t < \infty \\ u(s,-s^3) = \sin(s), & -\infty < s < \infty. \end{cases} \tag{7.3}$$

Now we are prescribing the behavior of u on the cubic $t = -x^3$. As before, the solution, if it exists, must have the form $u(x,t) = f(x - 3t)$ for some function f, and here we want

$$\sin(s) = u(s,-s^3) = f(s - 3(-s^3)) = f(s + 3s^3).$$

We would like to try the same algebraic trickery as before and put $\sigma = s + s^3$ and solve for s in terms of σ, but it is not at all apparent how to do that. (Perhaps a formula for roots of a cubic would help, but who knows what that says.)

Instead, as detailed below, we can appeal to the **INVERSE FUNCTION THEOREM** to argue that there exists a function $h \in \mathcal{C}^1(\mathbb{R})$ such that $\sigma = s + s^3$ if and only if $s = h(\sigma)$. We therefore put

$$f(\sigma) := \sin(h(\sigma)) \quad \text{and} \quad u(x,t) := \sin(h(x - 3t))$$

to obtain a solution candidate. We leave checking that this actually is a solution as an exercise with the inverse function theorem.

7.7 Problem. The following two statements are true.

(i) *Suppose that $\sigma \in \mathcal{C}(\mathbb{R})$ is strictly monotonic (i.e., σ is either strictly increasing or strictly decreasing). Then there exists $h \in \mathcal{C}(\mathbb{R})$ such that*

$$h(\sigma(s)) = s \quad \text{and} \quad \sigma(h(S)) = S \quad \text{for all} \quad s,\ S \in \mathbb{R}.$$

Such a function h is, of course, the **INVERSE** *of σ; this result says that a continuous strictly monotonic function on $\mathbb{R}$ has a continuous inverse.*

(ii) *Let $\sigma \in \mathcal{C}^1(\mathbb{R})$ and $h \in \mathcal{C}(\mathbb{R})$ such that $\sigma'(s) \neq 0$ for all $s \in \mathbb{R}$ and $\sigma(h(S)) = S$ for all $S \in \mathbb{R}$. Then $h \in \mathcal{C}^1(\mathbb{R})$ and*

$$h'(S) = \frac{1}{\sigma'(h(S))} \tag{7.4}$$

for all $S \in \mathbb{R}$. (The identity (7.4) is, hopefully, exactly what we expect by differentiating both sides of $\sigma(h(S)) = S$ and using the chain rule. The novelty here is that h is not initially assumed to be differentiable.)

Use these facts to justify the claims at the end of Example 7.6. That is, using these two facts, explain why there exists a function $h \in \mathcal{C}^1(\mathbb{R})$ such that putting

$$u(x, t) := \sin(h(x - 3t))$$

solves the problem (7.3).

Day 8: Wednesday, August 28.

Our work with side conditions was strictly algebraic; now we consider the interaction of the side condition curve with the geometry of the PDE.

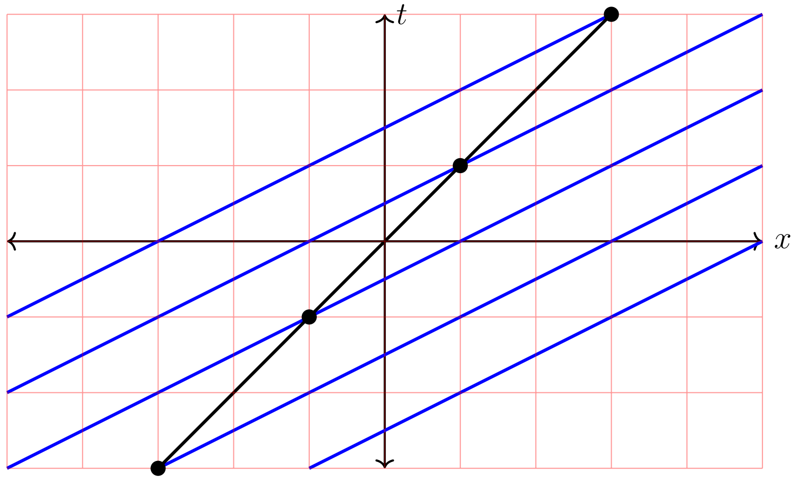
8.1 Example. We revisit the side conditions of Examples 7.4 and 7.6 more geometrically. Recall that all solutions to $u_t + 3u_x = 0$ have the form $u(x, t) = f(x - 3t)$ for some $f \in \mathcal{C}^1(\mathbb{R})$, and, since this transport equation is equivalent to

$$0 = \begin{pmatrix} u_x \\ u_t \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \nabla u \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

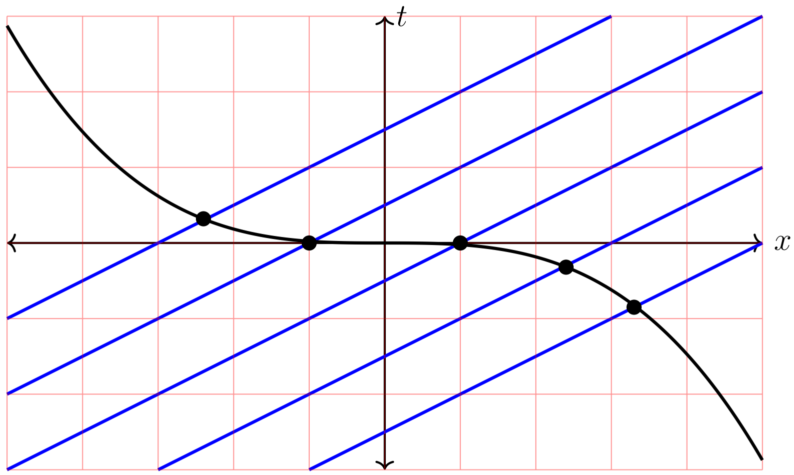
solutions u are constant on lines parallel to $(3, 1)$, i.e., lines with slope $1/3$.

(i) We graph in blue lines with slope $1/3$. Any solution u to $u_t + 3u_x = 0$ is constant on these lines. We graph in black the line parametrized by (s, s) with $s \in \mathbb{R}$, i.e., the line

$t = x$. We note that this black line intersects each blue line exactly once.



(ii) Again we graph in blue lines with slope $1/3$. Again, any solution u to $u_t + 3u_x = 0$ is constant on these lines. We graph in black the curve parametrized by $(s, -s^3)$ with $s \in \mathbb{R}$, i.e., the cubic $t = -x^3$. We note that this black curve intersects each blue line exactly once.



We will explore these graphical phenomena more generally later in the context of *characteristics* as part of our study of variable-coefficient linear problems, e.g., PDE of the form $u_t + c(x,t)u_x = 0$. For now, we just want to observe the very nice interaction of the side condition curves with the lines of slope $1/3$ that govern our solutions.

Here is a PDE with a side condition that does not admit any solution.

8.2 Example. Suppose that u solves

$$\begin{cases} u_t + 3u_x = 0, & -\infty < x, \ t < \infty \\ u(3s, s) = \sin(s), & -\infty < s < \infty. \end{cases}$$

Then u has the form $u(x,t) = f(x-3t)$ for some $f \in \mathcal{C}^1(\mathbb{R})$, and this f must satisfy

$$\sin(s) = u(3s,s) = f(3s-3s) = f(0) \tag{8.1}$$

for all $s \in \mathbb{R}$. This is impossible, as the sine is not constant. (For example, (8.1) would require $0 = \sin(0) = f(0) = \sin(\pi/2) = 1$.)

Algebraically, the problem simply fails. Geometrically, we note that the side condition curve is the line $t = x/3$, and u must be constant on this line. But the side condition says that u cannot be constant on this line! In contrast to the geometry of Example 8.1, the side condition curve intersects a line of slope $1/3$ more than once (in fact, infinitely often).

8.3 Problem. What goes wrong if you try to solve

$$\begin{cases} u_t + 3u_x = 0, \ -\infty < x, \ t < \infty \\ u(s,s^2) = \sin(s), \ -\infty < s < \infty? \end{cases}$$

Discuss the failure of this problem algebraically (the values $s = 0$ and $s = 1/3$ will be useful) and geometrically; include a sketch of how the side condition curve interacts with lines of slope $1/3$. Contrast that interaction with the situation in Example 8.2.

Now we return to the dangling problem of solving the more general transport equation. Consider the IVP

$$\begin{cases} au_t + bu_x = 0, \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty. \end{cases} \tag{8.2}$$

Here a, $b \neq 0$ to avoid the trivial case of a PDE that is really an ODE. Everything that we did for $u_t + u_x = 0$ could be replicated by recognizing that this transport equation is equivalent to

$$\nabla u \cdot \begin{pmatrix} b \\ a \end{pmatrix} = 0.$$

The only challenge would be the extra notation of a and b throughout.

However, to illustrate a valuable PDE technique that will serve us well with more complicated problems, we do not do this. Instead, suppose that we only know our previous result that

$$\begin{cases} v_t + v_x = 0, \ -\infty < x, \ t < \infty \\ v(x,0) = g(x), \ -\infty < x < \infty. \end{cases} \iff v(x,t) = g(x-t). \tag{8.3}$$

How can we use (8.3) to solve (8.2)? (In (8.3), we are using v and g, not u and f, in an effort not to overwork notation.)

This technique is **RESCALING**. First, we simplify the problem as much as possible by noting that, since $b \neq 0$, the IVP (8.2) is equivalent to

$$\begin{cases} u_t + cu_x = 0, \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty, \end{cases} \qquad c = \frac{b}{a}. \tag{8.4}$$

Now we assume that u solves (8.4). The key step is to define a new function via

$$U(X,T) := u(\alpha X, \beta T),$$

where α, $\beta \in \mathbb{R}$ are fixed constants whose value we will determine later. Specifically, we would like to choose them conveniently so that U solves an IVP like (8.3), which we fully understand.

We compute

$$U_X(X,T) = \alpha u_x(\alpha X, \beta T) \quad \text{and} \quad U_T(X,T) = \beta u_t(\alpha X, \beta T).$$

We hope that $U_T + U_X = 0$. We compute further

$$U_T(X,T) + U_X(X,T) = \beta u_t(\alpha X, \beta T) + \alpha u_x(\alpha X, \beta T).$$

Since we know

$$u_t(x,t) + u_x(x,t) = 0$$

for all $(x,t) \in \mathbb{R}^2$, if we take $\beta = 1$ and $\alpha = c$, then we have

$$U_T(X,T) + U_X(X,T) = u_t(cX,T) + u_x(cX,T) = 0.$$

And since $c \neq 0$, we can always express u in terms of U. That is, we have

$$U(X,T) = u(cX,T) \quad \text{and} \quad u(x,t) = U\left(\frac{x}{c}, t\right). \tag{8.5}$$

We are just missing an initial condition. We want to prescribe $U(X,0) = F(X)$ for some function F, and this means

$$F(X) = U(X,0) = u(cX,0) = f(cX).$$

To avoid overworking our variables, maybe we should define F via another symbol entirely, like $F(S) = f(cS)$.

Then U satisfies

$$\begin{cases} U_T + U_X = 0, & -\infty < X,\ T < \infty \\ U(X,0) = F(X), & -\infty < X < \infty, \end{cases} \qquad F(S) := f(cS),$$

and so by (8.3) we have

$$U(X,T) = F(X-T) = f(c(X-T)).$$

By (8.5), we conclude

$$u(x,t) = U\left(\frac{x}{c}, t\right) = f\left(c\left(\frac{x}{c} - t\right)\right) = f(x - ct).$$

And if we really want to go back to (8.2), we find

$$u(x,t) = f\left(x - \frac{b}{a}t\right) = f\left(\frac{ax - bt}{a}\right).$$

This rescaling trick can be employed more generally as follows. Suppose that $u = u(x,t)$ solves a "complicated" PDE. Put $U(X,T) = \gamma u(\alpha X, \beta T)$ and choose α, β, and γ (above $\gamma = 1$ because the transport equation was linear) so that U solves a "simpler" PDE. Use the relationship $u(x,t) = \gamma^{-1} U(\alpha^{-1}x, \beta^{-1}t)$ to recover u from knowledge of U.

8.4 Problem. The HEAT EQUATION for $u = u(x,t)$ is

$$u_t - \kappa u_{xx} = 0, \ -\infty < x < \infty, \ t \geq 0,$$

where $\kappa > 0$. (The importance of nonnegative time will be discussed later.) Suppose that u solves the heat equation and define $U(X,T) = u(\alpha X, \beta T)$ for α, $\beta \in \mathbb{R}$. What values of α and β make U solve the "simpler" heat equation

$$U_T - U_{XX} = 0?$$

8.5 Problem. Let a, b, c, A, B, $C \neq 0$. The most general version of the KORTEWEG–DE VRIES (KDV) EQUATION for $u = u(x,t)$ is

$$au_t + bu_{xxx} + cuu_x = 0, \ -\infty < x, \ t < \infty.$$

Suppose that u solves the KdV equation and define $U(X,T) = \gamma u(\alpha X, \beta T)$. What values of α, β, and γ make U solve the KdV equation

$$AU_T + BU_{XXX} + CUU_X = 0?$$

The point of this change of variables is that if we know how to solve KdV with one set of coefficients, then we know how to solve it with any other.

Day 9: Friday, August 30.

No class.

Day 10: Wednesday, September 4.

No class.

Day 11: Friday, September 6.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Appendix A.3 discusses Leibniz's rule at length (and in more detail than you are required to know). The examples on pp. 683–684 show how the rule can fail if the integrand is not sufficiently nice. A more general version of the rule appears on p. 687 and encompasses improper integrals, which we will eventually find useful. Lemma 1 on p. 177 gives a proof similar to ours for calculating $\partial_t\left[\int_0^t f(t,s)\, ds\right]$. A generalization of this appears in equation (12) on p. 688.

And we're back.

We now consider the **NONHOMOGENEOUS TRANSPORT EQUATION**:

$$\begin{cases} u_t + u_x = g(x,t), \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty. \end{cases}$$

Going back to the derivation of the (homogeneous) transport equation, one can think of g as a "source" (or "sink") term for the substance moving along the path—if the substance is cars and the path is a road, a nonzero g corresponds to on/off ramps along the road. This problem will be valuable to us for at least three reasons: (1) it illustrates and motivates some useful techniques with definite integrals, (2) its solution will be a key step in solving the (homogeneous) wave equation later, and (3) its solution form will motivate a surprisingly helpful idea for solving nonhomogeneous wave and heat equations later, too.

We get down to business and repeat our prior successful strategy. Fix x, $t \in \mathbb{R}$ and set

$$v(s) := u(x+s, t+s),$$

so

$$v'(s) = u_x(x+s,t+s) + u_t(x+s,t+s) = g(x+s,t+s) \quad \text{and} \quad v(0) = u(x,t).$$

Direct integration implies

$$v(s) = v(0) + \int_0^s v'(\sigma)\ d\sigma = u(x,t) + \int_0^s g(x+\sigma, t+\sigma)\ d\sigma.$$

That is,

$$u(x+s,t+s) = u(x,t) + \int_0^s g(x+\sigma,t+\sigma)\ d\sigma$$

for all x, t, $s \in \mathbb{R}$.

As before, we choose s conveniently with $s = -t$ to make the initial condition at $u(x,0)$ show up:

$$u(x-t,0) = u(x,t) + \int_0^{-t} g(x+\sigma,t+\sigma)\ d\sigma,$$

and so

$$f(x-t) = u(x,t) + \int_0^{-t} g(x+\sigma,t+\sigma)\ d\sigma.$$

One more rearrangement yields

$$u(x,t) = f(x-t) - \int_0^{-t} g(x+\sigma,t+\sigma)\ d\sigma.$$

It will pay off to clean up the integral a bit. The following is the nonobvious result of trial and error, but one motivation is that it would be nice to see the "$x-t$" structure in the integrand as well as in f. We can get this by substituting $\tau = t + \sigma$ (for lack of a better variable of integration), so

$$\tau(0) = t, \qquad \tau(-t) = 0, \qquad d\tau = d\sigma, \quad \text{and} \quad \sigma = \tau - t.$$

Then

$$-\int_0^{-t} g(x+\sigma, t+\sigma)\ d\sigma = -\int_t^0 g(x-t+\tau, \tau)\ d\tau = \int_0^t g(x-t+\tau,\tau)\ d\tau.$$

We summarize our work.

11.1 Theorem. *Let $f \in \mathcal{C}^1(\mathbb{R})$ and $g \in \mathcal{C}(\mathbb{R}^2)$ and suppose that u solves*

$$\begin{cases} u_t + u_x = g(x,t), \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty. \end{cases} \tag{11.1}$$

Then

$$u(x,t) = f(x-t) + \int_0^t g(x-t+\tau,\tau)\ d\tau. \tag{11.2}$$

However, we did not show that any function u in the form (11.2) actually solves (11.1). This requires computing both

$$\partial_x\left[\int_0^t g(x-t+\tau,\tau)\ d\tau\right] \quad \text{and} \quad \partial_t\left[\int_0^t g(x-t+\tau,\tau)\ d\tau\right].$$

We did something like the x-derivative in (6.1) when deriving the transport equation, but we never justified it, and the t-derivative looks even more complicated, since t appears in both the limit of integration and the integrand.

The time has come to sort this out. Consider the more abstract situation of calculating the derivative

$$\partial_x\left[\int_a^b f(x,s)\ ds\right].$$

Here h is defined on

$$\left\{(x,s) \in \mathbb{R}^2 \ \middle|\ a \le s \le b, \ x \in J\right\},$$

where J is some interval. For the integral to exist, we want the map

$$[a,b] \to \mathbb{R} \colon s \mapsto f(x,s)$$

to be continuous for each $x \in J$. We might abbreviate this map by $f(x,\cdot)$ and say that we want $f(x,\cdot) \in \mathcal{C}([a,b])$.

So what is the derivative, assuming that we do not recall (6.1)? The integral is approximately a Riemann sum, and derivatives and sums interact nicely:

$$\int_a^b f(x,s)\ ds \approx \sum_{k=1}^n f(x,s_k)(s_k - s_{k-1})$$

for a partition $\{s_k\}_{k=1}^n$ of the interval $[a,b]$. Certainly

$$\partial_x\left[\sum_{k=1}^n f(x,s_k)(s_k - s_{k-1})\right] = \sum_{k=1}^n f_x(x,s_k)(s_k - s_{k-1}),$$

and

$$\sum_{k=1}^{n} f_x(x, s_k)(s_k - s_{k-1}) \approx \int_a^b f_x(x, s)\ ds,$$

so perhaps

$$\partial_x \left[\int_a^b f(x, s)\ ds \right] = \int_a^b f_x(x, s)\ ds?$$

With some extra hypotheses, and work, this turns out to be true. The crux of the problem is an "interchange of limits" argument, the sort that permeates much of analysis. Using the definition of the derivative (and algebraically rearranging some terms on the left), this boils down to showing

$$\lim_{h \to 0} \int_a^b \frac{f(x+h, s) - f(x, s)}{h}\ ds = \int_a^b \lim_{h \to 0} \frac{f(x+h, s) - f(x, s)}{h}\ ds. \tag{11.3}$$

What properties of integrals give us the right to do this?

11.2 Theorem (Leibniz's rule for differentiating under the integral). *Let $J \subseteq \mathbb{R}$ be an interval and $a, b \in \mathbb{R}$ with $a \le b$. Let $\mathcal{D} := \{(x, s) \in \mathbb{R}^2 \mid x \in J,\ a \le s \le b\}$. Suppose that $f \in \mathcal{C}(\mathcal{D})$ and that f_x exists on $\mathcal{D}$ with $f_x \in \mathcal{C}(\mathcal{D})$. Then the map*

$$\mathcal{I} \colon J \to \mathbb{R} \colon x \mapsto \int_a^b f(x, s)\ ds$$

is defined and differentiable on J and

$$\mathcal{I}'(x) = \int_a^b f_x(x, s)\ ds.$$

11.3 Problem. Here is a sketch of the proof, up to some tricky estimates.

(i) Chase through the algebra of difference quotients and integrals to show that it suffices to establish (11.3) to prove Leibniz's rule.

(ii) Go further and show (using, perhaps, Problem 3.2) that it suffices to establish

$$\lim_{h \to 0} \int_a^b \int_0^1 \big[f_x(x + th, s) - f_x(x, s) \big]\ dt\ ds = 0. \tag{11.4}$$

(iii) Proving (11.4) takes some careful work with uniform continuity on compact subsets of $\mathbb{R}^2$, and that is beyond the scope of our class. However, show that if f_{xx} exists and is continuous on $\mathcal{D}$ and if there is $M > 0$ such that $|f_{xx}(x, s)| \le M$ for all $(x, s) \in \mathcal{D}$, then (11.4) holds. [Hint: *use Problem 3.2 again and watch out for the triple integral that shows up.*]

11.4 Problem. Let

$$\phi(x) := \int_0^1 s\cos(s^2+x)\ ds.$$

Calculate ϕ' in two ways in two ways: first by evaluating the integral with FTC2 and differentiating the result and second by differentiating under the integral and then simplifying the result with FTC2. (The point is to convince you that differentiating under the integral gives the right answer.)

If $g \in \mathcal{C}^1(\mathbb{R}^2)$, then Leibniz's rule justifies the calculation

$$\partial_x\left[\int_0^t g(x-t+\tau,\tau)\ d\tau\right] = \int_0^t g_x(x-t+\tau,\tau)\ d\tau$$

by taking $f(x,s) = g(x-t+s,s)$ with $t \in \mathbb{R}$ fixed. The hypothesis $g \in \mathcal{C}^1(\mathbb{R}^2)$ is, by the way, stronger than what we had in Theorem 11.1. (It is also asking more of g than we did of the forcing term in the ODE from Theorem 3.8. PDE are hard.)

We still need to calculate

$$\partial_t\left[\int_0^t g(x-t+\tau,\tau)\ d\tau\right],$$

and now the variable of differentiation appears in both the limit of integration (which should remind us of FTC1) and in the integrand (which should remind us of Leibniz's rule). To do this, it suffices to know how to compute

$$\partial_t\left[\int_0^t f(t,s)\ ds\right],$$

as we could then take $f(t,s) = g(x-t+s,s)$ with x fixed.

Here is the trick: we introduce a fake variable and set

$$F(x,t) := \int_0^x f(t,s)\ ds.$$

Then

$$\int_0^t f(t,s)\ ds = F(t,t),$$

so by the multivariable chain rule

$$\partial_t\left[\int_0^t f(t,s)\ ds\right] = F_x(t,t) + F_t(t,t).$$

But

$$F_x(t,t) = \partial_x\left[\int_0^x f(t,s)\ ds\right]\bigg|_{x=t} = f(t,t)$$

by FTC1 and

$$F_t(t,t) = \partial_t\left[\int_0^x f(t,s)\ ds\right]\bigg|_{x=t} = \int_0^t f_t(t,s)\ ds$$

by Leibniz's rule.

We have proved the following.

11.5 Lemma. *Let $f \in \mathcal{C}^1(\mathbb{R}^2)$. Then*

$$\partial_t \left[\int_0^t f(t,s)\ ds\right] = f(t,t) + \int_0^t f_t(t,s)\ ds$$

for all $t \in \mathbb{R}$.

11.6 Problem. Use this lemma to show that if $g \in \mathcal{C}^1(\mathbb{R}^2)$, then

$$u(x,t) := \int_0^t g(x - t + \tau, \tau)\ d\tau$$

solves

$$\begin{cases} u_t + u_x = g(x,t), \ -\infty < x, \ t < \infty \\ u(x,0) = 0, \ -\infty < x < \infty. \end{cases}$$

11.7 Problem. Find all solutions to

$$\begin{cases} u_t + cu_x + ru = g(x,t), \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty. \end{cases}$$

where $f \in \mathcal{C}^1(\mathbb{R})$, $g \in \mathcal{C}^1(\mathbb{R}^2)$, and c, $r \in \mathbb{R}$. [Hint: *as always, start with $v(s) := u(x + cs, t + s)$ for x, $t \in \mathbb{R}$ fixed and find an ODE for v.*] This transport equation models the propagation of a substance where the amount of the substance on the path can change both from the "source/sink" term g and in proportion r to the amount of substance on the path.

Day 12: Monday, September 9.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Page 281 provides cultural and historical context for the wave equation. Pages 282–285 exhaustively derive the wave equation from physical principles. Pages 300–302 derive D'Alembert's formula using a slightly different approach from ours in class. Read Examples 3 and 4 on pp. 303–304.

We commence our study of a new PDE: the **WAVE EQUATION**. In the immortal words of G. B. Whitham from his staggering *Linear and Nonlinear Waves*,

> *"[A]* ***wave*** *is any recognizable signal that is transferred from one part of [a] medium, to another with a recognizable velocity of propagation. The signal may be any feature of the disturbance, such as a maximum or an abrupt change in some quantity, provided that it can be clearly recognized and its location at any time can be determined. The signal may distort, change its magnitude, and change its velocity provided it is still recognizable."*

The initial value problem (IVP) for the wave equation on $\mathbb{R}$ reads

$$\begin{cases} u_{tt} = u_{xx}, & -\infty < x, \ t < \infty \\ u(x,0) = f(x), & -\infty < x < \infty \\ u_t(x,0) = g(x), & -\infty < x < \infty. \end{cases}$$

Here f, $g \colon \mathbb{R} \to \mathbb{R}$ are given functions. This IVP models the motion of an infinitely long string that moves in the vertical direction only: let $u(x,t)$ be the displacement of the string from its rest position at position x along its length and time t. The function f models the initial displacement and g the initial velocity. While a finite string is of course physically much more realistic, we will see that finite length leads to some complicated, and possibly unsatisfying, boundary conditions; mathematically, the infinite string is rather "nicer" (if more unrealistic physically).

We can solve the IVP by noticing a formal similarity to the difference of perfect squares: u solves the wave equation if and only if

$$u_{tt} - u_{xx} = 0,$$

and we might rewrite this in "operator" notation as

$$(\partial_t^2 - \partial_x^2)u = 0,$$

and then factor that as

$$(\partial_t - \partial_x)(\partial_t + \partial_x)u = 0.$$

What this means is that if u solves $u_{tt} = u_{xx}$, and if we define $v := u_t + u_x$, then v solves

$$v_t - v_x = 0.$$

12.1 Problem. Prove that.

The function v therefore solves a transport equation. Since

$$v(x,0) = u_t(x,0) + u_x(x,0) = g(x) + \partial_x[u](x,0) = g(x) + \partial_x[f](x) = g(x) + f'(x),$$

the function v really solves

$$\begin{cases} v_t - v_x = 0, & -\infty < x, \ t < \infty \\ v(x,0) = g(x) + f'(x). \end{cases}$$

We know that the solution to this problem is

$$v(x,t) = g(x+t) + f'(x+t).$$

Consequently, the solution u to the original wave equation $u_{tt} = u_{xx}$ must also solve

$$u_t + u_x = v(x,t) = g(x+t) + f'(x+t).$$

Since $u(x,0) = f(x)$, we meet another transport equation:

$$\begin{cases} u_t + u_x = g(x+t) + f'(x+t), \ -\infty < x, \ t < \infty \\ u(x,0) = f(x). \end{cases}$$

We know from Theorem 11.1 that the solution to

$$\begin{cases} u_t + u_x = h(x,t), \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty \end{cases}$$

is

$$u(x,t) = f(x-t) + \int_0^t h(x-t+\tau,\tau) \ d\tau.$$

With

$$h(x,t) = g(x+t) + f'(x+t),$$

we have

$$h(x-t+\tau,\tau) = g((x-t+\tau)+\tau) + f'((x-t+\tau)+\tau) = g(x-t+2\tau) + f'(x-t+2\tau).$$

Thus the solution u to the wave equation $u_{tt} = u_{xx}$ is

$$u(x,t) = f(x-t) + \int_0^t \big[g(x-t+2\tau) + f'(x-t+2\tau)\big] \ d\tau.$$

We change variables in the integral with

$$s = x-t+2\tau, \qquad ds = 2 \ d\tau, \qquad s(0) = x-t, \qquad s(t) = x+t,$$

to find

$$\begin{aligned} \int_0^t \big[g(x-t+2\tau) + f'(x-t+2\tau)\big] \ d\tau &= \frac{1}{2}\int_{x-t}^{x+t} \big[g(s) + f'(s)\big] \ ds \\ &= \frac{f(x+t) - f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(s) \ ds. \end{aligned}$$

We conclude

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

Here is a slightly more general result.

12.2 Theorem (D'Alembert's formula). *Let $f \in \mathcal{C}^2(\mathbb{R})$ and $g \in \mathcal{C}^1(\mathbb{R})$ and $c > 0$. The only solution $u \in \mathcal{C}^2(\mathbb{R}^2)$ to*

$$\begin{cases} u_{tt} = c^2 u_{xx}, & -\infty < x,\ t < \infty \\ u(x,0) = f(x), & -\infty < x < \infty \\ u_t(x,0) = g(x), & -\infty < x < \infty \end{cases} \tag{12.1}$$

is the function

$$u(x,t) = \frac{f(x+ct) + f(x-ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\ ds. \tag{12.2}$$

12.3 Problem. Prove this.

(i) First, check that u as defined in (12.2) actually solves the wave IVP (12.1). Explain why the regularity assumptions $f \in \mathcal{C}^2(\mathbb{R})$ and $g \in \mathcal{C}^1(\mathbb{R})$ are necessary.

(ii) Next, develop the result for $c \neq 1$ from the work above by assuming that u solves (12.1) and setting $U(X,T) = u(\alpha X, \beta T)$ for some α, $\beta \in \mathbb{R}$. Choose α and β so that U solves $U_{TT} = U_{XX}$ and use the work above (updating the initial conditions as needed) to find a formula for U. From that, develop the formula (12.2) for u.

12.4 Example. We solve the wave IVP (12.1) for $c = 1$ and some choices of f and g and graph some results.

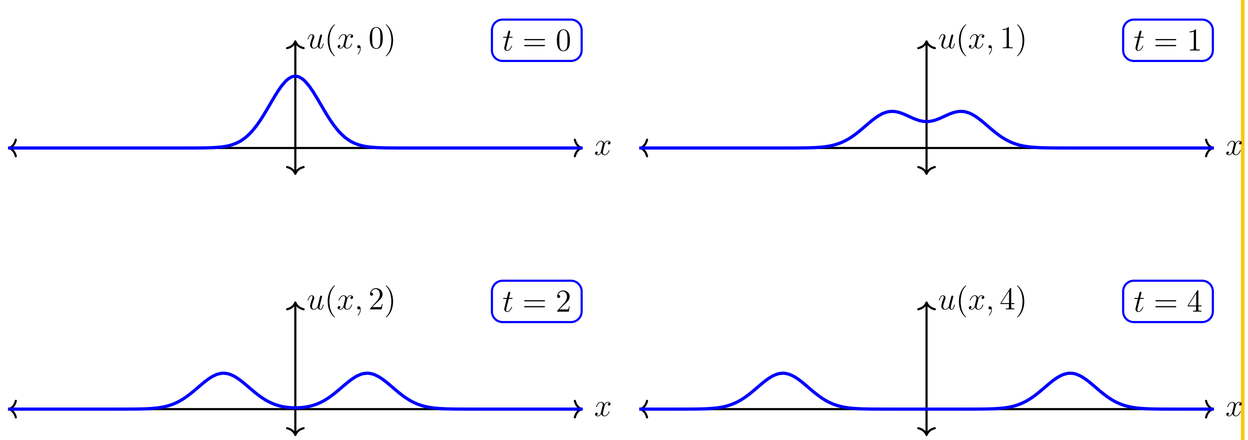
(i) Take

$$f(x) = 2e^{-x^2} \quad \text{and} \quad g(x) = 0.$$

D'Alembert's formula tells us that the solution is

$$u(x,t) = \frac{2e^{-(x+t)^2} - 2e^{-(x-t)^2}}{2} + \frac{1}{2}\int_{x-t}^{x+t} 0\ ds = e^{-(x+t)^2} + e^{-(x-t)^2}.$$

Here are some plots.



It looks like the initial condition $u(x,0) = 2e^{-x^2}$ has split into two smaller "pulses," one moving to the right and the other to the left. This is exactly what the formula $u(x,t) = e^{-(x+t)^2} - e^{-(x-t)^2}$ says: as t increases, the graph of $x \mapsto e^{-(x+t)^2}$ moves to the left, while $x \mapsto e^{-(x-t)^2}$ moves to the right. However, the graph of $u(\cdot, t)$ is not really just the graph of $x \mapsto e^{-(x+t)^2}$ superimposed on the graph of $x \mapsto e^{-(x-t)^2}$; there is an interaction between the two graphs due to the sum in the definition of u. Nonetheless, this interaction is very "weak" for x or t large because e^{-s^2} is very small when $|s|$ is very large.

(ii) Take

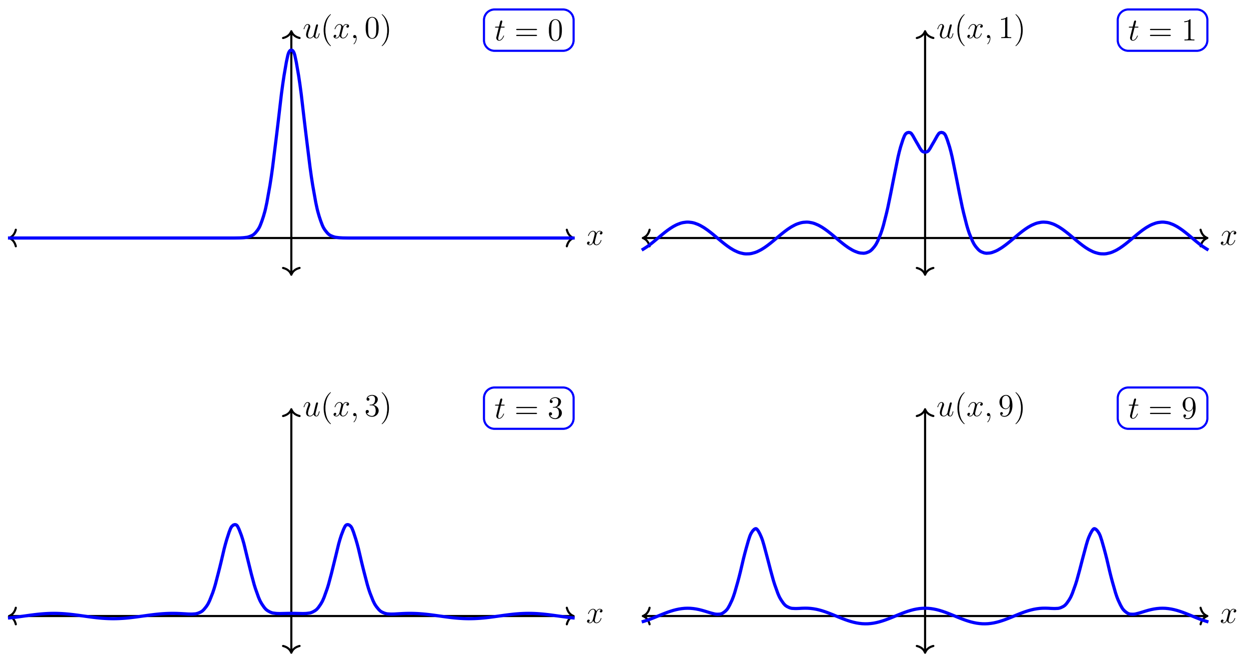
$$f(x) = 10e^{-x^2} \quad \text{and} \quad g(x) = \cos(x).$$

D'Alembert's formula tells us that the solution is

$$u(x,t) = \frac{10e^{-(x+t)^2} + 10e^{-(x-t)^2}}{2} + \frac{1}{2}\int_{x-t}^{x+t} \cos(s)\, ds$$

$$= 5\big(e^{-(x+t)^2} + e^{-(x-t)^2}\big) + \frac{\big(\sin(x+t) - \sin(x-t)\big)}{2}.$$

Here are some graphs.



Again, it looks like the initial condition "splits" into two "smaller" pulses that travel to the right and left; now there is more "noise" between them due to the nonzero initial condition on u_t. In particular, the pulses are not nearly as "identical" as they were for the previous initial data; contrast times 1, 3, and 9 with the previous pulses for times 1, 2, and 4.

Here is why this "counterpropagating pulse" phenomenon happens. Rewrite D'Alembert's formula as

$$\frac{f(x+ct)+f(x-ct)}{2}+\frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\ ds=\frac{1}{2}\left(f(x+ct)+\int_0^{x+ct} g(s)\ ds\right)$$
$$+\frac{1}{2}\left(f(x-ct)+\int_{x-ct}^{0} g(s)\ ds\right)$$

and abbreviate

$$F(X):=\frac{1}{2}\left(f(X)+\frac{1}{c}\int_0^X g(s)\ ds\right) \quad \text{and} \quad G(X):=\frac{1}{2}\left(f(X)+\frac{1}{c}\int_X^0 g(s)\ ds\right).$$

Then if u solves $u_{tt}=c^2u_{xx}$, we can write

$$u(x,t)=F(x+ct)+G(x-ct). \tag{12.3}$$

This is the superposition of the "profiles" F and G with F translated left with "speed" c and G translated "right." And this is why the graphs in Example 12.4 break up into two "counterpropagating" profiles.

12.5 Remark. *The profiles F and G above are definitely not the initial data f and g in general. In fact, the formula* (12.3) *makes sense without any initial data. Just assume that u solves $u_{tt}=c^2u_{xx}$ and artificially introduce the initial conditions $f(x):=u(x,0)$ and $g(x):=u_t(x,0)$. Then the work above shows that u satisfies* (12.3)*, and we can forget about f and g if we want.*

The structure in (12.3) is really a sum of traveling waves.

12.6 Definition. *A function $u\colon \mathbb{R}^2\to\mathbb{R}$ is a* **TRAVELING WAVE** *if there exist a function $p\colon\mathbb{R}\to\mathbb{R}$ and $c\in\mathbb{R}$ such that*

$$u(x,t)=p(x-ct)$$

for all x, $t\in\mathbb{R}$. The function p is the **PROFILE** *and the scalar c is the* **WAVE SPEED***.*

The idea of a traveling wave is that the profile p is translated, or "travels," via the shift by $-ct$. In particular, if $c>0$, then as time increases, the graph of $x\mapsto u(x,t)$ is just the graph of p translated to the right by ct units.

12.7 Problem. Explain why all solutions to the homogeneous transport equation $u_t+u_x=0$ are traveling waves but solutions to the wave equation are typically *not* traveling waves.

Day 13: Wednesday, September 11.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Example 6 on p. 307 and the following remark and Example 7 on pp. 308–310 present the "method of images" for the semi-infinite string.

When studying a PDE in the unknown function $u = u(x,t)$, the process of guessing that u is a traveling wave of the form $u(x,t) = p(x-ct)$ and then figuring out the permissible profile(s) p and wave speed(s) c, if any, is called making a **TRAVELING WAVE ANSATZ** for that PDE. (In general, an **ANSATZ** for a PDE is an educated guess that a solution has a particular form.)

13.1 Example. For the sake of a toy problem, we pause from our study of the wave equation and consider a nonlinear transport equation:

$$u_x + u_t + u^2 = 0.$$

We make the traveling wave ansatz $u(x,t) = p(x-ct)$ for a profile function $p = p(X)$ and a wave speed $c \in \mathbb{R}$. The multivariable chain rule tells us that

$$u_x(x,t) = p'(x-ct) \quad \text{and} \quad u_t(x,t) = -cp'(x-ct).$$

Thus p and c must satisfy

$$p'(x-ct) - cp'(x-ct) + [p(x-ct)]^2 = 0$$

for all x, $t \in \mathbb{R}$. If we take $x = X$ and $t = 0$, which we are free to do, we see that p must satisfy

$$(1-c)p'(X) + [p(X)]^2 = 0,$$

or, more succinctly,

$$(1-c)p' + p^2 = 0.$$

This is actually a separable ODE, and we can rewrite it as

$$(1-c)\frac{dp}{dX} = -p^2.$$

The equilibrium solution is $p(X) = 0$. When $c = 1$, we have $-p^2 = 0$, and so again $p(X) = 0$. From now on, assume $c \neq 1$ and $p \neq 0$. Then we formally separate variables and integrate:

$$(1-c)\frac{dp}{dX} = -p^2 \Longrightarrow \frac{c-1}{p^2}\frac{dp}{dX} = 1 \Longrightarrow (c-1)\int p^{-2}\,dp = \int dX \Longrightarrow (1-c)p^{-1} = X+K.$$

Here we are using K, not C, for the constant of integration, to avoid confusion with c. Since $1-c \neq 0$, we can solve for p:

$$p(X) = \frac{1-c}{X+K}.$$

Actually, this encodes the zero solution resulting from $c = 1$, and so all traveling waves are

$$u(x,t) = \frac{1-c}{x-ct+K}. \tag{13.1}$$

13.2 Problem. Find all other solutions to $u_t + u_x + u^2 = 0$. [Hint: *put* $v(s) = u(x+s, t+s)$ *and find a separable ODE for* v.]

13.3 Problem. We have said (and proved) that all solutions to the transport equation $u_t + u_x = 0$ are traveling waves, but make a traveling wave ansatz $u(x,t) = p(x - ct)$ anyway and solve for p and c. What is special about the case $c = -1$?

13.4 Problem. Make a traveling wave ansatz $u(x,t) = p(x - ct)$ for the KdV equation $u_t + u_{xxx} + uu_x = 0$ and find, but do not solve, an ODE that p must satisfy.

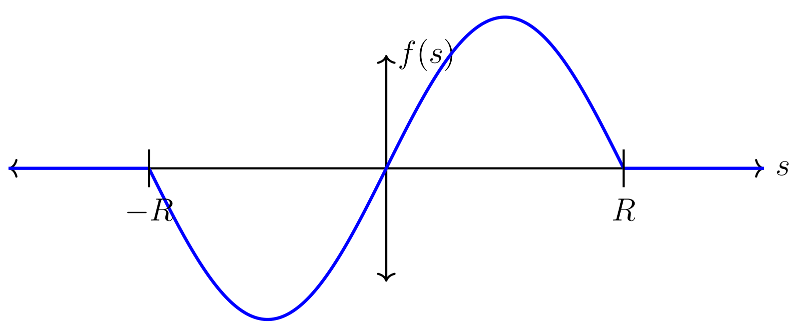
We will continue making traveling wave ansatzes for other PDE that we meet and interpreting those solutions physically and mathematically in the broader context of those equations. Now we return to the wave equation and tease out more properties from D'Alembert's formula.

Common jargon for the wave equation is that it "exhibits finite propagation speed." Physically, this means that data or disturbances in one part of the fictitious infinite string take some time to affect other parts of the string. Here is what this means mathematically.

Suppose that the initial data f and g have **COMPACT SUPPORT** in the sense that there is $R > 0$ such that

$$f(s) = 0 \quad \text{and} \quad g(s) = 0 \quad \text{for} \quad |s| > R.$$

In other words, f and g can only be nonzero on the interval $[-R, R]$. (Here we are using s for the independent variable of f and g to avoid confusion with x.) In more other words, the only "data" carried by f and g exists on this finite interval.



If the string is governed by $u_{tt} = c^2 u_{xx}$, then we expect that $c > 0$ is the speed of the wave(s) moving through the string. After t units of time, data or disturbances should only propagate ct units along the x-axis from where they were at time 0. This is born out by D'Alembert's formula.

Fix $t > 0$ and suppose that $R + ct < x$. Then this position x is more than ct units outside the "support" of f and g. We do not expect the data or disturbances from f and g to reach position x in only this time t. Now here is the math: since $R + ct < x$ and c, $t > 0$, we have

$$R < R + 2ct < x + ct, \qquad R < x - ct, \qquad \text{and} \qquad R < x - ct < x + ct.$$

Since $f(s) = 0$ for $s > R$, we have

$$\frac{f(x + ct) + f(x - ct)}{2} = 0.$$

Also, since $g(s) = 0$ on (R, ∞) and $[x - ct, x + ct] \subseteq (R, \infty)$, we have

$$\int_{x-ct}^{x+ct} g(s)\ ds = 0.$$

D'Alembert's formula then implies that $u(x, t) = 0$.

Here is what we have proved.

13.5 Corollary (Finite propagation speed for the wave equation). *Let $f \in \mathcal{C}^2(\mathbb{R})$ and $g \in \mathcal{C}^1(\mathbb{R})$ have compact support with $f(s) = g(s) = 0$ for $|s| > R$. Let $c > 0$. If u solves the wave IVP (12.1), then $u(x, t) = 0$ for $|x| > R + c|t|$.*

13.6 Problem. Review the work preceding the corollary and check that it holds for $|x| > R + c|t|$, not just for $x > R + ct$ as we actually worked out above.

13.7 Problem. Formulate and prove a finite propagation speed result for the transport IVP

$$\begin{cases} u_t + cu_x = 0, \ -\infty < x, \ t < \infty \\ u(x, 0) = f(x), \ -\infty < x < \infty \end{cases}$$

that is similar to Corollary 13.5.

As an illustration of more properties of D'Alembert's formula (and, really, more properties of functions and integrals), we introduce our first boundary condition and study the "semi-infinite" string. Suppose that one end of the string is fixed at $x = 0$, so $u(0, t) = 0$ for all t, and the string extends infinitely to the right for $x > 0$. We take initial data valued only for $x \geq 0$ and consider the IVP-BVP

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \leq x < \infty, \ -\infty < t < \infty \\ u(x, 0) = f(x), \ 0 \leq x < \infty \\ u_t(x, 0) = g(x), \ 0 \leq x < \infty \\ u(0, t) = 0, \ -\infty < t < \infty. \end{cases} \tag{13.2}$$

Here we assume $f \in \mathcal{C}^2([0, \infty))$ and $g \in \mathcal{C}^1([0, \infty))$. As usual, at the left endpoint 0 we only assume that limits from the right hold, e.g.,

$$\lim_{x \to 0^+} f(x) = f(0), \qquad \lim_{h \to 0^+} \frac{f(0 + h) - f(0)}{h} = f'(0), \qquad \lim_{x \to 0^+} f'(x) = f'(0), \qquad \text{and so on.}$$

There are two new wrinkles in this problem. The first is the presence of the boundary condition $u(0, t) = 0$. We call this a boundary condition because it specifies what the solution is doing at the left endpoint, or "boundary," of its x-domain.

The second is that f and g are only defined on $[0, \infty)$. If f and g were defined on all of $\mathbb{R}$, we could reduce this to the wave IVP previously solved and use D'Alembert's formula. The right idea is to *extend* f and g to $\mathbb{R}$. That is, we want functions $\widetilde{f} \in \mathcal{C}^2(\mathbb{R})$ and $\widetilde{g} \in \mathcal{C}^1(\mathbb{R})$ such that

$$\widetilde{f}(x) = f(x) \quad \text{and} \quad \widetilde{g}(x) = g(x) \quad \text{for} \quad x \geq 0.$$

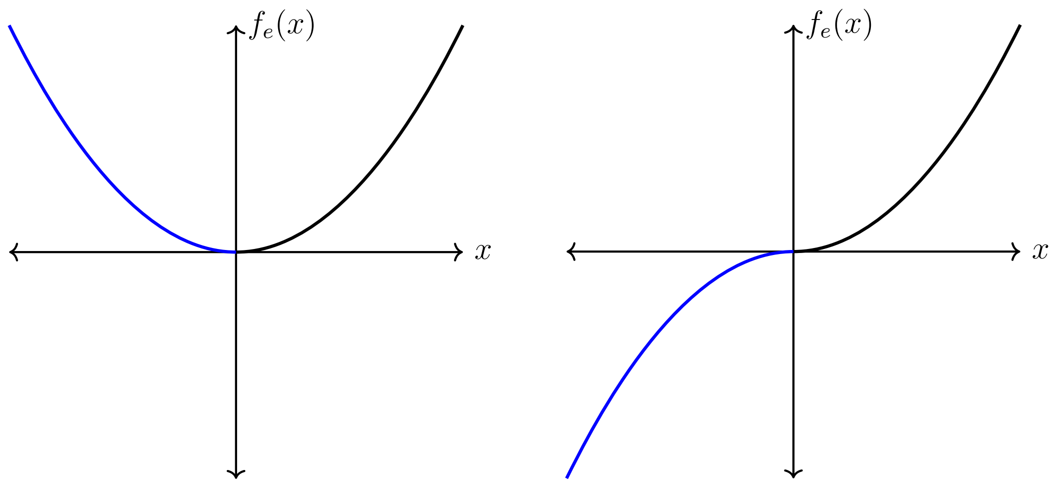
One way to make f and $\widetilde{g}$ well-behaved on $(-\infty, 0)$ is to exploit the good behavior of f and g and just "reflect" f and g across the vertical axis. That is, we could make f the **EVEN REFLECTION OF** f

$$f_e(x) = \begin{cases} f(x), & x \geq 0 \\ f(-x), & x < 0 \end{cases}$$

or the **ODD REFLECTION OF** f

$$f_o(x) = \begin{cases} f(x), & x \geq 0 \\ -f(-x), & x < 0. \end{cases} \tag{13.3}$$

Here are pictures with the original graph of f in black and the graph of the extension added in blue.



We could do the same reflections for g. The question is then what behavior we can get at $x = 0$: will the reflections be sufficiently differentiable there?

Here is where we need to think about the data in our problem. The boundary condition $u(0, t) = 0$ talks to the initial conditions and implies

$$f(0) = u(0, 0) = 0 \quad \text{and} \quad g(0) = u_t(0, 0) = \partial_t[u(0, t)]\big|_{t=0} = \partial_t[0] = 0.$$

To be clear, we are assuming that the semi-infinite IVP-BVP has a solution, and we are concluding that $f(0) = g(0) = 0$. We will henceforth *require* this in the initial data; the problem does not make sense without it.

We know that odd functions are 0 at 0, so this suggests that we use the odd reflections f_o and $\widetilde{g}_o$.

13.8 Problem. Check this claim: if $h: \mathbb{R} \to \mathbb{R}$ is odd, i.e., if $h(s) = h(-s)$ for all $s \in \mathbb{R}$, then $h(0) = 0$. Also show that if h is odd, then h'' is odd, so $h''(0) = 0$.

We hope, then, that D'Alembert's formula with the initial data given by f_o and g_o will provide a solution to (13.2) when restricted to $x \geq 0$ and $t = 0$:

$$u(x,t) = \frac{f_o(x+t) + f_o(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g_o(s)\ ds.$$

Will it? There are several things to check.

Day 14: Friday, September 13.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

The remark on p. 308 discusses how to solve the semi-infinite string problem with the boundary condition $u_x(0,t) = 0$. Pages 326–327 give physical motivation for the driven wave equation.

First we claim that if $g \in \mathcal{C}([0,\infty))$ and $g(0) = 0$, then $g_o \in \mathcal{C}(\mathbb{R})$, so the integral is at least defined.

14.1 Problem. Check this continuity claim.

Next, we check the newest and shiniest part of the IVP-BVP: the boundary condition. We want $u(0,t) = 0$ for all t, so we need

$$\frac{f_o(t) + f_o(-t)}{2} + \frac{1}{2}\int_{-t}^{t} g_o(s)\ ds = 0.$$

Since f_o is odd, $f_o(t) = -f_o(-t)$, so the first term is 0.

14.2 Problem. Show that if $h \in \mathcal{C}(\mathbb{R})$ is odd, then

$$\int_{-a}^{a} h = 0$$

for any $a \in \mathbb{R}$. [Hint: *substitute.*] Draw a picture indicating why this should be true in general (caution: picture $\neq$ proof).

With this calculation, the integral term is also odd, so the boundary condition does work out.

What about differentiability? We want $f_o \in \mathcal{C}^2(\mathbb{R})$ and $g_o \in \mathcal{C}^1(\mathbb{R})$. But since f_o is odd, we need $f_o''(0) = 0$. Will this happen? We expect

$$f_o'(x) = \begin{cases} f'(x), & x > 0 \\ f'(-x), & x < 0 \end{cases} \qquad \text{and} \qquad f_o''(x) = \begin{cases} f''(x), & x > 0 \\ -f''(-x), & x < 0. \end{cases}$$

For f_o'' to be continuous, we therefore want

$$0 = f_o''(0) = \lim_{x \to 0^+} f_o''(x) = \lim_{x \to 0^+} f''(x) = f''(0),$$

where the second-to-last equality is from the piecewise formula for f_o'', and the last equality is from the continuity of f'' on $[0, \infty)$. Does the IVP-BVP structure of (13.2) guarantee this?

14.3 Problem. **(i)** Check that if u solves the semi-infinite problem (13.2), then we do have $f''(0) = 0$. [Hint: *how does the boundary data talk to the initial data and the PDE to help you calculate f''?*] The conclusion here is that requiring $f''(0) = 0$ is a natural, unrestrictive condition for the semi-infinite wave IVP-BVP to make sense.

(ii) Suppose that $f \in \mathcal{C}^2([0, \infty)$ with $f(0) = f''(0) = 0$. With the odd reflection f_o defined in (13.3), show that $f_o \in \mathcal{C}^2(\mathbb{R})$. [Hint: *the hard work is at $x = 0$: use left and right limits for the difference quotients to show that $f_o'(0)$ and $f_o''(0)$ are defined, and then show that* $\lim_{s \to 0} f_o''(s) = f_o''(0)$.]

The conclusion is success: if $f \in \mathcal{C}^2([0, \infty))$ with $f(0) = f''(0) = 0$ and $g \in \mathcal{C}^1([0, \infty))$ with $g(0) = 0$, then we can construct a solution to the semi-infinite wave IVP-BVP (13.2) out of the odd reflections and D'Alembert's formula. The nuance here is that we have these "compatibility conditions" $f(0) = f''(0) = g(0) = 0$ for the problem to make sense. Nothing in the original statement of (13.2) gave those conditions explicitly; they were lurking hidden in the background.

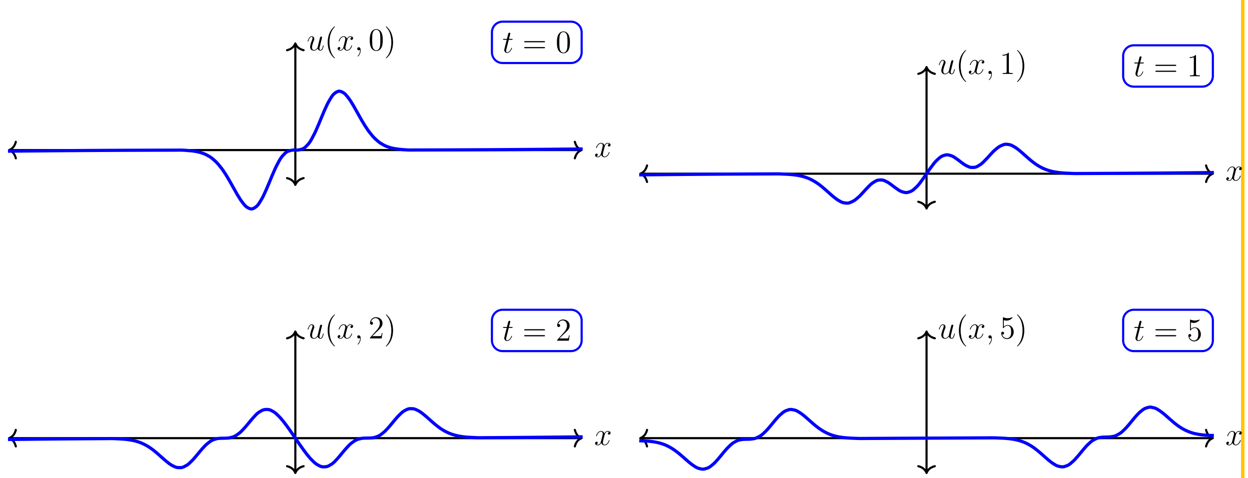
14.4 Example. Here is the solution to the semi-infinite wave IVP-BVP with $c = 1$ when

$$f(x) = 4x^3 e^{-x^2} \quad \text{and} \quad g(x) = 0.$$

Here f and g are already odd, so we do not need to go to any great lengths to calculate their odd reflections. Indeed, we just have

$$u(x,t) = 2(x+t)^3 e^{-(x+t)^2} + 2(x-t)^3 e^{-(x-t)^2}.$$

Below we graph the solution for various time values.



As before, we see that the solution splits up into two counterpropagating pulses, but now they are clearly reflections or "images" of each other through the vertical axis.

14.5 Problem. Prove that if f and g are odd, then D'Alembert's formula (12.2) yields an odd function in x, i.e., $u(-x,t) = -u(x,t)$. This justifies the "reflection" remark in the example above.

14.6 Problem (Optional, involved). Solve the problem

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \le x < \infty, \ -\infty < t < \infty \\ u(x,0) = f(x), \ 0 \le x < \infty \\ u_t(x,0) = g(x), \ 0 \le x < \infty \\ u_x(0,t) = 0, \ -\infty < t < \infty, \end{cases}$$

where $f \in \mathcal{C}^2([0,\infty))$ and $g \in \mathcal{C}^1([0,\infty))$. This problem models a semi-infinite string where the left endpoint is allowed to move vertically. [Hint: *try even extensions for f and g. What "compatibility" conditions arise?*]

Now we take up the study of the driven or nonhomogeneous wave equation:

$$\begin{cases} u_{tt} = u_{xx} + h(x,t), \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty \\ u_t(x,0) = g(x), \ -\infty < x < \infty, \end{cases} \tag{14.1}$$

where we assume, as usual, $f \in \mathcal{C}^2(\mathbb{R})$ and $g \in \mathcal{C}^1(\mathbb{R})$ and, at the minimum, $h \in \mathcal{C}(\mathbb{R}^2)$. We develop our solution method by first noting some (probably non-obvious) patterns among the driven linear equations that we have previously solved.

1. The first-order linear nonhomogeneous IVP at the ODE level "splits" into the sum of two "easier" problems:

$$\begin{cases} y' = ay + f(t) \\ y(0) = y_0 \end{cases} \quad = \quad \begin{cases} y' = ay \\ y(0) = y_0 \end{cases} \quad + \quad \begin{cases} y' = ay + f(t) \\ y(0) = 0. \end{cases}$$

This sum is wholly euphemistic; the point is that the solution to the "full" IVP is the sum of solutions to the "simpler" IVP. They are "simpler" because the first has no driving term (but has a "harder" initial condition), while the second has an "easier" initial condition (but a "harder" driving term).

Of course, the solution to

$$\begin{cases} y' = ay \\ y(0) = y_0 \end{cases}$$

is

$$y(t) = e^{at}y_0,$$

and the solution to

$$\begin{cases} y' = ay + f(t) \\ y(0) = 0 \end{cases}$$

is

$$y(t) = e^{at} \int_0^t e^{-a\tau} f(\tau) \, d\tau.$$

The key to everything is rewriting this second solution:

$$e^{at} \int_0^t e^{-a\tau} f(\tau) \, d\tau = \int_0^t e^{a(t-\tau)} f(\tau) \, d\tau.$$

We recognize the presence of the first solution within the second solution via a notational sleight-of-hand: put

$$\mathcal{P}(t) := e^{at},$$

so

$$e^{at} y_0 = \mathcal{P}(t) y_0 \quad \text{and} \quad \int_0^t e^{a(t-\tau)} f(\tau) \, d\tau = \int_0^t \mathcal{P}(t-\tau) f(\tau) \, d\tau.$$

We think of $\mathcal{P}$ as a "propagator operator" for the homogeneous problem in that it "propagates" the initial data y_0 to where it should be at time t (namely, to $e^{at} y_0$). The solution to the full nonhomogeneous IVP is therefore

$$y(t) = \mathcal{P}(t) y_0 + \int_0^t \mathcal{P}(t-\tau) f(\tau) \, d\tau.$$

2. The nonhomogeneous transport IVP similarly "splits":

$$\begin{cases} u_t = -u_x + g(x,t) \\ u(x,0) = f(x) \end{cases} = \begin{cases} u_t = -u_x \\ u(x,0) = f(x) \end{cases} + \begin{cases} u_t = -u_x + g(x,t) \\ u(x,0) = 0. \end{cases}$$

Here we are writing the $-u_x$ term on the right to suggest that these problems are really "families" of ODE in t "indexed" by $x \in \mathbb{R}$. For example, if we fix $x \in \mathbb{R}$ and put $v(t) = u(x,t)$, then the transport equation is $v' = -u_x + g(x,t)$, which is morally an ODE in t.

Our hard work has shown that the solution to

$$\begin{cases} u_t = -u_x \\ u(x,0) = f(x) \end{cases}$$

is

$$u(x,t) = f(x-t),$$

while the solution to

$$\begin{cases} u_t = -u_x + g(x,t) \\ u(x,0) = 0 \end{cases}$$

is

$$u(x,t) = \int_0^t g(x - t + \tau, \tau) \, d\tau.$$

We introduce a new "propagator" that is "indexed" by x via

$$\mathcal{P}(t, x)f := f(x - t).$$

Then the solution to

$$\begin{cases} u_t = -u_x \\ u(x, 0) = f(x) \end{cases}$$

is

$$u(x, t) = \mathcal{P}(t, x)f.$$

Now fix τ and denote by $g(\cdot, \tau)$ the map

$$g(\cdot, \tau) \colon \mathbb{R} \to \mathbb{R} \colon X \mapsto g(X, \tau).$$

Then we can recognize the propagator in the solution to

$$\begin{cases} u_t = -u_x + g(x, t) \\ u(x, 0) = 0 \end{cases}$$

via

$$u(x, t) = \int_0^t g(x - t + \tau, \tau)\ d\tau = \int_0^t g(x - (t - \tau), \tau)\ d\tau = \int_0^t \mathcal{P}(t - \tau, x)g(\cdot, \tau)\ d\tau.$$

The solution to the full nonhomogeneous transport IVP is therefore

$$u(x, t) = \mathcal{P}(t, x)f + \int_0^t \mathcal{P}(t - \tau, x)(\cdot, \tau)\ d\tau.$$

Hopefully we see a pattern: the solution to the nonhomogeneous problem is the sum of the propagator applied to the initial data and the integral of the propagator "shifted by $t - \tau$" applied to the driving term.

This pattern is not wholly helpful for the driven wave equation, however, because that problem has two initial conditions. The right idea is to turn to the dreaded variation of parameters formula for second-order linear ODE. Here is a version of that formula that we typically do *not* see in standard ODE classes, as checking it requires differentiating under the integral.

14.7 Theorem (Variation of parameters). *Let b, $c \in \mathbb{R}$ and let $f \in \mathcal{C}(\mathbb{R})$. Suppose that $\mathcal{P} \in \mathcal{C}^2(\mathbb{R})$ solves*

$$\begin{cases} \mathcal{P}'' + b\mathcal{P}' + c\mathcal{P} = 0 \\ \mathcal{P}(0) = 0 \\ \mathcal{P}'(0) = 1. \end{cases}$$

Then for y_0, $y_1 \in \mathbb{R}$, the only solution to the IVP

$$\begin{cases} y'' + by' + cy = f(t) \\ y(0) = y_0 \\ y'(0) = y_1 \end{cases} \tag{14.2}$$

is

$$y(t) = \mathcal{P}'(t)y_0 + \mathcal{P}(t)(y_1 + y_0 b) + \int_0^t \mathcal{P}(t-\tau)f(\tau)\ d\tau. \tag{14.3}$$

In particular, the functions

$$z(t) := \mathcal{P}'(t)y_0 + \mathcal{P}(t)(y_1 + y_0 b) \quad \text{and} \quad y_\star(t) := \int_0^t \mathcal{P}(t-\tau)f(\tau)\ d\tau$$

solve the respective IVP

$$\begin{cases} z'' + bz' + cz = 0 \\ z(0) = y_0 \\ z'(0) = y_1 \end{cases} \quad \text{and} \quad \begin{cases} y_\star'' + by_\star' + cy_\star = f(t) \\ y_\star(0) = 0 \\ y_\star'(0) = 0. \end{cases} \tag{14.4}$$

Proving this theorem is challenging. First, one needs a uniqueness result for second-order linear IVP to guarantee the "only" result; we will not pursue that here. Second (or maybe first), what is the motivation for this formula? It is much less obvious than variation of parameters for first-order linear IVP, which effectively falls out from the product rule. The slickest way of proceeding for the second-order case is to convert that problem into a first-order linear *system*, which then has much in common with first-order (scalar) problems.

14.8 Problem. **(i)** Check that the formula (14.3) does yield a solution to (14.2). [Hint: *Lemma 11.5.*]

(ii) Let $\lambda \in \mathbb{R} \setminus \{0\}$. What does Theorem 14.7 say about the solution to

$$\begin{cases} y'' + \lambda^2 y = f(t) \\ y(0) = y_0 \\ y'(0) = y_1? \end{cases}$$

How does this resemble Problem 5.5?

Day 15: Monday, September 16.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 327–328 give a different motivation for Duhamel's formula. Page 329 states and proves the formula, and Example 6 on page 330 goes through the calculations for concrete initial and driving data.

Inspired by the propagators for ODE and the transport equation, we revisit the wave equation. The solution to the homogeneous problem

$$\begin{cases} u_{tt} = u_{xx}, & -\infty < x,\ t < \infty \\ u(x,0) = f(x), & -\infty < x < \infty \\ u_t(x,0) = g(x), & -\infty < x < \infty, \end{cases}$$

is given by D'Alembert's formula:

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

This morally resembles the first two terms of the solution (14.3) to the homogeneous second-order linear ODE (in the case $b = 0$) in that the initial data appears in each term separately. If we stare a little longer, we might see a resemblance between the terms in that

$$\partial_t\left[\frac{1}{2}\int_{x-t}^{x+t} g(s)\ ds\right] = \frac{g(x+t) + g(x-t)}{2}.$$

This is a consequence of a more general FTC identity.

15.1 Problem. Let I, $J \subseteq \mathbb{R}$ be intervals. Let $f \in \mathcal{C}(I)$ and a, $b \in \mathcal{C}^1(J)$ with $a(t)$, $b(t) \in I$ for all $t \in J$. Show that

$$\partial_t\left[\int_{a(t)}^{b(t)} f\right] = f(b(t))b'(t) - f(a(t))a'(t).$$

[Hint: *FTC1 + properties of integrals + chain rule.*]

Now define

$$\mathcal{P}(t,x)g := \frac{1}{2}\int_{x-t}^{x+t} g(s)\ ds. \tag{15.1}$$

The result above shows

$$\partial_t[\mathcal{P}(t,x)f] = \frac{f(x+t) + f(x-t)}{2},$$

and so D'Alembert's formula compresses to

$$u(x,t) = \partial_t[\mathcal{P}(t,x)f] + \mathcal{P}(t,x)g.$$

This strongly resembles the first two terms in (14.3)!

Consequently, by analogy with (14.4) we are led to conjecture that

$$u(x,t) := \int_0^t \mathcal{P}(t-\tau, x)h(\cdot,\tau)\ d\tau \tag{15.2}$$

solves

$$\begin{cases} u_{tt} = u_{xx} + h(x,t),\ -\infty < x,\ t < \infty \\ u(x,0) = 0,\ -\infty < x < \infty \\ u_t(x,0) = 0,\ -\infty < x < \infty, \end{cases}$$

We check the PDE and leave the initial conditions as an exercise. To do this, we need the identities

$$\partial_x \left[\int_0^t \phi(x,t,\tau)\ d\tau\right] = \int_0^t \phi_x(x,t,\tau)\, d\tau \quad \text{and} \quad \partial_t \left[\int_0^t \phi(x,t,\tau)\ d\tau\right] = \phi(x,t,t) + \int_0^t \phi_t(x,t,\tau)\, d\tau$$

for suitably well-behaved ϕ.

Then with u from (15.2), we have

$$u_{xx}(x,t) = \partial_x^2 \left[\int_0^t \mathcal{P}(t-\tau,x)h(\cdot,\tau)\ d\tau\right] = \int_0^t \partial_x^2\big[\mathcal{P}(t-\tau,x)h(\cdot,\tau)\big]\ d\tau.$$

Here we use the formula (15.1) to compute

$$\mathcal{P}(t-\tau,x)h(\cdot,\tau) = \frac{1}{2}\int_{x-(t-\tau)}^{x+(t-\tau)} h(s,\tau)\ ds = \frac{1}{2}\int_{x-t+\tau}^{x+t-\tau} h(s,\tau)\ ds,$$

and therefore, for τ fixed,

$$\partial_x\big[\mathcal{P}(t-\tau,x)h(\cdot,\tau)\big] = \frac{1}{2}\partial_x\left[\int_{x-t+\tau}^{x+t-\tau} h(s,\tau)\ ds\right] = \frac{h(x+t-\tau,\tau) - h(x-t+\tau,\tau)}{2}$$

and

$$\partial_x^2\big[\mathcal{P}(t-\tau,x)h(\cdot,\tau)\big] = \partial_x\left[\frac{h(x+t-\tau,\tau) - h(x-t+\tau,\tau)}{2}\right] = \frac{h_x(x+t-\tau,\tau) - h_x(x-t+\tau,\tau)}{2}.$$

Thus

$$u_{xx}(x,t) = \frac{1}{2}\int_0^t \big[h_x(x+t-\tau,\tau) - h_x(x-t+\tau,\tau)\big]\ d\tau. \tag{15.3}$$

Now we work on the time derivative. We have

$$u_t(x,t) = \partial_t\left[\int_0^t \mathcal{P}(t-\tau,x)h(\cdot,\tau)\ d\tau\right] = \mathcal{P}(t-t,x)h(\cdot,t) + \int_0^t \partial_t\big[\mathcal{P}(t-\tau,x)h(\cdot,\tau)\big]\ d\tau.$$

We compute

$$\mathcal{P}(t-t,x)h(\cdot,t) = \mathcal{P}(0,x)h(\cdot,t) = \frac{1}{2}\int_{x-0}^{x+0} h(s,t)\ ds = \frac{1}{2}\int_x^x h(s,t)\ ds = 0$$

and, for τ fixed,

$$\partial_t\big[\mathcal{P}(t-\tau,x)h(\cdot,\tau)\big] = \frac{1}{2}\partial_t\left[\int_{x-t+\tau}^{x+t-\tau} h(s,\tau)\ ds\right] = \frac{h(x+t-\tau,\tau)+h(x-t+\tau,\tau)}{2}.$$

Then

$$u_t(x,t) = \frac{1}{2}\int_0^t \big[h(x+t-\tau,\tau)+h(x-t+\tau,\tau)\big]\ d\tau,$$

so

$$u_{tt}(x,t) = \frac{h(x+t-t,t)+h(x-t+t,t)}{2} + \frac{1}{2}\int_0^t \partial_t\big[h(x+t-\tau,\tau)+h(x-t+\tau,\tau)\big]\ d\tau.$$

Certainly

$$\frac{h(x+t-t,t)+h(x-t+t,t)}{2} = \frac{h(x,t)+h(x,t)}{2} = h(x,t),$$

while

$$\int_0^t \partial_t\big[h(x+t-\tau,\tau)+h(x-t+\tau,\tau)\big]\ d\tau = \int_0^t \big[h_x(x+t-\tau,\tau)-h_x(x-t+\tau,\tau)\big]\ d\tau.$$

All together,

$$u_{tt}(x,t) = h(x,t) + \frac{1}{2}\int_0^t \big[h_x(x+t-\tau,\tau)-h_x(x-t+\tau,\tau)\big]\ d\tau = h(x,t)+u_{xx}(x,t),$$

after comparison to (15.3).

15.2 Problem. **(i)** Show that the function u defined in (15.2) satisfies

$$u(x,0) = u_t(x,0) = 0$$

for all $x \in \mathbb{R}$, and conclude that the function

$$u(x,t) = \partial_t[\mathcal{P}(t,x)f] + \mathcal{P}(t,x)g + \int_0^t \mathcal{P}(t-\tau,x)h(\cdot,\tau)\ d\tau \qquad (15.4)$$

solves the driven wave equation

$$\begin{cases} u_{tt} = u_{xx} + h(x,t), \ -\infty < x,\ t < \infty \\ u(x,0) = f(x),\ -\infty < x < \infty \\ u_t(x,0) = g(x),\ -\infty < x < \infty. \end{cases}$$

(ii) Show that the solution to the driven wave equation is unique. [Hint: *if u and v both solve it, what IVP does their difference $w := u - v$ solve, and why does that imply $w = 0$?*]

Any actual calculations with the formula (15.4) for concrete initial and driving terms f, g, and h boil down to computing antiderivatives, and there is probably not much insight to be gained from such manipulations at this point in life. Instead, here is a way to recognize the formula (15.2) as a double integral.

15.3 Problem. Let $h \in \mathcal{C}(\mathbb{R}^2)$ and let $x, t \in \mathbb{R}$. Let $\mathcal{D}(x,t)$ be the region in $\mathbb{R}^2$ consisting of the boundary and interior of the triangle whose endpoints are $(x-t,0)$, $(x+t,0)$, and (x,t). Show that

$$\int_0^t \mathcal{P}(t-\tau,x)h(\cdot,\tau)\,d\tau = \frac{1}{2}\iint_{\mathcal{D}(x,t)} h,$$

with the propagator $\mathcal{P}$ defined in (15.1). [Hint: *start by drawing $\mathcal{D}(x,t)$. For simplicity in solving this problem, you may assume $x > 0$ and $t > 0$, although the result is valid for all x and t.*]

Day 16: Wednesday, September 18.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 310–311 discuss solving the finite string wave equation with the "method of images." Theorem 3 contains the main result. It is not really necessary to assume that $\widetilde{f}_o$ and $\widetilde{g}_o$ are as regular as the theorem does; such regularity is forced on them by the "compatibility conditions," without which the problem really does not make sense.

Our best success with the wave equation involved having it posed spatially (in x) on $(-\infty,\infty)$. That gave us D'Alembert's formula, from which all good things flow. The next best situation was the semi-infinite string, which we reduced to the infinite string on $(-\infty,\infty)$ by carefully extending the initial data. We now consider the most physically realistic, but also most mathematically complicated, situation: the finite string.

Assume that a string of length $L > 0$ is constrained to move vertically with its endpoints fixed. If $u(x,t)$ is the displacement of the string from its equilibrium position at time t and spatial position $x \in [0,L]$, this means $u(0,t) = u(L,t) = 0$ for all t. We arrive at the initial-boundary value problem (IVP-BVP)

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ u(x,0) = f(x), \ 0 \le x \le L \\ u_t(x,0) = g(x), \ 0 \le x \le L \\ u(0,t) = u(L,t) = 0, \ -\infty < t < \infty. \end{cases} \tag{16.1}$$

As usual, f and g are the initial data, and we assume $f \in \mathcal{C}^2([0,L])$ and $g \in \mathcal{C}^1([0,L])$. (Why? We want a solution u to this problem to be twice continuously differentiable on $\{(x,t) \in \mathbb{R}^2 \mid 0 \le x \le L, \ t \in \mathbb{R}\}$. This forces $f = u(\cdot,0) \in \mathcal{C}^2([0,L])$ and the same for g.)

Our success with the semi-infinite string suggests that we extend f and g *carefully* to $(-\infty,\infty)$ and use D'Alembert's formula. By "carefully," we mean that the extensions should be sufficiently differentiable. We fool around with some values for f and g to see what extension might be the right one. Assume that the IVP-BVP (16.1) has a solution u. First,

$$f(0) = u(0,0) \text{ by the initial condition } f(x) = u(x,0), \ 0 \le x \le L$$

$$= 0 \text{ by the boundary condition } u(0,t) = 0, \ -\infty < t < \infty.$$

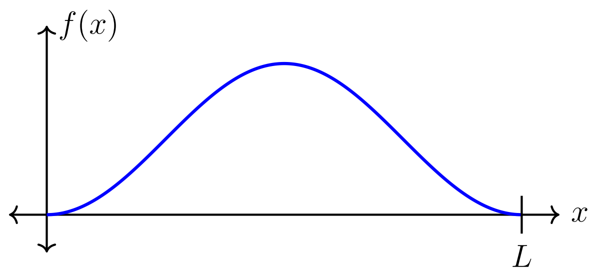
Next,

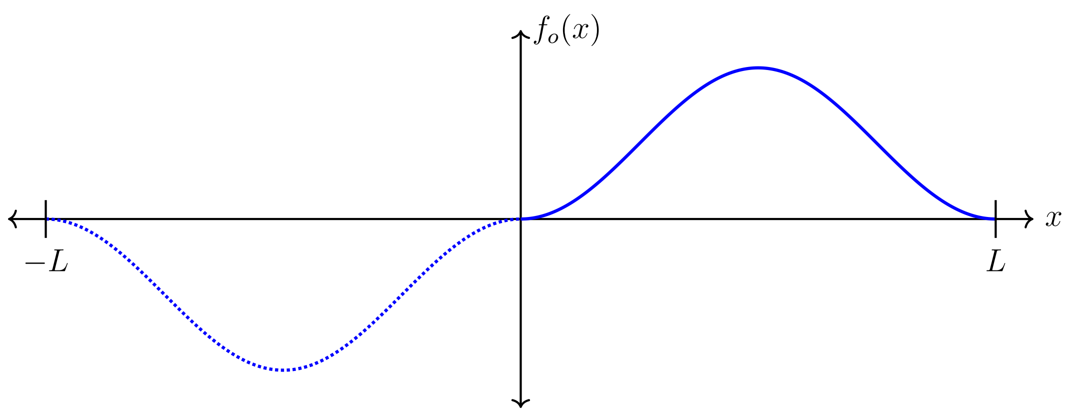
$$f''(x) = u_{xx}(x,0) \text{ by differentiating the initial condition } f(x) = u(x,0) \text{ twice in } x$$

$$= u_{tt}(x,0) \text{ because } u \text{ solves } u_{tt} = u_{xx}.$$

Thus $f''(0) = u_{tt}(0,0)$. Now we differentiate both sides of $u(0,t) = 0$ twice with respect to t to find $u_{tt}(0,t) = 0$ for all t. Thus $u_{tt}(0,0) = 0$, and so $f''(0) = 0$.

We conclude that if the IVP-BVP (16.1) has a solution, then the initial data f must meet the "compatibility conditions" $f(0) = f''(0) = 0$. Here is the picture; note how f resembles the graph of $y = x^2$ at the endpoints.





16.1 Problem. Continue to suppose that (16.1) has a solution.

(i) Show that $f(L) = f''(L) = 0$.

(ii) Show that $g(0) = 0$.

(iii) What happens if you try to get information on f' and g'?

These results suggest that we try to use the odd extensions (or reflections) of f and g. The challenge here is that because f and g are only defined on $[0, L]$, at best we could extend them to be odd on $[-L, L]$. We simply do not have enough data to continue that reflection outside this interval, since we only know the values of f and g on $[0, L]$. So, start with, as before

$$f_o \colon [-L, L] \to \mathbb{R} \colon x \mapsto \begin{cases} f(x), & 0 \le x \le L \\ -f(-x), & -L \le x < 0 \end{cases}$$

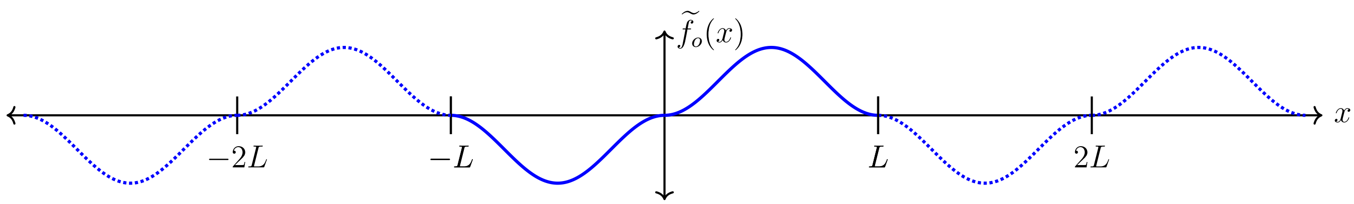
and

$$g_o \colon [-L, L] \to \mathbb{R} \colon x \mapsto \begin{cases} g(x), & 0 \le x \le L \\ -g(-x), & -L \le x < 0. \end{cases}$$

Now here is the clever part. We are going to extend f_o and g_o from $[-L, L]$ to $(-\infty, \infty)$ *periodically*. Informally, we "copy and paste" the graphs from $[-L, L]$ to the intervals $[(2k+1)L, (2k+3)L]$ for $k \in \mathbb{Z}$. Call these periodic extensions $\widetilde{f}_o$ and $\widetilde{g}_o$; we require them to satisfy

$$\widetilde{f}_o(x + 2L) = \widetilde{f}_o(x) \quad \text{and} \quad \widetilde{g}_o(x + 2L) = \widetilde{g}_o(x)$$

for all $x \in \mathbb{R}$.

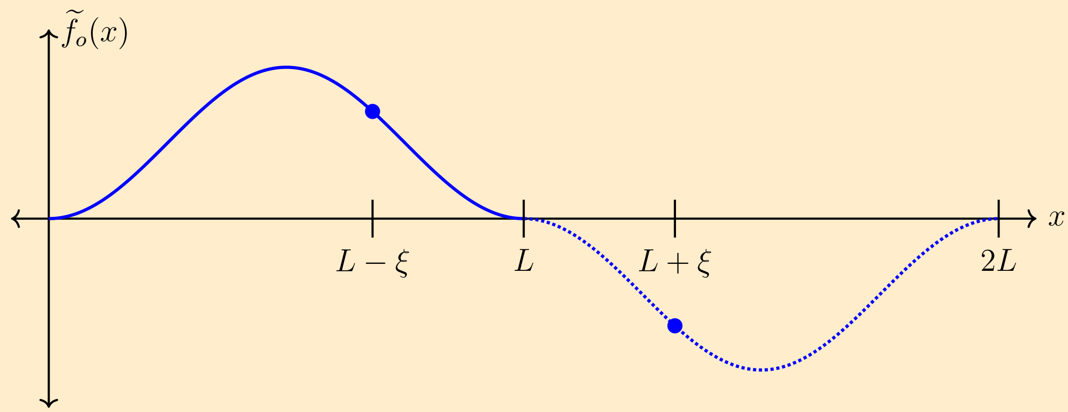


We are going to apply D'Alembert's formula to the problem

$$\begin{cases} u_{tt} = u_{xx}, & -\infty < x, \ t < \infty \\ u(x, 0) = \widetilde{f}_o(x), & -\infty < x < \infty \\ u_t(x, 0) = \widetilde{g}_o(x), & -\infty < x < \infty. \end{cases} \tag{16.2}$$

To do this, we need to be sure that $\widetilde{f}_o \in \mathcal{C}^2(\mathbb{R})$ and $\widetilde{g}_o \in \mathcal{C}^1(\mathbb{R})$, and for this to be worthwhile for our original finite string problem, we need to be sure that the solution produced by D'Alembert's formula meets the boundary conditions $u(0, t) = u(L, t) = 0$.

16.2 Problem (Optional, possibly annoying). **(i)** Show that $\widetilde{f}_o$ is twice differentiable at $x = L$ and that $\widetilde{f}_o''$ is continuous at $x = L$. It may help to argue first that $\widetilde{f}_o$ is "odd" about L in the sense that $\widetilde{f}_o(L-\xi) = -\widetilde{f}_o(L+\xi)$ for $0 \le \xi \le L$.



From this, argue by periodicity that $\widetilde{f}_o \in \mathcal{C}^2(\mathbb{R})$.

(ii) Solve (16.2) with D'Alembert's formula. Check the boundary conditions $u(0,t) = u(L,t) = 0$. [Hint: *use the oddness of $\widetilde{f}_o$, and also of $\widetilde{g}_o$, at L, as discussed above.*]

Day 17: Friday, September 20.

We took Exam 1.

Day 18: Monday, September 23.

Much of our work has concerned initial value problems. We are given initial-in-time data, and we build solutions out of that data. Often we obtain uniqueness results: there is only one solution to the differential equation at hand with the given initial data (Theorem 3.6, Theorem 3.8, Problem 5.4, Theorem 7.1, Theorem 11.1, Theorem 12.2, Theorem 14.7, Problem 15.2). Once uniqueness is established, a natural follow-up question is that of "continuous dependence on initial conditions." Very informally, this is motivated by the slogan *if two things start "close together" and move according to the "same rules," then they should remain "close together" at least for "some time."*

18.1 Problem. Suppose that y_1 and y_2 both solve

$$y' = ay + f(t)$$

for some $a \in \mathbb{R}$ and $f \in \mathcal{C}(\mathbb{R})$. Show that if $0 \le t \le T$, then

$$|y_1(t) - y_2(t)| \le e^{aT}|y_1(0) - y_2(0)|.$$

This gives a measure of how "close" y_1 and y_2 are on the interval $[0,T]$ in terms of T and the initial data.

We study this in the context of the wave equation. First, for functions f, $g \in \mathcal{C}(\mathbb{R})$, we define the "wave operator" $\mathcal{W}[f,g]$ by

$$\mathcal{W}[f,g](x,t) := \frac{f(x+t)+f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(s)\ ds. \tag{18.1}$$

Now let f_1, $f_2 \in \mathcal{C}^2(\mathbb{R})$ and g_1, $g_2 \in \mathcal{C}^1(\mathbb{R})$. Suppose that u and v solve the wave IVP

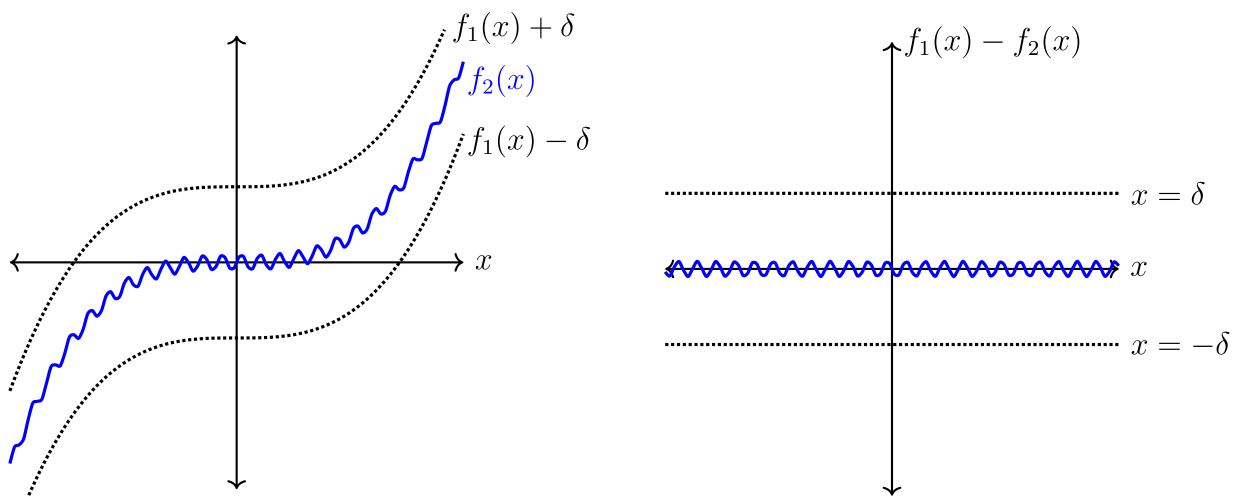
$$\begin{cases} u_{tt} = u_{xx}, & -\infty < x,\ t < \infty \\ u(x,0) = f_1(x), & -\infty < x < \infty \\ u_t(x,0) = g_1(x), & -\infty < x < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, & -\infty < x,\ t < \infty \\ v(x,0) = f_2(x), & -\infty < x < \infty \\ v_t(x,0) = g_2(x), & -\infty < x < \infty. \end{cases} \tag{18.2}$$

If f_1 and f_2 are "close," and if g_1 and g_2 are "close," will u and v be "close"?

First we spell out what we mean by "close." We assume there are δ, $\epsilon > 0$ such that

$$|f_1(x) - f_2(x)| < \delta \quad \text{and} \quad |g_1(x) - g_2(x)| < \epsilon$$

for all $x \in \mathbb{R}$. This means that the graph of f_2 lies between the graphs of $f_1 - \delta$ and $f_1 + \delta$, a sort of "δ-tube" centered on the graph of f_1; equivalently, the graph of $f_1 - f_2$ lies in the "strip" between $-\delta$ and δ. The same, of course, holds for g_1 and g_2 with δ replaced by ϵ. (Later we will see that there are other ways of measuring closeness of functions via different "norms" on function spaces—many involve integrals as a measurement of "averaging.")



18.2 Problem. Suppose that u and v solve the wave IVP in (18.2). Let $w = u - v$, $f = f_1 - f_2$, and $g = g_1 - g_2$. Show that $w = \mathcal{W}[f,g]$ with $\mathcal{W}$ defined in (18.1).

Our task is now to control the size of w, ideally in terms of δ and ϵ. We use the notation of the preceding problem. Since $w = \mathcal{W}[f,g]$, we have

$$|w(x,t)| \leq \frac{|f(x+t)+f(x-t)|}{2} + \frac{1}{2}\left|\int_{x-t}^{x+t} g(s)\ ds\right|.$$

The triangle inequality on the first term implies

$$|f(x+t)+f(x-t)| \le |f(x+t)|+|f(x-t)|,$$

and then the triangle inequality on f implies

$$|f(x+t)| = |f_1(x+t)-f_2(x+t)| < \delta,$$

and the same for $|f(x-t)|$. All together,

$$\frac{|f(x+t)+f(x-t)|}{2} < \frac{\delta+\delta}{2} = \delta.$$

We estimate the integral with the triangle inequality for integrals (Problem 2.4):

$$\left|\int_{x-t}^{x+t} g(s)\ ds\right| \le \int_{x-t}^{x+t} |g(s)|\ ds,$$

at least if $x-t \le x+t$, i.e., if $t \ge 0$. Since

$$|g(s)| = |g_1(s)-g_2(s)| < \epsilon$$

for all $s \in \mathbb{R}$, this implies

$$\left|\int_{x-t}^{x+t} g(s)\ ds\right| \le \int_{x-t}^{x+t} \epsilon\ ds = 2t\epsilon$$

when $t \ge 0$.

18.3 Problem. Show that if $t < 0$, then

$$\left|\int_{x-t}^{x+t} g(s)\ ds\right| < 2|t|\epsilon.$$

We conclude

$$|u(x,t)-v(x,t)| = |w(x,t)| = |\mathcal{W}[f,g](x,t)| < \delta + |t|\epsilon. \tag{18.3}$$

This shows that for any fixed time $t \in \mathbb{R}$, the solutions u and v are uniformly close in x in a manner depending precisely on how close the initial conditions are.

However, this estimate is less than ideal because it depends on time t. As $t \to \pm\infty$, $\delta + |t|\epsilon \to \infty$, and so perhaps over long times the solutions u and v could grow apart.

18.4 Problem. Here is a somewhat silly example of how this could occur. Let δ, $\epsilon > 0$. Take $f_1 = g_1 = 0$ and $f_2(x) = \delta/2$ and $g_2(x) = \epsilon/2$. Show that if u and v solve (18.2), then

$$u(x,t) = 0 \quad \text{and} \quad v(x,t) = \frac{\delta+\epsilon t}{2}.$$

Check explicitly that (18.3) still holds, but explain informally how u and v "grow apart" in time.

The factor of $|t|$ in (18.3) arose from from estimating the integral term in $\mathcal{W}[f, g]$. A recurring tension in analysis is whether estimates or equalities are preferable; perhaps, depending on g, we could get sharper control over $\int_{x-t}^{x+t} g(s)\ ds$ by actually computing it. It turns out that we can get a better estimate than (18.3) if we ask a different question, and so we focus on the finite string problem. (The question of continuous dependence on initial conditions for the semi-infinite string would yield the same estimate as above.)

Let $L > 0$ and let u and v now solve

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ u(x,0) = f_1(x), \ 0 \le x \le L \\ u_t(x,0) = g_1(x), \ 0 \le x \le L \\ u(0,t) = u(L,t) = 0, \ -\infty < t < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ v(x,0) = f_2(x), \ 0 \le x \le L \\ v_t(x,0) = g_2(x), \ 0 \le x \le L \\ v(0,t) = v(L,t) = 0, \ -\infty < t < \infty. \end{cases} \tag{18.4}$$

Put $f = f_1 - f_2$ and $g = g_1 - g_2$, and let $\widetilde{f}_o$ and $\widetilde{g}_o$ be the $2L$-periodic, odd extensions, i.e.,

$$\widetilde{f}_o(x) = \begin{cases} f(x), \ 0 \le x \le L \\ -f(-x), \ -L \le x < 0 \end{cases} \quad \text{and} \quad \widetilde{f}_o(x + 2L) = \widetilde{f}_o(x), \ x \in \mathbb{R}.$$

Assume that the initial data satisfies all the hypotheses necessary for $w = \mathcal{W}[\widetilde{f}_o, \widetilde{g}_o]$ to solve

$$\begin{cases} w_{tt} = w_{xx}, \ -\infty < x, \ t < \infty \\ w(x,0) = \widetilde{f}_o(x), \ -\infty < x < \infty \\ w_t(x,0) = \widetilde{f}_o(x), \ -\infty < x < \infty, \end{cases}$$

so, restricted to $[0, L]$, w also solves

$$\begin{cases} w_{tt} = w_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ w(x,0) = f(x), \ 0 \le x \le L \\ w_t(x,0) = g(x), \ 0 \le x \le L \\ w(0,t) = w(L,t) = 0, \ -\infty < t < \infty. \end{cases}$$

And now we start to estimate. Assume there are δ, $\epsilon > 0$ such that

$$|f_1(x) - f_2(x)| < \delta \quad \text{and} \quad |g_1(x) - g_2(x)| < \epsilon$$

for all $x \in [0, L]$.

18.5 Problem. Explain why

$$|\widetilde{f}_o(x)| < \delta \quad \text{and} \quad |\widetilde{g}_o(x)| < \epsilon$$

for all $x \in \mathbb{R}$.

It follows as before that

$$\left|\frac{\widetilde{f}_o(x+t)+\widetilde{f}_o(x-t)}{2}\right| < \delta$$

for all x, $t \in \mathbb{R}$. The difference is that the integral term in $\mathcal{W}[\widetilde{f}_o, \widetilde{g}_o]$ will be much better behaved.

Day 19: Wednesday, September 25.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

The corollary on p. 314 deduces continuous dependence on initial conditions for the finite string problem from Theorem 5 on pp. 313–314. That theorem proves the hard-won estimate with the independent-of-time upper bound $\delta + 2L\epsilon$ that we eke out. Theorem 1 on p. 289 proves uniqueness for the finite string problem via energy estimates. The remark on pp. 290–291 explains how to interpret that energy integral in terms of classical kinetic + potential energy.

Here is how we do *not* get that better behavior: do what we did before and expect something to change. We could estimate

$$\left|\int_{x-t}^{x+t} \widetilde{g}_o(s)\ ds\right| < 2|t|\epsilon$$

exactly as for the infinite string using the triangle inequality for integrals, and that still produces the annoying factor of t in the estimate. We can do better by using the special structure of $\widetilde{g}_o$ here: it is odd and $2L$-periodic in addition to enjoying the estimate $|\widetilde{g}_o(s)| < \epsilon$ for all s.

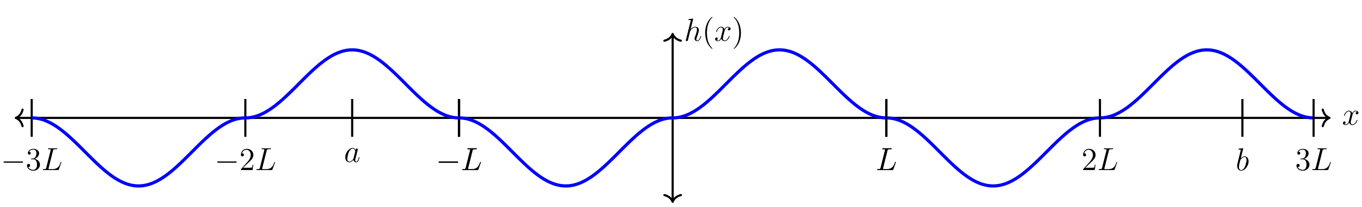
To cut down on writing, we let $h \in \mathcal{C}(\mathbb{R})$ be odd and $2L$-periodic with $|h(x)| < \epsilon$ for all x. We claim that

$$\int_c^{c+2L} h = 0 \tag{19.1}$$

for all $c \in \mathbb{R}$; in words, the integral of h over any interval of length $2L$ vanishes.

19.1 Problem. Prove this. [Hint: *use the results of Problems 3.3 and 14.2.*]

Here is what we will show: the value of $\int_a^b f$ is bounded by a constant multiple of ϵ independent of a and b (but dependent on L). We start with a suggestive proof by picture. Here $-2L < a < -L$ and $2L < b < 3L$.



We expand

$$\int_a^b h = \int_a^{-L} h + \int_{-L}^{L} h + \int_L^b h. \tag{19.2}$$

By (19.1), or the cancelation of positive and negative areas from the picture, $\int_{-L}^{L} h = 0$. Thus by the triangle inequality for real numbers and the triangle inequality for integrals,

$$\left|\int_a^b h\right| = \left|\int_a^{-L} h + \int_L^b h\right| \le \left|\int_a^{-L} h\right| + \left|\int_L^b h\right| \le \int_a^{-L} |h| + \int_L^b |h|.$$

Now we use the estimate on h and actually evaluate some integrals:

$$\left|\int_a^b h\right| \le \int_a^{-L} |h| + \int_L^b |h| < \int_a^{-L} \epsilon + \int_L^b \epsilon = \epsilon(-L-a) + \epsilon(b-L).$$

Since $-2L < a < -L$, we have $L < -a < 2L$, and so $0 < -L - a < L$. Since $L < b < 3L$, we have $0 < b - L < 2L$. And so

$$\left|\int_a^b h\right| < \epsilon(-L-a) + \epsilon(b-L) < L\epsilon + 2L\epsilon = 3L\epsilon. \tag{19.3}$$

Here is what happens more generally, beyond the special case of this picture. Let a, $b \in \mathbb{R}$ with $a < b$. Divide $\mathbb{R}$ into intervals of the form $[(2j+1)L, (2j+3)L)$ with $j \in \mathbb{Z}$. Then there are j, $k \in \mathbb{Z}$ such that

$$(2j+1)L \le a < (2j+3)L \quad \text{and} \quad (2k+1)L \le b < (2k+3)L. \tag{19.4}$$

In the picture above, we have $-3L \le a < -L$ and $L \le b < 3L$, so there $j = -2$ and $k = 0$. In the general case, since $a < b$, it follows that $j \le k$.

19.2 Problem. Does it? If $a < b$, then the inequalities above imply $(2j+1)L \le a < b < (2k+3)L$. Manipulate this into $j < k + 1$. Since j and k are integers, this means $j \le k$.

Now we expand the integral again:

$$\int_a^b h = \int_a^{(2j+3)L} h + \left(\int_{(2j+3)L}^{(2j+5)L} h + \int_{(2j+5)L}^{(2j+7)L} h + \cdots + \int_{(2k-1)L}^{(2k+1)L} h + \int_{(2k+1)L}^{(2k+3)L} h\right) + \int_{(2k+3)L}^{b} h.$$

The parenthetical sum here boiled down to the single integral $\int_{-L}^{L} h$ in the toy calculation (19.2). Every integral in the parenthetical sum is 0 by (19.1). Thus

$$\left|\int_a^b h\right| = \left|\int_a^{(2j+3)L} h + \int_{(2k+3)L}^{b} h\right| \le \left|\int_a^{(2j+3)L} h\right| + \left|\int_{(2k+3)L}^{b} h\right| \le \int_a^{(2j+3)L} |h| + \int_{(2k+3)L}^{b} |h|$$

$$< \int_a^{(2j+3)L} \epsilon + \int_{(2k+3)L}^{b} \epsilon = \epsilon((2j+3)L - a) + \epsilon(b - (2k+3)L).$$

The estimates (19.4) imply

$$(2j+3)L - a < 2L \quad \text{and} \quad b - (2k+3)L < 2L.$$

All together,

$$\left|\int_a^b h\right| < \epsilon((2j+3)L - a) + \epsilon(b - (2k+3)L) < 2L\epsilon + 2L\epsilon = 4L\epsilon.$$

This is a slightly worse estimate (in that the right side is larger) than our toy calculation that gave us (19.3).

19.3 Problem. Why? What was special about the positioning of a in that toy drawing? Why will that not always be the case, as compared to (19.4)?

But it is not a big deal. The point is that the size of $\int_a^b h$ is indeed controlled by a constant multiple of ϵ, with the constant independent of a and b.

At last, here is how this is useful. All along the goal has been to estimate $\int_{x-t}^{x+t} \widetilde{g}_o(s)\ ds$. We know that $\widetilde{g}_o$ is continuous, odd, and $2L$-periodic with $|\widetilde{g}_o(s)| < \epsilon$ for all $s \in \mathbb{R}$. Our work above therefore implies (with $a = x - t$ and $b = x + t$) that

$$\left|\int_{x-t}^{x+t} \widetilde{g}_o(s)\ ds\right| < 4L\epsilon.$$

With u and v as solutions to (18.4), all of our work implies

$$|u(x,t) - v(x,t)| < \delta + 2L\epsilon$$

for all $x \in [0, L]$ and $t \in \mathbb{R}$. This is the uniform-in-time estimate that we were lacking for the infinite string wave equation.

It has taken us some time, but now we can state a general result for wave IVP.

19.4 Theorem. **(i)** *Let $f_1,\ f_2 \in \mathcal{C}^2(\mathbb{R})$ and $g_1,\ g_2 \in \mathcal{C}^1(\mathbb{R})$. Suppose that $\delta,\ \epsilon > 0$ with*

$$|f_1(x) - f_2(x)| < \delta \quad \text{and} \quad |g_1(x) - g_2(x)| < \epsilon$$

for all $x \in \mathbb{R}$. Let u and v solve

$$\begin{cases} u_{tt} = u_{xx}, & -\infty < x,\ t < \infty \\ u(x,0) = f_1(x), & -\infty < x < \infty \\ u_t(x,0) = g_1(x), & -\infty < x < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, & -\infty < x,\ t < \infty \\ v(x,0) = f_2(x), & -\infty < x < \infty \\ v_t(x,0) = g_2(x), & -\infty < x < \infty. \end{cases} \tag{19.5}$$

Then

$$|u(x,t) - v(x,t)| < \delta + |t|\epsilon$$

for all $x,\ t \in \mathbb{R}$.

(ii) *Let* $L > 0$ *and* $f_1,\ f_2 \in \mathcal{C}^2([0,L])$ *and* $g_1,\ g_2 \in \mathcal{C}^1([0,L])$ *with*

$$f_1(x) = f_1''(x) = f_2(x) = f_2''(x) = g_1(x) = g_2(x) = 0$$

for $x = 0,\ L$. *Suppose that* $\delta,\ \epsilon > 0$ *with*

$$|f_1(x) - f_2(x)| < \delta \quad \text{and} \quad |g_1(x) - g_2(x)| < \epsilon$$

for all $x \in \mathbb{R}$. *Let* u *and* v *solve*

$$\begin{cases} u_{tt} = u_{xx},\ 0 \le x \le L,\ -\infty < t < \infty \\ u(x,0) = f_1(x),\ 0 \le x \le L \\ u_t(x,0) = g_1(x),\ 0 \le x \le L \\ u(0,t) = u(L,t) = 0,\ -\infty < t < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx},\ 0 \le x \le L,\ -\infty < t < \infty \\ v(x,0) = f_2(x),\ 0 \le x \le L \\ v_t(x,0) = g_2(x),\ 0 \le x \le L \\ v(0,t) = v(L,t) = 0,\ -\infty < t < \infty. \end{cases} \tag{19.6}$$

Then

$$|u(x,t) - v(x,t)| < \delta + 2L\epsilon$$

for all $x,\ t \in \mathbb{R}$.

There is just one embarrassing gap in our results above. We have long since proved uniqueness for the infinite string wave equations in (19.5) via D'Alembert's formula. However, for the finite string problems in (19.6), we merely used the odd periodic extensions to construct *a* solution via D'Alembert. We never proved that it was unique.

19.5 Problem (Optional). To be fair, we never proved uniqueness for the semi-infinite string, either. Here is how to do that. Let $f \in \mathcal{C}^2([0,\infty))$ and $g \in \mathcal{C}^1([0,\infty))$ with

$$f(0) = f''(0) = g(0) = 0.$$

Suppose that u solves

$$\begin{cases} u_{tt} = u_{xx},\ 0 \le x < \infty,\ -\infty < t < \infty \\ u(x,0) = f(x),\ 0 \le x < \infty \\ u_t(x,0) = g(x),\ 0 \le x < \infty. \end{cases}$$

Define

$$u_o(x,t) := \begin{cases} u(x,t),\ x \ge 0,\ t \in \mathbb{R} \\ u(-x,t),\ x < 0,\ t \in \mathbb{R}. \end{cases}$$

Check that u_o solves

$$\begin{cases} (u_o)_{tt} = (u_o)_{xx},\ -\infty < x,\ t < \infty \\ u_o(x,0) = f_o(x),\ -\infty < x < \infty \\ (u_o)_t(x,0) = g_o(x), -\infty < x < \infty, \end{cases}$$

where f_o and g_o are the usual odd extensions of f and g to $\mathbb{R}$. [Hint: *the real work is making sure u_o is sufficiently differentiable at $x = 0$.*] Conclude that u_o is given by D'Alembert's formula, and so this determines the values of u uniquely in terms of f and g.

D'Alembert's formula will not help us get uniqueness for the finite string problem. Instead, we introduce a totally new method, which will reappear for other PDE in the future. We study what is called an "energy integral." In the mathematical jargon, an "energy integral" refers to the integral (definite or improper) of some nonnegative function that, through the right lens, might represent some physical notion of "energy," kinetic or potential (whatever that means).

Here is how this arises. Suppose that u and v both solve finite string problems with the same initial data:

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ u(x,0) = f(x), \ 0 \le x \le L \\ u_t(x,0) = g(x), \ 0 \le x \le L \\ u(0,t) = u(L,t) = 0 \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ v(x,0) = f(x), \ 0 \le x \le L \\ v_t(x,0) = g(x), \ 0 \le x \le L \\ v(0,t) = v(L,t) = 0. \end{cases}$$

Put $w := u - v$.

19.6 Problem. Check that

$$\begin{cases} w_{tt} = w_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ w(x,0) = 0, \ 0 \le x \le L \\ w_t(x,0) = 0, \ 0 \le x \le L \\ w(0,t) = w(L,t) = 0. \end{cases} \tag{19.7}$$

We would like to show that $w = 0$. To do this, we set

$$E(t) := \int_0^L \left[w_t(x,t)^2 + w_x(x,t)^2\right] dx.$$

This is our "energy integral"; it is the integral of a nonnegative quantity. We claim that E is differentiable and $E'(t) = 0$ for all t.

We will check this later. For now, here is how it helps. If $E' = 0$, this means that E is constant; one helpful value is probably $t = 0$, so we compute

$$E(t) = E(0) = \int_0^L \left[w_t(x,0)^2 + w_x(x,0)^2\right] dx$$

for all t. From the initial conditions, $w_t(x,0) = 0$ and, since $w(x,0) = 0$ for all x, we have $w_x(x,0) = 0$ for all x, too. Thus $E(0) = 0$. And, since E is constant, so too do we have $E(t) = 0$ for all t.

Now, observe that each $E(t)$ is the integral of a nonnegative function. This is important.

19.7 Problem. Let $a, b \in \mathbb{R}$ with $a < b$ and $f \in \mathcal{C}([a,b])$ with $f(x) \geq 0$ for each $x \in [a,b]$. If $\int_a^b f = 0$, show that $f(x) = 0$ for all $x \in [a,b]$. [Hint: *suppose* $f(x_0) \neq 0$ *for some* $x_0 \in [a,b]$. *Draw a picture. What does this imply about the value of* $\int_a^b f$*? Turn the picture into a proof. Continuity will play a role.*]

This problem, together with the result $E(t) = 0$ and the definition of $E(t)$, implies

$$w_t(x,t)^2 + w_x(x,t)^2 = 0 \tag{19.8}$$

for all $x \in [0,L]$ and $t \in \mathbb{R}$. We have all but arrived at $w(x,t) = 0$ for all $x \in [0,L]$ and $t \in \mathbb{R}$.

19.8 Problem. Use (19.8), the boundary condition $w(0,t) = 0$, and FTC2 to show that $w(x,t) = 0$ for all $x \in [0,L]$ and $t \in \mathbb{R}$.

Our last task is to justify the earlier claim that $E' = 0$.

19.9 Problem. Compute $E'(t)$ by differentiating under the integral; use the identity $w_{tt} = w_{xx}$ to replace a factor that appears in the differentiated integrand; and then recognize the integrand as a perfect derivative in x. Be sure to explain why, if w solves (19.7), the hypotheses of Leibniz's rule (Theorem 11.2) are met. After that, use FTC2 and the boundary conditions to obtain $E'(t) = 0$.

Day 20: Friday, September 27.

No class due to university closure.

Day 21: Monday, September 30.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

While we only stated and did not really discuss the heat equation (yet), there is a wealth of information in the book. Pages 121–125 give a derivation of the heat equation from physical principles and present one very special solution.

We begin our study of the heat equation on the line:

$$\begin{cases} u_t = u_{xx}, \ -\infty < x < \infty, \ t > 0 \\ u(x,0) = f(x), \ -\infty < x < \infty. \end{cases}$$

Broadly, the heat equation models the distribution of heat in an infinitely long rod; the function f specifies the initial heat distribution along the rod. As with the wave equation, we start with this physically unrealistic situation of an infinite spatial domain, and eventually we will move to the more physically realistic (and mathematically complicated) "finite" rod.

The heat equation might look superficially similar to the wave equation; after all, both have the term u_{xx} on one side of the equation. We might even think that the heat equation is *simpler* than the wave equation in that only one time derivative appears. Not so! The "imbalance" of derivatives in the heat equation vastly complicates it. We will not have such a sweeping D'Alembert's formula for the heat equation, and both existence and uniqueness of solutions becomes much trickier here.

In fact, we need entirely new tools to tackle the heat equation. Our success with the transport and wave equations arose fundamentally from familiar calculus. Now we need unfamiliar calculus. We start by building some machinery in two areas that may appear to have nothing to do with the heat equation, or PDE in general: the essential calculus of complex-valued functions of a real variable (good news: it is the same as the essential calculus of real-valued functions of a real variable) and improper integrals.

Here is a terrible definition of complex numbers.

21.1 Undefinition. $\mathbb{C} = \{x + iy \mid x,\ y \in \mathbb{R},\ i^2 = -1\}$.

This definition is terrible because it provides no explanation of what the string of symbols $x + iy$ actually means or why such an object i actually exists. We just assume the existence of complex numbers and that their arithmetical properties act as they should.

21.2 Definition. *Let $z \in \mathbb{C}$ with $z = x + iy$ for some $x,\ y \in \mathbb{R}$. The* REAL PART *of z is $\text{Re}(z) := x$; the* IMAGINARY PART *of z is $\text{Im}(z) := y$; and the* MODULUS *of z is $|z| := \sqrt{\text{Re}(z)^2 + \text{Im}(z)^2}$. That is, $|x + iy| = \sqrt{x^2 + y^2}$.*

We define equality of $z,\ w \in \mathbb{C}$ as $z = w$ if and only if both $\text{Re}(z) = \text{Re}(w)$ and $\text{Im}(z) = \text{Im}(w)$.

21.3 Example. With $z = 2+i$ and $w = 1-3i$, we multiply as we would with real numbers and remember $i^2 = -1$:

$$\begin{aligned} zw = (2+i)(1-3i) = (2+i)1+(2+i)(-3i) = 2+i-6i-3i^2 = 2-5i-3(-1) = 2-5i+3 \\ = 5 - 5i = 5(1-i). \end{aligned}$$

Since the modulus satisfies $|zw| = |z||w|$, we have (with $z = 5$ and $w = 1 - i$, now)

$$|5(1-i)| = |5||1-i| = |5||1 + (-1)i| = |5|\sqrt{2}.$$

Here is a crash course in complex calculus. Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{C}$ be a function. Put

$$f_1(t) := \text{Re}[f(t)] \quad \text{and} \quad f_2(t) := \text{Im}[f(t)].$$

Then $f_1,\ f_2\colon I \to \mathbb{R}$ are functions, and *real-valued* functions at that, and $f(t) = f_1(t) + if_2(t)$. Now we do calculus.

21.4 Definition. *With the notation above, we say that*

(i) $\lim_{t\to a} f(t) = L$ *if* $\lim_{t\to a} f_1(t) = \mathrm{Re}[L]$ *and* $\lim_{t\to a} f_2(t) = \mathrm{Im}[L]$ *(with* $a = \pm\infty$ *allowed);*

(ii) f *is* CONTINUOUS *if* f_1 *and* f_2 *are continuous;*

(iii) f *is* DIFFERENTIABLE *if* f_1 *and* f_2 *are differentiable, and we define*

$$f'(t) := f_1'(t) + i f_2'(t);$$

(iv) *if* f *is continuous (in the sense of the above), then for any* a, $b \in I$, *we define*

$$\int_a^b f := \int_a^b f_1 + i \int_a^b f_2.$$

From these definitions, one can prove that all the familiar computational rules of real-valued calculus hold, e.g., the product and chain rules for differentiation, the linearity of the integral in the integrand, and the fundamental theorem of calculus. We will do none of that explicitly and just assume that everything works as it should.

Our most important complex-valued function of a real variable is the following version of the exponential.

21.5 Definition. *For* $t \in \mathbb{R}$, *let* $e^{it} := \cos(t) + i\sin(t)$.

Motivation for this definition comes from inserting it into the power series for the (real) exponential, doing some algebra, and recognizing the series for sine and cosine.

21.6 Example. Here is how calculus works for the exponential. Let $f(t) := e^{it}$. Then, with the notation above, $f_1(t) = \cos(t)$ and $f_2(t) = \sin(t)$, so

$$f'(t) = -\sin(t) + i\cos(t) = i^2\sin(t) + i\cos(t) = i[i\sin(t) + \cos(t)] = ie^{it}.$$

That is, the chain rule formula

$$f'(t) = \partial_t[e^{it}] = e^{it}\partial_t[it] = e^{it}i$$

works as we expect.

Now we integrate:

$$\int_0^{2\pi} f = \int_0^{2\pi} \cos(t)\ dt + i\int_0^{2\pi} \sin(t)\ dt = 0 + i0 = 0.$$

We also have

$$\int_0^{2\pi} f = \int_0^{2\pi} e^{it}\ dt = \frac{1}{i}\int_0^{2\pi} ie^{it}\ dt = \frac{1}{i}\int_0^{2\pi} f'(t)\ dt = \frac{1}{i}[f(2\pi) - f(0)] = \frac{1}{i}[1-1] = 0.$$

Here we are using the identity $e^{2\pi ik} = 1$ for all $k \in \mathbb{Z}$.

Now we develop further results on integrals.

21.7 Definition. *Let $f: \mathbb{R} \to \mathbb{C}$ be continuous. Suppose that both of the limits*

$$\int_{-\infty}^{0} f := \lim_{a \to -\infty} \int_{0}^{a} f \quad \text{and} \quad \int_{0}^{\infty} f := \lim_{b \to \infty} \int_{0}^{b} f$$

exist. Then we say that f is **INTEGRABLE**, *and we define*

$$\int_{-\infty}^{\infty} f := \int_{-\infty}^{0} f + \int_{0}^{\infty} f.$$

21.8 Example. Let $f(t) = e^{-|t|}$ and $a < 0$ and $b > 0$. We compute some integrals:

$$\int_{a}^{0} f = \int_{a}^{0} e^{-|t|}\, dt = \int_{a}^{0} e^{t}\, dt = e^{0} - e^{a} = 1 - e^{a}$$

and

$$\int_{0}^{b} f = \int_{0}^{b} e^{-|t|}\, dt = \int_{0}^{b} e^{-t}\, dt = -(e^{-b} - e^{-0}) = 1 - e^{-b}.$$

Then

$$\lim_{a \to -\infty} \int_{a}^{0} f = \lim_{a \to -\infty} (1 - e^{a}) = 1 \quad \text{and} \quad \lim_{b \to \infty} \int_{0}^{b} f = \lim_{b \to \infty} (1 - e^{-b}) = 1,$$

so $\int_{-\infty}^{0} f = \int_{0}^{\infty} f = 1$. Thus f is integrable and

$$\int_{-\infty}^{\infty} f = 1 + 1 = 2.$$

It is often both difficult to establish that f is integrable and unnecessary to calculate $\int_{-\infty}^{\infty} f$ exactly. Instead, the following tests usually suffice.

21.9 Theorem. *Let $f: \mathbb{R} \to \mathbb{C}$ be continuous.*

(i) [Absolute integrability implies integrability] *If $|f|$ is integrable, then so is f, and the* **TRIANGLE INEQUALITY** *holds:*

$$\left| \int_{-\infty}^{\infty} f \right| \le \int_{-\infty}^{\infty} |f|.$$

(ii) [Comparison test] *Suppose that $g: \mathbb{R} \to \mathbb{C}$ is continuous with $|g|$ integrable and $|f(t)| \le |g(t)|$ for all t. Then $|f|$, and thus f, are integrable, and*

$$\int_{-\infty}^{\infty} |f| \le \int_{-\infty}^{\infty} |g|.$$

21.10 Example. Let $f\colon \mathbb{R} \to \mathbb{C}$ be continuous with $|f|$ integrable. Let $k \in \mathbb{R}$ and put $h(t) := f(t)e^{ikt}$. Since $|e^{is}| = 1$ for all $s \in \mathbb{R}$ (check it), we have

$$|f(t)e^{ikt}| = |f(t)||e^{ikt}| = |f(t)|,$$

and so by the comparison test (with actual equality holding), the functions h and $|h|$ are integrable.

21.11 Problem. Let $a > 0$ and let $f(x) = e^{-ax^2}$. Show that f is integrable. [Hint: *first find $C > 0$ such that $e^{ax-ax^2} \le C$ for $0 \le x \le 1$. Then argue that $e^{-ax^2} \le e^{-ax}$ for $x \ge 1$. Put these estimates together to show $e^{-ax^2} \le (C+1)e^{-ax}$ for $x \ge 0$.*]

21.12 Problem. It is important in the definition of the improper integral to specify the convergence of the integrals $\int_{-\infty}^{0} f$ and $\int_{0}^{\infty} f$ separately. If $f\colon \mathbb{R} \to \mathbb{C}$ is continuous and if $\lim_{R\to\infty} \int_{-R}^{R} f$ exists, then we call this limit the **Cauchy principal value** of the improper integral of f over $(-\infty, \infty)$, and we might write

$$\mathrm{P.V.} \int_{-\infty}^{\infty} f := \lim_{R\to\infty} \int_{-R}^{R} f.$$

(i) Give an example of a continuous function $f\colon \mathbb{R} \to \mathbb{C}$ such that $\lim_{R\to\infty} \int_{-R}^{R} f$ exists and yet f is not integrable.

(ii) If, however, f is integrable, then $\int_{-\infty}^{\infty} f = \mathrm{P.V.} \int_{-\infty}^{\infty} f$. Here is why. Assume that $f\colon \mathbb{R} \to \mathbb{C}$ is integrable and let $\epsilon > 0$. Explain why there exists $R_0 > 0$ such that if $R > R_0$, then

$$\left|\int_{-\infty}^{0} f - \int_{-R}^{0} f\right| < \frac{\epsilon}{2} \quad \text{and} \quad \left|\int_{0}^{\infty} f - \int_{0}^{R} f\right| < \frac{\epsilon}{2}.$$

Use this to show that

$$\left|\int_{-\infty}^{\infty} f - \int_{-R}^{R} f\right| < \epsilon,$$

and conclude that $\int_{-\infty}^{\infty} f = \lim_{R\to\infty} \int_{-R}^{R} f$.

(iii) Something special happens when we try to integrate a nonnegative function. The following is true in general: if $g\colon [0,\infty) \to [0,\infty)$ is continuous, increasing ($g(x_1) \le g(x_2)$ for $0 \le x_1 \le x_2$), and bounded above (there is $M > 0$ such that $0 \le g(x) \le M$ for all $x \ge 0$), then $\lim_{x\to\infty} g(x)$ exists. The proof of this result depends on the completeness of the real numbers, but drawing a picture probably suggests why it is true. Draw such a picture. Then use this result to show that if $f\colon \mathbb{R} \to [0,\infty)$ is continuous, and if there is $M > 0$ such that $\left|\int_{-R}^{R} f\right| \le M$ for all $R \ge 0$, then f is integrable. [Hint: *apply the result to the functions $R \mapsto \int_{0}^{R} f$ and $R \mapsto \int_{-R}^{0} f$.*]

Day 22: Wednesday, October 2.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 415–418 give an overview of transforms, including but not limited to the Fourier. This is extremely worthwhile reading for the mathematical cultural background that it provides. Integrability and the Fourier transform are defined on p. 423; note the symmetric limit in (8), which is not how we defined improper integrals. See also the remark on the Cauchy principal value at the bottom of p. 423/top of p. 424.

Many of the "nice" function properties that we are assuming today are spelled out in Section 7.2. We will revisit quite a few of these as we layer more rigor over our Fourier analysis. Our derivation of the heat equation solution appears on pp. 460–461, with plenty of references to other parts of Chapter 7 that we have not quite discussed yet (including convolutions).

We introduce the critical tool of the Fourier transform and deploy it on the heat equation. We take an "eat dessert first" approach (inspired by Tim Hsu's *Fourier Series, Fourier Transforms, and Function Spaces: A Second Course in Analysis*). Specifically, here is our strategy.

1. We define the Fourier transform for continuous, absolutely integrable functions. Eventually we will relax the continuity requirement to piecewise continuity.

2. We apply the Fourier transform to the heat equation.

3. ???.

4. We get a *candidate* solution formula for the heat equation.

5. We check that this candidate is actually a solution (i.e., by doing calculus).

6. We use this solution to learn other interesting things about the heat equation.

7. We study other aspects of the heat equation that we will not be able to understand with the Fourier transform. In particular, we develop the machinery to prove a uniqueness theorem for solutions to the heat equation. (We got that more or less for free along with existence from D'Alembert's formula for the wave equation. The heat equation is harder.)

8. We fill in a variety of gaps in our understanding of the Fourier transform so that we can study other problems with it more rigorously.

9. We study other problems with it more rigorously.

Example 21.10 assures us that the following definition makes sense. (Does it?)

22.1 Definition. *Let $f\colon \mathbb{R} \to \mathbb{C}$ be continuous with $|f|$ integrable (Definition 21.7). The*

Fourier transform *of* f *at* $k \in \mathbb{R}$ *is*

$$\widehat{f}(k) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-ikx} \, dx.$$

We sometimes write $\mathfrak{F}[f](k) = \widehat{f}(k)$.

The factor of $1/\sqrt{2\pi}$ is a bit of a "fudge factor" that makes some calculations and identities later easier and more transparent, at the cost of making others harder and more opaque. Life is a series of compromises.

Previously we have said that integrals *extract* useful data about functions and also *represent* functions. We have not seen all that much extraction of useful data, but it turns out that the **Fourier modes** $\widehat{f}(k)$ will tell us a variety of useful facts about f. The Fourier transform also "represents" f in the following sense. Here, for the first of many times, we will use the weasel word "nice" to refer to a property of functions that we will fill in later in our subsequent, more rigorous treatment of Fourier transforms.

22.2 Untheorem. *Let* $f\colon \mathbb{R} \to \mathbb{C}$ *be "nice." Then*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)e^{ikx} \, dk.$$

That is, for suitable f, we can recover f from its Fourier transform.

Since this is a course in differential equations, we should wonder how the Fourier transform interacts with the derivative. Quite nicely, thank you for asking.

If f is differentiable, and if both f and f' are "nice," then we should be able to represent f' (not just f) via its Fourier transform:

$$f'(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f'}(k)e^{ikx} \, dk.$$

But we should also be able to calculate f' from the Fourier representation of f and differentiation under the integral:

$$f'(x) = \partial_x \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)e^{ikx} \, dk \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \partial_x[\widehat{f}(k)e^{ikx}] \, dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ik\widehat{f}(k)e^{ikx} \, dk.$$

Equating these two putative representations of f' and doing a little algebra, we find

$$\int_{-\infty}^{\infty} [\widehat{f'}(k) - ik\widehat{f}(k)]e^{ikx} \, dk = 0$$

for all $x \in \mathbb{R}$.

Now here is a "nice" property of Fourier integrals. We should think of the transform as an "instrument" that we apply to functions, and the results we get are those Fourier modes. If the results are always 0, the input should always be 0.

22.3 Untheorem. *Let $g\colon \mathbb{R} \to \mathbb{C}$ be "nice" and suppose that*

$$\int_{-\infty}^{\infty} g(k)e^{ikx}\, dk = 0$$

for all $x \in \mathbb{R}$. Then $g(k) = 0$ for all $k \in \mathbb{R}$.

It follows that

$$\widehat{f'}(k) = ik\widehat{f}(k).$$

This is immensely important: under the lens of the Fourier transform, differentiation becomes "multiply by ik." We might say

$$\widehat{\partial_x[\cdot]} = ik \times \widehat{(\cdot)}.$$

We can extend this to the second derivative (and higher derivatives) for "nice" functions:

$$\widehat{f''}(k) = \widehat{(f')'}(k) = ik\widehat{f'}(k) = (ik)^2\widehat{f}(k) = -k^2\widehat{f}(k).$$

This is all that we need to know about the Fourier transform to apply it with abandon to the heat equation. Suppose that u solves

$$\begin{cases} u_t = u_{xx}, & -\infty < x < \infty,\ t > 0 \\ u(x,0) = f(x), & -\infty < x < \infty \end{cases}$$

and that u and f are "nice." We apply the Fourier transform to u "spatially" or "in the x-variable." Consequently, "nice" should mean, at least, that $u(\cdot, t)$ is integrable for each $t > 0$ (where $u(\cdot, t)$ is the map $x \mapsto u(x,t)$) and also that f is integrable.

Put

$$\widehat{u}(k,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} u(x,t)e^{-ikx}\, dx.$$

We should think of t as just a parameter in the integrand; all of the action is happening with x. Then

$$\widehat{u_{xx}}(k,t) = -k^2\widehat{u}(k,t).$$

In the time variable, we recognize differentiation under the integral:

$$\begin{aligned}\widehat{u_t}(k,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} u_t(x,t)e^{-ikx}\, dx &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \partial_t[u(x,t)e^{-ikx}]\, dx \\ &= \partial_t\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} u(x,t)e^{-ikx}\, dx\right] = \partial_t[\widehat{u}](k,t).\end{aligned}$$

To avoid confusion, we will not write this as $\widehat{u}_t(k,t)$. All together, we expect that a "nice" solution u to the heat equation with "nice" initial data f will satisfy

$$\begin{cases} \partial_t[\widehat{u}](k,t) = -k^2\widehat{u}(k,t) \\ \widehat{u}(k,0) = \widehat{f}(k). \end{cases}$$

This is really a family of IVP at the ODE level parametrized in $k \in \mathbb{R}$. (We posed the heat equation only for $t > 0$, but we can solve this IVP for all t, so we might as well consider all t here.) The notation may be burdensome, but all this is asking us to do is solve

$$\begin{cases} y' = -k^2 y \\ y(0) = \widehat{f}(k) \end{cases}$$

for each $k \in \mathbb{R}$. Certainly we know how to do that: $y(t) = \widehat{f}(k)e^{-k^2 t}$. And so $\widehat{u}$ should satisfy

$$\widehat{u}(k,t) = \widehat{f}(k)e^{-k^2 t}.$$

Now we can recover u from $\widehat{u}$ by Untheorem 22.2:

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{u}(k,t)e^{ikx}\,dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)e^{-k^2 t}e^{ikx}\,dk. \tag{22.1}$$

This may well be a valid candidate for a solution formula!

22.4 Problem. **(i)** Fix $t > 0$ and $x \in \mathbb{R}$ and define $g(k) := \widehat{f}(k)e^{-k^2 t}e^{ikx}$. Show that if $\widehat{f}$ is integrable or bounded (bounded meaning the existence of $M > 0$ such that $|\widehat{f}(k)| \le M$ for all k), then g is integrable, and so the integral on the right in (22.1) converges. (It will turn out that if f is integrable, then $\widehat{f}$ is always bounded, although not necessarily integrable.)

(ii) Assume that we may differentiate under the integral on the right in (22.1) with respect to x and t as much as we want for $x \in \mathbb{R}$ and $t > 0$. Show that u as defined by (22.1) satisfies $u_t = u_{xx}$.

(iii) Show that u as defined by (22.1) meets $u(x,0) = f(x)$. [Hint: *Untheorem 22.2.*]

22.5 Problem. Repeat the work above for the transport IVP

$$\begin{cases} u_t + u_x = 0, \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty \end{cases}$$

and recover the expected, beloved formula $u(x,t) = f(x-t)$. [Hint: *apply the Fourier transform to u in x and get an ODE-type IVP for $\widehat{u}$. Solve it. Then recover u from its Fourier transform via Untheorem 22.2. Do some algebra in the integrand and recognize the integral as the Fourier transform of f.*]

Now we begin the laborious process of verifying that (22.1) actually gives a formula for a solution to the heat equation. Problem 22.4 ensures that, if $|f|$ is integrable, then the formula actually converges to a real number for each $x \in \mathbb{R}$ and $t > 0$. (What goes wrong if $t \le 0$? This is one mathematical reason to take $t > 0$ in our statement of the heat equation—we do it because it leads to a problem that we can solve!)

We first replace $\widehat{f}(k)$ in (22.1) by its integral definition and find

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y)e^{-iky} \ dy \right) e^{-k^2 t} e^{ikx} \ dk.$$

Here we are writing the variable of integration in the definition of $\widehat{f}(k)$ as y so as not to overwork x. This cleans up slightly to

$$u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y)e^{-iky}e^{-k^2 t}e^{ikx} \ dy \ dk. \tag{22.2}$$

We might note that the factor of $f(y)$ is the only factor in the integrand that does not depend on k. If we interchange the order of integration (a dicey move—is Fubini's theorem valid for double improper integrals?), then we could probably pull it out of one integral:

$$u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y)e^{-iky}e^{-k^2 t}e^{ikx} \ dk \ dy = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) \left(\int_{-\infty}^{\infty} e^{-iky}e^{-k^2 t}e^{ikx} \ dk \right) \ dy.$$

We focus on the integral in parentheses. Collect the complex exponentials into one:

$$\int_{-\infty}^{\infty} e^{-iky}e^{-k^2 t}e^{ikx} \ dk = \int_{-\infty}^{\infty} e^{-k^2 t}e^{ik(x-y)} \ dk.$$

Pull in that factor of $1/2\pi$ and define

$$H(s,t) := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-k^2 t}e^{iks} \ dk.$$

Problem 21.11 and the comparison test ensure that this integral converges. Then our solution candidate should be

$$u(x,t) = \int_{-\infty}^{\infty} H(x-y,t)f(y) \ dy. \tag{22.3}$$

Now we need to check that *this* integral converges and that it is sufficiently differentiable in x and t. Doing so will require a much deeper understanding of H, which turns out to be quite a nice function.

Day 23: Friday, October 4.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Example 6 on pp. 425–426 computes the Fourier transform of the Gaussian. Example 1 on pp. 124–125 discusses the heat kernel, and p. 461 shows how the heat kernel satisfies the heat equation itself.

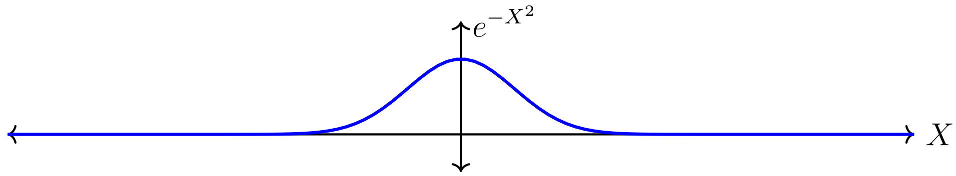
We start by cleverly rewriting H:

$$H(s,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-k^2 t}e^{iks} \ dk = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(k\sqrt{t})^2}e^{-i(-s)k} \ dk \right).$$

While this may not have been the obvious move, it shows that $H(s,t)$ is basically a Fourier transform (with the unusual notational choice of using s for the Fourier variable but k for the variable of integration). Specifically, put

$$\mathcal{G}(X) := e^{-X^2}.$$

This is a "Gaussian"-type function, and one of its chief virtues is that it decays *extremely* fast as $X \to \pm\infty$.



Now let $\mathcal{G}(\sqrt{t}\cdot)$ be the function

$$\mathcal{G}(\sqrt{t}\cdot)\colon \mathbb{R} \to \mathbb{R}\colon k \mapsto e^{-(k\sqrt{t})^2}.$$

Then

$$H(s,t) = \frac{1}{\sqrt{2\pi}}\widehat{\mathcal{G}(\sqrt{t}\cdot)}(-s). \tag{23.1}$$

Problem 22.4 ensures that this Fourier transform really is defined. So what is it?

The form of this transform first motivates us to think about transforms of "scaled" functions. Let $g\colon \mathbb{R} \to \mathbb{C}$ be continuous with $|g|$ integrable, and let $\alpha \in \mathbb{R}$. Denote by $g(\alpha\cdot)$ the map

$$g(\alpha\cdot)\colon \mathbb{R} \to \mathbb{C}\colon x \mapsto g(\alpha x).$$

23.1 Problem. Explain why $|g(\alpha\cdot)|$ is integrable.

Then, by definition,

$$\widehat{g(\alpha\cdot)}k = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(\alpha x)e^{-ikx}\,dx.$$

How can we relate $\widehat{g(\alpha\cdot)}$ to $\widehat{g}$? One idea is to make just g show up in the integrand. Substitute $u = \alpha x$ to find, formally,

$$\int_{-\infty}^{\infty} g(\alpha x)e^{-ikx}\,dx = \frac{1}{\alpha}\int_{\alpha\cdot(-\infty)}^{\alpha\cdot\infty} g(u)e^{-i(k/\alpha)u}\,du.$$

If $\alpha > 0$, we should then expect

$$\widehat{g(\alpha\cdot)}k = \frac{1}{\alpha}\widehat{g}\left(\frac{k}{\alpha}\right). \tag{23.2}$$

23.2 Problem (Nonzero scaling preserves integrability). Clean this up using the following more general approach. Let $h\colon \mathbb{R} \to \mathbb{C}$ be continuous and integrable and let $\alpha \in \mathbb{R}\setminus\{0\}$. Prove that

$$\int_{-\infty}^{\infty} h(\alpha x)\ dx = \frac{1}{|\alpha|}\int_{-\infty}^{\infty} h(s)\ ds.$$

What does this say about $\widehat{g(\alpha\cdot)}$ for $\alpha \neq 0$ and $|g|$ integrable? [Hint: *study the integrals* $\int_a^0 h(\alpha x)\ dx$ *and* $\int_0^b h(\alpha x)\ dx$. *Change variables and pay attention to how the sign of* α *affects the limits of integration.*]

23.3 Problem (Horizontal translation preserves integrability). Let $g\colon \mathbb{R} \to \mathbb{C}$ be continuous with $|g|$ integrable. Let $d \in \mathbb{R}$. Prove that the "shifted map"

$$S^d g\colon \mathbb{R} \to \mathbb{C}\colon x \mapsto g(x+d)$$

is integrable with

$$\int_{-\infty}^{\infty} g(x+d)\ dx = \int_{-\infty}^{\infty} g(u)\ du \quad \text{and} \quad \widehat{S^d g}(k) = e^{ikd}\widehat{g}(k).$$

[Hint: *for integrability, it may be easier to prove that the limits in Definition 21.7 exist and then use Problem 21.12 to express* $\int_{-\infty}^{\infty} S^d f = \lim_{R\to\infty}\int_{-R}^{R} S^d f$.]

We combine (23.1) and (23.2) to obtain

$$H(s,t) = \frac{1}{\sqrt{2\pi}}\widehat{\mathcal{G}(\sqrt{t}\cdot)}(-s) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{t}}\right)\widehat{\mathcal{G}}\left(-\frac{s}{\sqrt{t}}\right). \tag{23.3}$$

So, what is $\widehat{\mathcal{G}}$?

By definition, it is

$$\widehat{\mathcal{G}}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2}e^{-ikx}\ dx.$$

This is one of those times when brute force is not, in fact, the best force, and we need some tricks to evaluate this integral. We start by thinking about what $\mathcal{G}$ does (possibly from our separable ODE days):

$$\mathcal{G}'(x) = -2xe^{-x^2}(= -2x\mathcal{G}(x)).$$

We also expect

$$\widehat{\mathcal{G}'}(k) = ik\widehat{\mathcal{G}}(k).$$

Then

$$ik\widehat{\mathcal{G}}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} -2xe^{-x^2}e^{-ikx}\ dx.$$

We work quite a bit at the integral on the right:

$$\int_{-\infty}^{\infty} -2xe^{-x^2}e^{-ikx}\ dx = \frac{2}{i}\int_{-\infty}^{\infty} -ixe^{-x^2}e^{-ikx}\ dx = \frac{2}{i}\int_{-\infty}^{\infty} \partial_k[e^{-x^2}e^{-ikx}]\ dx.$$

If we can interchange the derivative and integral, then

$$ik\widehat{\mathcal{G}}(k) = \frac{2}{i}\partial_k\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2}e^{-ikx}\ dx\right] = \frac{2}{i}\partial_k[\widehat{\mathcal{G}}](k). \tag{23.4}$$

Be careful with notation: we are writing $\mathcal{G}' = \partial_x[\mathcal{G}]$ for the "ordinary" derivative of $\mathcal{G}$ but $\partial_k[\widehat{\mathcal{G}}]$ for the derivative of the Fourier transform of $\mathcal{G}$. (Strictly speaking, we have not proved that the transform is differentiable, because we did not justify interchanging the improper integral and ∂_k above. We will.) Now rejoice at (23.4): this is really an ODE for $\widehat{\mathcal{G}}$, and it reads

$$\partial_k[\widehat{\mathcal{G}}](k) = -\frac{k}{2}\widehat{\mathcal{G}}(k).$$

Perhaps it would look better as

$$y'(t) = -\frac{t}{2}y(t)?$$

Then $y(t) = y(0)e^{-t^2/4}$, and so

$$\widehat{\mathcal{G}}(k) = \widehat{\mathcal{G}}(0)e^{-k^2/4}.$$

So what is $\widehat{\mathcal{G}}(0)$? By definition,

$$\widehat{\mathcal{G}}(0) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2}\ dx,$$

and it turns out that

$$\int_{-\infty}^{\infty} e^{-x^2}\ dx = \sqrt{\pi}, \tag{23.5}$$

an identity that we will not prove here. Thus

$$\widehat{\mathcal{G}}(k) = \frac{\sqrt{\pi}}{\sqrt{2\pi}}e^{-k^2/4} = \frac{e^{-k^2/4}}{\sqrt{2}}. \tag{23.6}$$

All together, we conclude from (23.3) and (23.6) that

$$H(s,t) = \frac{1}{\sqrt{2\pi t}}\widehat{\mathcal{G}}\left(-\frac{s}{\sqrt{t}}\right) = \frac{1}{\sqrt{2\pi t}}\left(\frac{1}{\sqrt{2}}\right)e^{-(-s/\sqrt{t})^2/4} = \frac{1}{\sqrt{4\pi t}}e^{-s^2/4t}.$$

We call this function H the **HEAT KERNEL**.

This is a very nice expression for H—no more integrals! Be aware that $s \in \mathbb{R}$ can be arbitrary, but we need $t > 0$. Also, it is traditional to leave the 4 inside the square root.

23.4 Problem. Check that H satisfies the heat equation in the sense that $H_t = H_{ss}$.

At last we return to our candidate solution (22.3) for the heat equation:

$$u(x,t) = \int_{-\infty}^{\infty} H(x-y,t)f(y)\ dy = \int_{-\infty}^{\infty} \frac{e^{-(x-y)^2/4t}}{\sqrt{4\pi t}}f(y)\ dy,\ x \in \mathbb{R},\ t > 0. \tag{23.7}$$

With this explicit formula for H, we can prove the convergence of this integral with *two* different hypotheses on f.

23.5 Problem. Prove that the (second) integral in (23.7) converges in each of the following cases, assuming $x \in \mathbb{R}$ and $t > 0$. [Hint: *use the comparison test—the function that you "compare the integrand to" will be different in each case.*]

(i) f is bounded in the sense that there exists $M > 0$ such that $|f(y)| \le M$ for all $y \in \mathbb{R}$.

(ii) $|f|$ is integrable.

This assures us that the function u in (23.7) is defined: $u(x,t) \in \mathbb{R}$ for all $x \in \mathbb{R}$ and $t > 0$. Is u differentiable, and does u satisfy the heat equation? If differentiation under the improper integral is justified (something we really do need to think about), then Problem 23.4 implies

$$u_t(x,t) = \partial_t\left[\int_{-\infty}^{\infty} H(x-y,t)f(y)\,dy\right] = \int_{-\infty}^{\infty} \partial_t[H(x-y,t)f(y)]\,dy = \int_{-\infty}^{\infty} H_{ss}(x-y,t)f(y)\,dy$$
$$= \int_{-\infty}^{\infty} \partial_x^2[H(x-y,t)f(y)]\,dy = \partial_x^2\left[\int_{-\infty}^{\infty} H(x-y,t)f(y)\,dy\right] = u_{xx}(x,t).$$

Day 24: Monday, October 7.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Example 1 on p. 459 is Tychonov's example for nonuniqueness in the heat equation. Pages 462–463 outline the ϵ–δ-style argument that $\lim_{t\to0^+}\int_{-\infty}^{\infty} H(x-y,t)f(y)\,dy = f(x)$. This is the behavior that one expects with a delta function; see p. 471. Another "bounded in finite time" uniqueness result is Theorem 2 on p. 465.

Now that we have a solution to the heat equation, we start exploiting its properties. First, the formula

$$u(x,t) = \int_{-\infty}^{\infty} H(x-y,t)f(y)\,dy$$

is not defined at $t = 0$, but we want $u(x,0) = f(x)$. One can show that if f is continuous, then

$$\lim_{t\to0^+}\int_{-\infty}^{\infty} H(x-y,t)f(y)\,dy = f(x). \tag{24.1}$$

This is mostly a classical ϵ–δ argument that is a little too technical for us to do here but that does not require all that many fancy tools. Thus if we put instead

$$u(x,t) = \begin{cases} \int_{-\infty}^{\infty} H(x-y,t)f(y)\,dy, & x \in \mathbb{R},\ t > 0 \\ f(x), & x \in \mathbb{R}, t = 0, \end{cases} \tag{24.2}$$

then u is continuous on $\mathcal{D}_0 := \{(x,t) \in \mathbb{R}^2 \mid x \in \mathbb{R},\ t \ge 0\}$, and u solves $u_t = u_{xx}$ on $\mathcal{D} := \{(x,t) \in \mathbb{R}^2 \mid x \in \mathbb{R},\ t > 0\}$. All in all, this is a bit weaker than what we found for the transport and wave equations, where the solution formulas met the initial conditions immediately, without any extra work.

Next, what about uniqueness? We got that for free for transport and wave just from those happy solution formulas. Is there only one solution to

$$\begin{cases} u_t = u_{xx}, \ -\infty < x < \infty, \ t > 0 \\ u(x,0) = f(x), \ -\infty < x < \infty? \end{cases}$$

Remarkably, no! Even taking $f = 0$ does not force uniqueness.

24.1 Theorem (Tychonov). *There exists a function u that is continuous on $\mathcal{D}_0$ (as defined above) and that solves the heat equation $u_t = u_{xx}$ on $\mathcal{D}$ (as defined above) with $u(x,0) = 0$. However, u is not identically zero.*

With some effort, one can show that for $t > 0$, Tychonov's solution is not bounded in x by any exponentially growing function. Physically this unboundedness is wholly unrealistic: how could the temperature in a rod soar to ∞ without help from any extra heat source? (There is no heat source in the heat IVP beyond the initial temperature.)

So what does force uniqueness? It turns out that if we build a hypothesis onto the *solution's* behavior, not just the initial temperature, then we get uniqueness.

24.2 Theorem. *Suppose that u solves the heat equation*

$$\begin{cases} u_t = u_{xx}, \ -\infty < x < \infty, \ t > 0 \\ u(x,0) = 0, \ -\infty < x < \infty \end{cases}$$

and is continuous at $t = 0$. Suppose also that u is **BOUNDED IN FINITE TIME** *in the sense that for all $T > 0$, there is $M_T > 0$ such that $|u(x,t)| \le M_T$ for all $x \in \mathbb{R}$ and $t \in [0,T]$. Then $u(x,t) = 0$ for all $x \in \mathbb{R}$ and $t \ge 0$.*

We will prove this theorem later after we develop some new tools. Here is how it helps.

24.3 Problem. Suppose that u and v solve the same heat equation

$$\begin{cases} u_t = u_{xx}, \ -\infty < x < \infty, \ t > 0 \\ u(x,0) = f(x), \ -\infty < x < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_t = v_{xx}, \ -\infty < x < \infty, \ t > 0 \\ v(x,0) = f(x), \ -\infty < x < \infty \end{cases}$$

and are both bounded in finite time. Prove that $u = v$. [Hint: *consider $w = u - v$, show that w is bounded in finite time (the triangle inequality), and apply Theorem 24.2.*]

We might then ask what conditions on f guarantee that the solution to the heat IVP as given by (24.2) guarantee that u is bounded in finite time. Given $T > 0$, we want $M_T > 0$ such that $|u(x,t)| \le M_T$ for all $x \in \mathbb{R}$ and $t \in [0,T]$. Taking, say, $T = 1$ and $t = 0$, this shows that we want $M_1 > 0$ such that $|u(x,0)| \le M_1$ for all $x \in \mathbb{R}$. That means we want $|f(x)| \le M_1$ for all x. And so the initial temperature distribution must be bounded.

Is that enough? We drop the subscript and assume $|f(x)| \leq M$ for all $x \in \mathbb{R}$ and some $M > 0$. The previous paragraph shows that with u defined by (24.2), we have $|u(x,0)| \leq M$ for all $x \in \mathbb{R}$. Now we estimate $u(x,t)$ for $t > 0$. The triangle inequality for integrals implies

$$|u(x,t)| \leq \frac{M}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4t} \, dy = \frac{M}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-[(x-y)/2\sqrt{t}]^2} \, dy.$$

24.4 Problem. Substitute $s = -(x-y)/2\sqrt{t}$ in the integral on the right and use (23.5) to derive a bound for u that is independent of t.

Now we have an existence *and* uniqueness result for the heat equation (pending some housekeeping with unproven results).

24.5 Theorem. *Let $f \in \mathcal{C}(\mathbb{R})$ be bounded in the sense that there exists $M > 0$ such that $|f(x)| \leq M$ for all $x \in \mathbb{R}$. Then the only solution to the heat IVP*

$$\begin{cases} u_t = u_{xx}, \ -\infty < x < \infty, \ t > 0 \\ u(x,0) = f(x), \ -\infty < x < \infty \end{cases} \tag{24.3}$$

is

$$u(x,t) = \begin{cases} \int_{-\infty}^{\infty} H(x-y,t) f(y) \, dy, \ x \in \mathbb{R}, \ t > 0 \\ f(x), \ x \in \mathbb{R}, t = 0, \end{cases} \qquad H(s,t) = \frac{e^{-s^2/4t}}{\sqrt{4\pi t}}.$$

So what else is this solution *doing*? First, no matter what the initial temperature distribution is, eventually everything "cools all the way down."

24.6 Problem. Let $f \in \mathcal{C}(\mathbb{R})$ be bounded and let $|f|$ be integrable. Let u solve (24.3). Prove that

$$\lim_{t \to \infty} u(x,t) = 0$$

for each $x \in \mathbb{R}$. Go further and explain how this limit is "uniform" in x by finding a bound $|u(x,t)| \leq M(t)$ valid for all $x \in \mathbb{R}$ and $t > 0$ with $\lim_{t \to \infty} M(t) = 0$.

This is physically reasonable, yes? Now we prove a physically baffling result. Suppose that $f(x) \geq 0$ for all $x \in \mathbb{R}$ but $f(x_0) > 0$ for some $x_0 \in \mathbb{R}$. Assuming, as usual, that f is continuous, there is $\delta > 0$ such that $f(x) > 0$ for $x_0 - \delta < x < x_0 + \delta$. Then, for $t > 0$,

$$u(x,t) = \int_{-\infty}^{\infty} H(x-y,t) f(y) \, dy \geq \int_{x_0-\delta}^{x_0+\delta} H(x-y,t) f(y) \, dy > 0.$$

The second, strict inequality is just monotonicity of the integral. The first, nonstrict inequality is a consequence of the following (and the nonnegativity of H and f).

24.7 Problem. Let a, b, c, $d \in \mathbb{R}$ with $a < c < d \le b$. Suppose that $f \in \mathcal{C}([a,b])$ with $f(x) \ge 0$ for $x \in [a,b]$. Prove that

$$\int_c^d f \le \int_a^b f. \tag{24.4}$$

If $f(x) > 0$ for $x \in [a,b]$, show that the nonstrict inequality in (24.4) becomes strict. [Hint: *start with a picture and check that* $\int_c^d f = \int_a^b f - \int_a^c f - \int_d^b f$.]

Here is what we have shown: if the initial temperature distribution f is nonnegative on $\mathbb{R}$ but positive at some point x_0 (and maybe 0 elsewhere on $\mathbb{R}$), then u is positive for all x and all $t > 0$. Informally, "if f is positive somewhere, then u is positive everywhere." If we think about the heat equation as modeling the temperature of an infinite rod, then the initial heat contribution from f gets "transported instantly" to all of the rod, even if that initial heat contribution is localized over a small spatial interval, like $(x_0 - \delta, x_0 + \delta)$. This is an "infinite propagation speed" result for the heat equation, and it stands in marked contrast to the "finite propagation speeds" for the transport and wave equations.

24.8 Problem. Let $f \in \mathcal{C}(\mathbb{R})$ be bounded and nonnegative and suppose that u solves (24.3). Prove that if $u(x_0, t_0) = 0$ for some $x_0 \in \mathbb{R}$ and $t_0 > 0$, then $u(x,t) = 0$ for all $x \in \mathbb{R}$ and $t > 0$. [Hint: *contrapositive, quantifiers.*]

24.9 Problem. Prove the following "comparison" principle for the heat equation. Suppose that f_1, $f_2 \in \mathcal{C}(\mathbb{R})$ are bounded with $f_1(x) \le f_2(x)$ for all x and u and v solve

$$\begin{cases} u_t = u_{xx}, & x \in \mathbb{R},\ t > 0 \\ u(x,0) = f_1(x), & x \in \mathbb{R} \end{cases} \qquad \text{and} \qquad \begin{cases} v_t = v_{xx}, & x \in \mathbb{R},\ t > 0 \\ v(x,0) = f_2(x), & x \in \mathbb{R}. \end{cases}$$

Prove that either $u(x,t) = v(x,t)$ for all $x \in \mathbb{R}$ and $t > 0$ or that $u(x,t) < v(x,t)$ for all $x \in \mathbb{R}$ and $t > 0$.

Day 25: Wednesday, October 9.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Theorem 2 on p. 143 states the maximum principle; see also the minimum principle on the following page and Example 2 on pp. 145–146. The maximum principle is proved on pp. 148–149.

We turn our attention to proving Theorem 24.2. Our main tool in this proof is a result worthwhile in and of itself: a "maximum principle." Here we no longer work with the heat equation on the whole real line, just on a finite spatial subinterval.

25.1 Theorem (Maximum principle). *Suppose that u solves*

$$u_t = u_{xx}$$

for $a \le x \le b$ and $0 < t \le T$ and that u is continuous for $a \le x \le b$ and $0 \le t \le T$. Let

$$\mathcal{D} := \left\{(x,t) \in \mathbb{R}^2 \;\middle|\; a \le x \le b,\ 0 \le t \le T\right\}$$

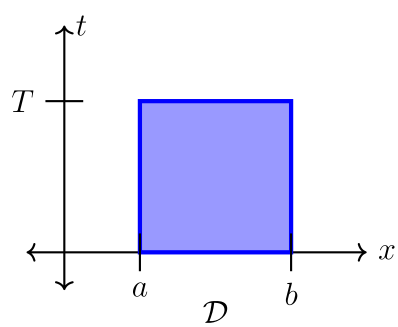
and

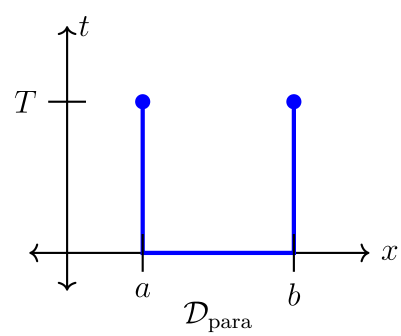
$$\mathcal{D}_{\text{para}} := \left\{(a,t) \in \mathbb{R}^2 \;\middle|\; 0 \le t \le T\right\} \cup \left\{(x,0) \in \mathbb{R}^2 \;\middle|\; a \le x \le b\right\} \cup \left\{(b,t) \in \mathbb{R}^2 \;\middle|\; 0 \le t \le T\right\}.$$

Then

$$\max_{(x,t)\in\mathcal{D}} u(x,t) = \max_{(x,t)\in\mathcal{D}_{\text{para}}} u(x,t).$$

Proof. First we draw some pictures: the set $\mathcal{D}$ and its "parabolic boundary" $\mathcal{D}_{\text{para}}$.





Since u is continuous on the closed, bounded set $\mathcal{D}$, the extreme value theorem implies that u attains its maximum somewhere on $\mathcal{D}$: there is $(x_0, t_0) \in \mathcal{D}$ such that

$$M := u(x_0, t_0) = \max_{(x,t)\in\mathcal{D}} u(x,t).$$

Likewise, since $\mathcal{D}_{\text{para}}$ is also closed and bounded, u attains a maximum somewhere on $\mathcal{D}_{\text{para}}$, and, since $\mathcal{D}_{\text{para}} \subseteq \mathcal{D}$, we have

$$m := \max_{(x,t)\in\mathcal{D}_{\text{para}}} u(x,t) \le M.$$

Our goal is to show $m = M$; we proceed by contradiction and assume $m < M$.

We need some results from single-variable calculus, which we briefly review here (and discuss further in Problem 25.4 below). Suppose $f \in \mathcal{C}^2([c,d])$, so f achieves its maximum at some $x_\star \in [c,d]$. If $x_\star \in (c,d)$, then $f'(x_\star) = 0$ and $f''(x_\star) \le 0$. If, however, all we know is that $x_\star \in [c,d]$, i.e., $x_\star$ is one of the endpoints, then all we know is that $f'(x_\star) \ge 0$.

Here is how these results motivate the maximum principle. If $(x_0, t_0) \in \mathcal{D} \setminus \mathcal{D}_{\text{para}}$, then $a < x_0 < b$ and $0 < t \le T$. We have $u_t(x_0, t_0) \ge 0$, since the function $u(x_0, \cdot)$ achieves its maximum on $[0, T]$ at $t_0 \in (0, T]$ and $u_{xx}(x_0, t_0) \le 0$, since the function $u(\cdot, t_0)$ achieves its maximum on $[a, b]$ at $x_0 \in (a, b)$. But we also know $u_t(x_0, t_0) - u_{xx}(x_0, t_0) = 0$. If we knew more, like $u_t(x_0, t_0) > 0$, or $u_{xx}(x_0, t_0) < 0$, then we would get a contradiction. If we could rule out $t_0 = T$, then we would have $0 < t_0 < T$, and so $u_t(x_0, t_0) > 0$. That would be enough for the contradiction. Or, if we could ensure $u_{xx}(x_0, T) < 0$, that would also get us the contradiction. However, we do not have enough information to do any of that.

Instead, the trick of the proof is to modify u into a new function v, which has a slightly more tractable second derivative in x. Probably the simplest function that has a nontrivial second derivative is $x \mapsto x^2$, or a multiple thereof. So, we put

$$v(x, t) := u(x, t) + \epsilon x^2,$$

where we will specify $\epsilon > 0$ shortly.

Now we think about extreme values. First, if $(x, t) \in \mathcal{D}_{\text{para}}$, then $x^2 \le \max\{a^2, b^2\}$. (This is sensitive if $a < 0$ or $b < 0$.) So,

$$\max_{(x,t)\in\mathcal{D}_{\text{para}}} v(x, t) \le m + \epsilon \max\{a^2, b^2\}. \tag{25.1}$$

Since we are assuming $m < M$, if we take $\epsilon > 0$ small enough relative to m, M, a^2, and b^2, then

$$m + \epsilon \max\{a^2, b^2\} < M. \tag{25.2}$$

And

$$v(x_0, t_0) = M + \epsilon x_0^2 \ge M.$$

(We have to keep the nonstrict inequality just in case $x_0 = 0$.) So, since $(x_0, t_0) \in \mathcal{D}$,

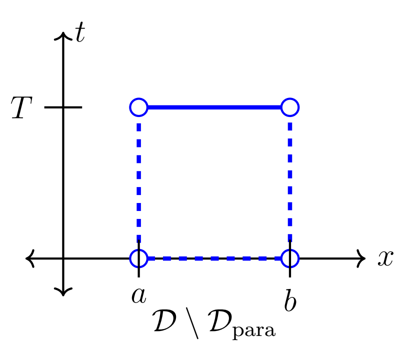
$$\max_{(x,t)\in\mathcal{D}} v(x, t) \ge M. \tag{25.3}$$

By the way, this maximum exists by the extreme value theorem, just as it did for u.

We combine (25.1), (25.2), and (25.3) to conclude

$$\max_{(x,t)\in\mathcal{D}_{\text{para}}} v(x, t) < \max_{(x,t)\in\mathcal{D}} v(x, t) =: v(x_1, t_1)$$

for some $(x_1, t_1) \in \mathcal{D} \setminus \mathcal{D}_{\text{para}}$. Here is a sketch of $\mathcal{D} \setminus \mathcal{D}_{\text{para}}$.



We then have $a < x_1 < b$ and $0 < t_1 \le T$. By the reasoning above, $v_{xx}(x_1, t_1) \le 0$ and $v_t(x_1, t_1) \ge 0$. Thus

$$v_t(x_1, t_1) \ge v_{xx}(x_1, t_1). \tag{25.4}$$

But we know more about v: for $a \le x \le b$ and $0 < t \le T$,

$$v_t(x,t) = u_t(x,t) \quad \text{and} \quad v_{xx}(x,t) = u_{xx}(x,t) + 2\epsilon.$$

Then the inequality (25.4) reads

$$u_t(x_1,t_1) \ge u_{xx}(x_1,t_1) + 2\epsilon,$$

and that simplifies to $2\epsilon \le 0$, a contradiction. ∎

25.2 Problem. Reread the preceding proof and convince yourself that at no point was differentiability of u at $t = 0$ used. (This is important, because we really cannot assume that u is differentiable at $t = 0$ in practice!)

25.3 Problem. Suppose that the heat equation is modeling the temperature distribution of a finite rod with endpoints at $x = a$ and $x = b$. Explain how the maximum principle implies that the maximum temperature that the rod reaches between times 0 and T occurs either at the endpoints at some point in time from times 0 to T or somewhere within the rod but only at time 0.

25.4 Problem. In calculus, we usually apply derivative tests for extreme values occurring at interior points of intervals, so here is a chance to think about what happens at the endpoints. Let $f \in \mathcal{C}^2([a,b])$ and suppose that

$$f(a) = \max_{a \le x \le b} f(x).$$

(i) Use the definition of the derivative to prove that $f'(a) \le 0$.

(ii) Give an example (start by drawing a picture) to show that we may have $f'(a) < 0$, in contrast to our likely calculus intuition that $f'(a) = 0$.

(iii) Give examples (start, again, by drawing pictures) to show that any of the possibilities $f''(a) > 0$, $f''(a) = 0$, or $f''(a) < 0$ are possible. When drawing, remember that $f'(a) \le 0$.

25.5 Problem. Let u satisfy the hypotheses of the maximum principle. By considering $v := -u$, prove that u also achieves its minimum on the parabolic boundary.

Now we start to prove Theorem 24.2. The idea is that if u solves the heat equation for all $x \in \mathbb{R}$ and $t > 0$ and is continuous at $t = 0$ and bounded in finite time and if $u(x,0) = 0$, then $u = 0$ (which is what we expect with zero initial conditions). Let $T > 0$; we show that $u(x,t) = 0$ for $x \in \mathbb{R}$ and $0 < t \le T$. Since $T > 0$ is arbitrary, this shows $u(x,t) = 0$ for all $t > 0$ and $x \in \mathbb{R}$. (The case $t = 0$ is the initial condition, so we ignore that.) Fix $x_0 \in \mathbb{R}$ and $t_0 \in (0,T]$. Our goal is now to show $|u(x_0,t_0)| \le \epsilon$ for all $\epsilon > 0$; then we will have $u(x_0,t_0) = 0$. We will do this by introducing an comparator function (which depends on ϵ, x_0, and t_0) and applying the maximum principle to that function; it will turn out to be "easy" to compute maxima on the parabolic boundary of its domain, and those maxima will force the inequalities $-\epsilon \le u(x_0, t_0) \le \epsilon$.

The question is now what the right "comparator function" is and what is the right finite spatial interval on which to apply the maximum principle (be aware that now u is defined for $x \in \mathbb{R}$).

Day 26: Friday, October 11.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Theorem 1 on p. 140 proves uniqueness for the finite rod heat equation.

We finish the proof of Theorem 24.2. Here is the strategy.

1. We are going to construct a "comparator" function v such that $v(x,t) \le 0$ for all $x \in \mathbb{R}$ and $t \in [0,T]$ with $v(x_0,t_0) = u(x_0,t_0) - \epsilon$. This will yield $u(x_0,t_0) \le \epsilon$, and a similar construction will give the lower bound. We will achieve this inequality on v by finding that v solves the heat equation and applying the maximum principle.

2. One way to build v is to exploit the linearity of the heat equation. We will find a solution w of the heat equation ($w_t = w_{xx}$) and normalize it so $w(x_0,t_0) = 1$.

3. Then we will put $v(x,t) := u(x,t) - \epsilon w(x,t)$ and get $v(x,t) \le 0$. By the claims above, this shows $u(x,t) \le \epsilon$.

4. We will leave as a problem proving the inequality $-\epsilon \le u(x,t)$ via similar means.

Here is w:

$$w(x,t) := \frac{x^2 + 2t}{x_0^2 + 2t_0}. \tag{26.1}$$

This is something of a "miracle" function—it is just so simple!—and it is the sort of thing that one cooks up in a sudden 15 minutes of inspiration after a week of frustration.

26.1 Problem. Check that w is defined (i.e., no division by zero problems) and solves the heat equation $w_t = w_{xx}$ with $w(x_0,t_0) = 1$.

Now we need to find the right domain $\mathcal{D}$ on which we will apply the maximum principle to this v. Since v is defined for all $x \in \mathbb{R}$, we are free to choose any spatial interval that we like. Perhaps the simplest is symmetry: $-r \le x \le r$ for some $r > 0$. That is, we take

$$\mathcal{D} = \left\{(x,t) \in \mathbb{R}^2 \mid -r \le x \le r,\ 0 \le t \le T\right\}.$$

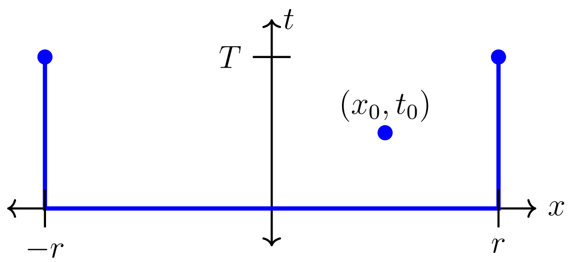
The maximum principle then guarantees that

$$\max_{(x,t)\in\mathcal{D}} v(x,t) = \max_{(x,t)\in\mathcal{D}_{\text{para}}} v(x,t),$$

where

$$\mathcal{D}_{\text{para}} = \left\{(-r,t) \in \mathbb{R}^2 \;\middle|\; 0 \le t \le T\right\} \cup \left\{(x,0) \in \mathbb{R}^2 \;\middle|\; -r \le x \le r\right\} \cup \{(r,t) \mid 0 \le t \le T\}$$

is the parabolic boundary. If we can show that the maximum of v on $\mathcal{D}_{\text{para}}$ is nonpositive, we are done. This will involve actually specifying r.



We start estimating. Because we are subtracting w in the definition of v, to get an upper bound on v on $\mathcal{D}_{\text{para}}$, we want *lower* bounds on w on $\mathcal{D}_{\text{para}}$. On the vertical sides of $\mathcal{D}_{\text{para}}$, we have $x = \pm r$ and $t \in [0, T]$, so

$$w(\pm r, t) = \frac{r^2 + 2t}{x_0^2 + 2t_0} \ge \frac{r^2}{x_0^2 + 2t_0},$$

so here, using the finite time bound on u (which says $u(x,t) \le |u(x,t)| \le M_T$ for $x \in \mathbb{R}$ and $t \in [0,T]$)

$$v(\pm r, t) \le u(\pm r, t) - \epsilon\frac{r^2}{x_0^2 + 2t_0} \le M_T - \epsilon\frac{r^2}{x_0^2 + 2t_0}.$$

If we take r sufficiently large relative to M_T, x_0, t_0, and ϵ, then

$$M_T - \frac{r^2}{x_0^2 + 2t_0} \le 0.$$

On the horizontal side, we have $|x| \le r$ and $t = 0$, so

$$w(x, 0) = \epsilon\frac{x^2}{x_0^2 + 2t_0} \ge 0,$$

and so here, using the initial condition $u(x,0) = 0$,

$$v(x,0) = u(x,0) - w(x,0) = -w(x,0) \le 0.$$

Thus $v(x,t) \le 0$ on the parabolic boundary, as desired. The reasoning above implies $u(x_0, t_0) \le \epsilon$, and we claim the other inequality holds.

26.2 Problem. Prove it. Specifically, by considering instead the "comparator" function $v(x,t) := -u(x,t) - \epsilon w(x,t)$, with w still defined as above, show that $-\epsilon \le u(x_0, t_0)$, which completes the argument.

We have done a lot of work on the heat equation posed spatially on $\mathbb{R}$, i.e., the "infinite rod" model. In particular, we had to assume an infinite spatial domain to use the Fourier transform. However, the maximum principle only required us to work on a finite domain. We might wonder, mathematically and physically, about the heat equation for a "finite rod."

As with the wave equation for a finite string, this involves boundary conditions:

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ u(x,0) = f(x), \ 0 \le x \le L \\ u(0,t) = a(t), \ u(L,t) = b(t), \ t \ge 0, \end{cases} \tag{26.2}$$

for given functions $f\colon [0,L] \to \mathbb{R}$ and a, $b\colon \mathbb{R} \to \mathbb{R}$. Now we require that the PDE hold at $t = 0$, unlike in our work on the line. It turns out that this is mathematically tractable in the sense that for the finite rod, we will eventually construct solutions that are genuinely differentiable at $t = 0$.

26.3 Problem. How differentiable do f, a, and b need to be for (26.2) to make sense? Recall that we assume that any solution to (26.2) is twice continuously differentiable for $0 \le x \le L$ and $t \ge 0$. Also, what are the values of $f(0)$, $f(L)$, $f''(0)$, and $f''(L)$?

Alternatively, we could pose a slightly different problem that demands continuity at $t = 0$:

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le L, \ t > 0 \\ u(x,0) = f(x), \ 0 \le x \le L \\ \lim_{(s,t)\to(x,0^+)} u(s,t) = f(x) \\ u(0,t) = a(t), \ u(L,t) = b(t), \ t \ge 0. \end{cases} \tag{26.3}$$

Either way, we can use an energy method to prove uniqueness.

26.4 Problem. Explain why to show uniqueness of (26.2), it suffices to show that if u solves

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ u(x,0) = 0, \ 0 \le x \le L \\ u(0,t) = u(L,t) = 0, \ t \ge 0, \end{cases} \tag{26.4}$$

then $u = 0$.

26.5 Theorem. *There is only one solution to* (26.4).

Proof. We assume that u solves (26.2) and study the "energy integral"

$$E(t) := \int_0^L u(x,t)^2 \, dx.$$

We have $E(t) \ge 0$ since the integrand is nonnegative and

$$E(0) = \int_0^L u(x,0)^2 \, dx = \int_0^L 0 \, dx = 0$$

from the initial condition $u(x,0) = 0$. We will show $E'(t) \le 0$ for all t. Then E is decreasing, so $E(t) \le 0$ for all $t \ge 0$. But $0 \le E(t)$ as well, so $0 \le E(t) \le 0$, and therefore $E(t) = 0$ for all t.

Onwards to differentiating. A solution to (26.4) is really twice continuously differentiable for $0 \le x \le L$ and $t \ge 0$, per our PDE conventions, so Leibniz's rule applies. We have

$$E'(t) = \partial_t \left[\int_0^L u(x,t)^2 \, dx\right] = \int_0^L \partial_t[u(x,t)^2] \, dx = \int_0^L 2u(x,t)u_t(x,t) \, dx$$

after differentiating under the integral and using the chain rule. This is an integral with respect to x, and we can connect u_t and u_{xx} via the heat equation, so we should probably do so and obtain

$$E'(t) = 2\int_0^L u(x,t)u_{xx}(x,t) \, dx.$$

If in a calculus class we encountered an antidifferentiation problem of the form $\int ff''$, we would probably integrate by parts with $u = f$ (terrible notation here) and $dv = f''$ to conclude $\int ff'' = ff' - \int f'f'$. Doing so here gives

$$E'(t) = 2u(x,t)u_x(x,t)\big|_{x=0}^{x=L} - 2\int_0^L u_x(x,t)^2 \, dx.$$

Since $u(0,t) = u(L,t) = 0$ by the boundary conditions, the first terms are 0, and so

$$E'(t) = -2\int_0^L u_x(x,t)^2 \, dx \le 0,$$

as desired. ■

26.6 Problem. To prove uniqueness for (26.3), we would want to show that if u solves

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le L, \ t > 0 \\ u(x,0) = 0, \ 0 \le x \le L \\ \lim_{(s,t)\to(x,0^+)} u(s,t) = 0 \\ u(0,t) = u(L,t) = 0, \ t \ge 0. \end{cases} \tag{26.5}$$

then $u = 0$. Reread the proof of the preceding theorem and explain why it still works for (26.5). [Hint: *the subtle point is that maybe E is now only differentiable on* $(0,\infty)$. *Why? Does that really matter?*]

For simplicity, we will usually work with $t \ge 0$ in the PDE part of the heat equation on finite spatial domains. That is, we will not consider (26.5) much further.

Day 27: Monday, October 14.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Remark (2) on pp. 141–142 discusses "continuous mean-square dependence" on initial data. The paragraph at the start of "The Maximum Principle and its consequences" on pp. 142–143 discusses how $\|\cdot\|_{L^2}$ does not imply control over $\|\cdot\|_{\infty}$. Continuous dependence on initial conditions in $\|\cdot\|_{\infty}$ appears in Theorem 3 on p. 147; see also the remark at the bottom of that page and Example 3 on p. 148.

The proof of Theorem 26.5 contains an important auxiliary result. Suppose that w solves

$$\begin{cases} w_t = w_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ w(0,t) = w(L,t) = 0, \ t \ge 0 \end{cases}$$

and is continuous for $0 \le x \le L$ and $t \ge 0$. Then the function

$$E[w] := \int_0^L w(x,t)^2 \, dx$$

is decreasing, and so $E[w](t) \le E[w](0)$ for all $t \ge 0$.

27.1 Problem. Reread the proof of Theorem 26.5 and convince yourself that this is true. (Here we are saying nothing about an initial condition for $w(x,0)$, and so we no longer conclude that $E[w](t) = 0$ for all t, nor do we want to.)

Now suppose we have two heat solutions with the same boundary conditions but possibly different initial conditions:

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ u(x,0) = a(t), \ t \ge 0 \\ u(L,t) = b(t), \ t \ge 0 \\ u(x,0) = f_1(x), \ 0 \le x \le L \end{cases} \qquad \text{and} \qquad \begin{cases} v_t = v_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ v(x,0) = a(t), \ t \ge 0 \\ v(L,t) = b(t), \ t \ge 0 \\ v(x,0) = f_2(x), \ 0 \le x \le L. \end{cases}$$

Put $w = u - v$, so w solves

$$\begin{cases} w_t = w_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ w(x,0) = 0, \ t \ge 0 \\ w(L,t) = 0, \ t \ge 0 \\ w(x,0) = f_1(x) - f_2(x), \ 0 \le x \le L. \end{cases}$$

Then the energy integral above implies $E[w](t) \le E[w](0)$ for all t, where

$$E[w](t) = \int_0^L w(x,t)^2 \, dx = \int_0^L \big[u(x,t) - v(x,t)\big]^2 \, dx$$

and

$$E[w](0) = \int_0^L w(x,0)^2 \, dx = \int_0^L \big[f_1(x) - f_2(x)\big]^2 \, dx.$$

That is,

$$\int_0^L \big[u(x,t) - v(x,t)\big]^2 \, dx \le \int_0^L \big[f_1(x) - f_2(x)\big]^2 \, dx \tag{27.1}$$

for all $t \ge 0$.

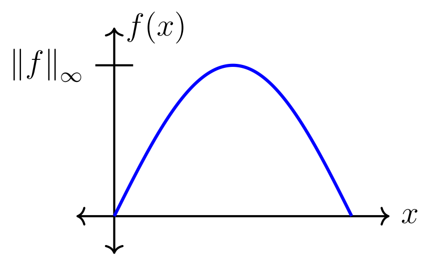
This should feel like our "continuous dependence on initial conditions" result for the wave equation. Both here and there, we bounded a difference of two solutions to the same PDE in terms of a difference of the initial conditions. The difference here is that the differences are no longer purely pointwise—they involve integrals.

How should we interpret these integrals? We need some new analytic tools. Recall that we have said, repeatedly, that integrals *represent* functions and *extract* and *measure* data about functions. We have seen such representations via FTC1 and the Fourier integral in Untheorem 22.2 (the latter still needing some patching up). We have seen data extracted via Fourier modes—such data is useful for representation purposes as in that touchy untheorem, and there are other uses to come.

One of the most natural measurements to desire about a function is its size: how "large" is it? That depends on one's perspective. Say that $f \colon [0, L] \to \mathbb{C}$ is continuous. From calculus, f has extreme values, and so

$$\|f\|_\infty := \max_{0 \le x \le L} |f(x)|$$

is defined.



We might call this number $\|f\|_\infty$ the "maximum norm" of f. This value $\|f\|_\infty$ measures how large in a "pointwise" sense f can be.

We might also think about average value. The number

$$\frac{1}{L}\int_0^L |f(x)| \, dx$$

is the average value of $|f|$ on $[0, L]$. Put

$$\|f\|_{L^1} := \int_0^L |f(x)| \, dx.$$

We call this the "L^1-norm" of f. (The unfortunate overworking of L as both an endpoint of the domain of f and part of the name of the norm is an accident of culture and bad writing.) When $L > 0$ is fixed, we might think that if $\|f\|_{L^1}$ is a "large" number, then "on average" f should be large, while if $\|f\|_{L^1}$ is small, then "on average" f should be small.

But neither $\|\cdot\|_\infty$ nor $\|\cdot\|_{L^1}$ appears in our estimates for the heat equation. Instead, we introduce the "L^2-norm":

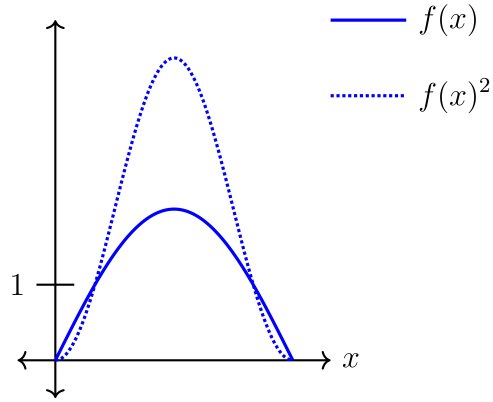
$$\|f\|_{L^2} := \left(\int_0^L |f(x)|^2 \, dx\right)^{1/2}.$$

The square root preserves a nice scaling property: $\|\alpha f\|_{L^2} = |\alpha| \, \|f\|_{L^2}$. Such a scaling property is already present in $\|\cdot\|_\infty$ and $\|\cdot\|_{L^1}$. Then (27.1) says

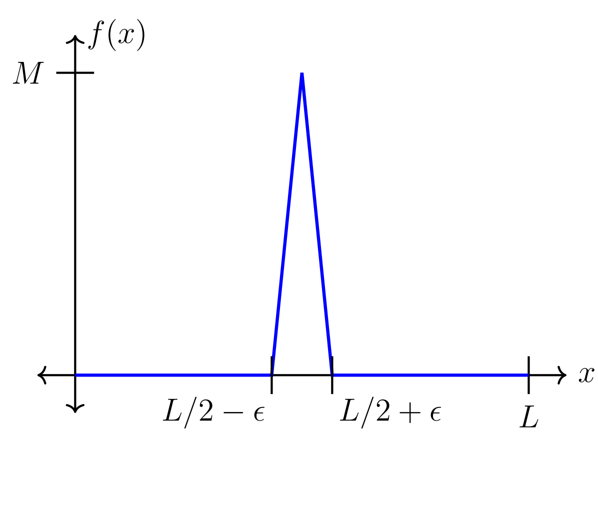
$$\|u(\cdot, t) - v(\cdot, t)\|_{L^2} \le \|f_1 - f_2\|_{L^2}.$$

As usual, $u(\cdot, t)$ is the function from $[0, L]$ to $\mathbb{R}$ given by $x \mapsto u(x, t)$, and the same for $v(\cdot, t)$.

If $\|\cdot\|_\infty$ measures pointwise extremes, and $\|\cdot\|_{L^1}$ measures average value, what is left for $\|\cdot\|_{L^2}$ to measure? The less satisfying answer is that $\|\cdot\|_{L^2}$ is simply a "mathematically nicer" norm for a variety of reasons that we have yet to encounter. The possibly satisfying "physical" reason is the slogan "squaring makes small things smaller and larger things larger." Recall that if $0 < y < 1$, then $0 < y^2 < y < 1$, while if $1 < y$, then $1 < y < y^2$. Perhaps a function f records the outputs of an experiment or the difference between two experiments; in comparing those differences, we might want an instrument that magnifies "large" differences but penalizes "small" differences less. Squaring in the L^2-norm introduces that magnification/penalization behavior while still retaining the "averaging" behavior of the L^1-norm.



Our gut instinct is probably to want estimates in $\|\cdot\|_\infty$, as this is what we know best from calculus so far. Unfortunately, an estimate in $\|\cdot\|_{L^1}$ or $\|\cdot\|_{L^2}$ need not imply an estimate in $\|\cdot\|_\infty$. Consider the following picture, where ϵ, L, $M > 0$ are fixed.



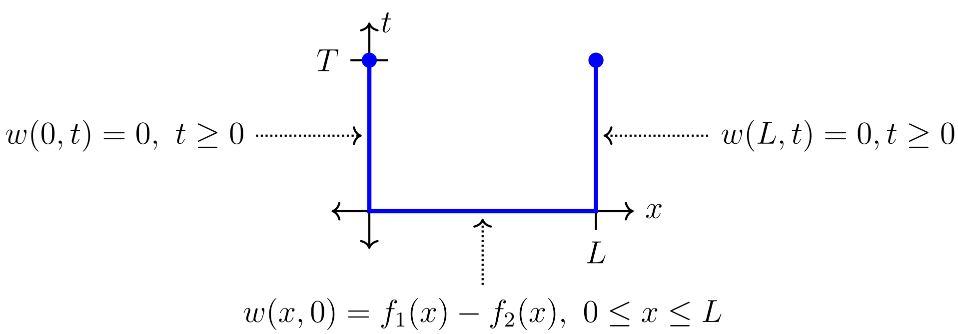
It should be the case that $\|f\|_\infty = M$ but $\|f\|_{L^1}$ and $\|f\|_{L^2}$ are "small" since the area under the graph of $f = |f|$ is quite small.

27.2 Problem. Quantify this. First, find a piecewise formula for f (you may assume that what look like line segments in the drawing above are actually line segments). Then calculate $\|f\|_{L^1}$ and $\|f\|_{L^2}$. Explain precisely how $\|f\|_\infty$ can be "large" even though $\|f\|_{L^1}$ and $\|f\|_{L^2}$ are "small." [Hint: *try something like* $M = 1/\epsilon$ *or* $M = 1/\epsilon^2$.]

Nonetheless, we can still obtain $\|\cdot\|_\infty$ estimates on differences of solutions to the heat boundary value problems. The maximum principle and Problem 25.5 together tell us that if

$$\begin{cases} w_t = w_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ w(0,t) = w(L,t) = 0, \ t \ge 0 \\ w(x,0) = f_1(x) - f_2(x), \ 0 \le x \le L, \end{cases}$$

then the maximum and minimum values of w over $0 \le x \le L$ and $0 \le t \le T$ (for any $T > 0$) occur on the "parabolic boundary" sketched below.



On the vertical sides of this boundary ($x = 0$ and $x = L$), $w = 0$, so those sides contribute nothing really interesting to the extreme values of w. Suppose $-\epsilon \le f_1(x) \le f_2(x) \le \epsilon$ for all $x \in [0, L]$. Then the minimum of w on the parabolic boundary is at least $-\epsilon$ and the maximum is at least ϵ, so $-\epsilon \le w(x,t) \le \epsilon$ for all $x \in [0, L]$ and $t \in [0, T]$. Even better, this estimate is independent of T, so it is true for all $t \ge 0$.

Here is what we have proved.

27.3 Theorem. *Suppose that u and v solve*

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ u(x,0) = a(t), \ t \ge 0 \\ u(L,t) = b(t), \ t \ge 0 \\ u(x,0) = f_1(x), \ 0 \le x \le L \end{cases} \quad \text{and} \quad \begin{cases} v_t = v_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ v(x,0) = a(t), \ t \ge 0 \\ v(L,t) = b(t), \ t \ge 0 \\ v(x,0) = f_2(x), \ 0 \le x \le L. \end{cases}$$

(i) *If $\|f_1 - f_2\|_{L^2} < \epsilon$, then $\|u(\cdot,t) - v(\cdot,t)\|_{L^2} < \epsilon$ for all $t \ge 0$.*

(ii) *If $\|f_1 - f_2\|_\infty < \epsilon$, then $\|u(\cdot,t) - v(\cdot,t)\|_\infty < \epsilon$ and $\|u(\cdot,t) - v(\cdot,t)\|_{L^2} < L\epsilon$ for all*

$t \geq 0$.

27.4 Problem. Prove the L^2-estimate in part (ii) of Theorem 27.3.

We could also ask for a continuous dependence on initial conditions result for the heat equation on the line. Since there are no inherent boundary conditions (notwithstanding a boundedness in finite time condition that we might impose to guarantee uniqueness), we just ask how we might estimate the solution to

$$\begin{cases} u_t = u_{xx}, \ x \in \mathbb{R}, \ t > 0 \\ u(x,0) = f(x), \ x \in \mathbb{R} \end{cases}$$

in terms of f. There is not much to do here: if there is $M > 0$ such that $|f(y)| \leq M$ for all $y \in \mathbb{R}$, then (for $t > 0$)

$$|u(x,t)| = \left| \int_{-\infty}^{\infty} H(x-y,t)f(y) \ dy \right| \leq M \int_{-\infty}^{\infty} H(x-y,t) \ dy.$$

We basically showed in Problem 24.4 that this integral is independent of both x and y. Call its value C, so we have shown

$$|u(x,t)| \leq CM,$$

where C is independent of x, t, and f, and M depends only on f.

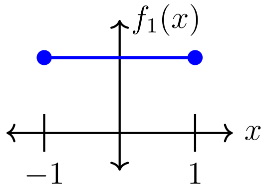
We could go further and ask about "average" behavior of solutions to the heat equation on the line. This would require introducing the improper integral analogues of the L^1- and L^2-norms, and generalizing $\|\cdot\|_\infty$ to functions defined on $\mathbb{R}$ (and note that continuous functions on $\mathbb{R}$ need not attain an absolute maximum or minimum there—think of arctan). This is where we are headed anyway, and it will be a natural part of our upcoming retread of the Fourier transform—we have eaten dessert first, and now it is time for better nutrition.

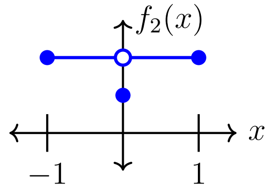
Day 28: Wednesday, October 16.

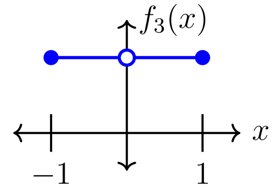
Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 222–223 discuss piecewise continuous functions. Examples 1, 2, and 3 on pp. 424–425 compute Fourier transforms from the definition. Problem 10 on p. 429 proves the Riemann–Lebesgue lemma.

An adequate theory of the Fourier transform requires us to integrate more than just continuous functions. Consider the following functions graphed below. The "area under the graph" between $x = -1$ and $x = 1$ of each function should be the same, even though the functions are quite different.







Any "good" notion of area should take into account that "the area under a point is 0." That f_2 disagrees with f_1 just at $x = 0$, and that f_3 is not even defined at $x = 0$ should not matter.

These three functions have two important features in common. First, they are "mostly" continuous in the sense that they are defined and continuous at all but finitely many points of $[-1, 1]$. The function f_1 is defined and continuous at all points in this interval; the function f_2 is defined everywhere but discontinuous at 0; and f_3 is not even defined at 0. Second, these functions have very good limit behaviors: the limit

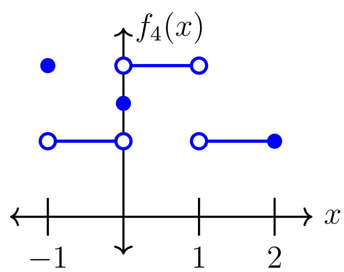
$$\lim_{s \to x^{\pm}} f_k(x)$$

exists for all $x \in (-1, 1)$ and each k, and "most of the time" these limits are equal (and equal $f_k(x)$ to boot) because the f_k are "mostly continuous. Also, the two limits

$$\lim_{s \to -1^+} f_k(x) \quad \text{and} \quad \lim_{s \to 1^-} f_k(x)$$

exist. The point is that the f_k are not all that badly behaved—in particular, there are no vertical asymptotes/blow-ups.

Here is a fourth function, now defined on (most of) $[-1, 2]$, that should have a well-defined area under its graph.



The only difference here is that sometimes $\lim_{s \to x^+} f_4(s) \neq \lim_{s \to x^-} f_4(s)$ for some $x \in (-1, 2)$.

All of these functions are what we want to call "piecewise continuous."

28.1 Definition. *Let a, $b \in \mathbb{R}$ with $a < b$ and suppose there is $\{x_k\}_{k=1}^n \subseteq [a, b]$ such that the function $f \colon [a, b] \setminus \{x_k\}_{k=1}^n \to \mathbb{C}$ is continuous and that all of the following limits exist:*

$$\lim_{s \to a^+} f(s), \qquad \lim_{s \to x^{\pm}} f(s), \ a < x < b, \qquad \text{and} \qquad \lim_{s \to b^-} f(s).$$

Then we say that f is **PIECEWISE CONTINUOUS** *on $[a, b]$. The set of all piecewise con-*

tinuous functions on $[a, b]$ is $\mathcal{C}_{\mathrm{pw}}([a, b])$.

We emphasize that a piecewise continuous function on $[a, b]$ need not be a function defined "on" $[a, b]$. The point is that any integral worth its salt should "forgive" the absence of definition at a few points and also jump discontinuities. We now give meaning to $\int_a^b f$ for $f \in \mathcal{C}_{\mathrm{pw}}([a, b])$, assuming that we only know how to integrate continuous functions, per Theorem 2.1.

Here are two illustrative examples of how this should proceed.

28.2 Example. **(i)** Let $f \in \mathcal{C}_{\mathrm{pw}}([a, b])$ be discontinuous and/or undefined at a. Define

$$\widetilde{f}\colon [a, b] \to \mathbb{C}\colon x \mapsto \begin{cases} \lim_{s \to a^+} f(s), & x = a \\ f(x), & a < x \le b. \end{cases}$$

Then $\widetilde{f} \in \mathcal{C}([a, b])$, and so we set $\int_a^b f := \int_a^b \widetilde{f}$.

(ii) Let $f \in \mathcal{C}_{\mathrm{pw}}([a, b])$ be discontinuous and/or undefined at $c \in (a, b)$. Define

$$\widetilde{f}_1\colon [a, c] \to \mathbb{C}\colon x \mapsto \begin{cases} f(x) & a \le x < c \\ \lim_{s \to c^-} f(s), & x = c \end{cases}$$

and

$$\widetilde{f}_2\colon [c, b] \to \mathbb{C}\colon x \mapsto \begin{cases} \lim_{s \to c^+} f(s), & x = c \\ f(x), & c < x \le b. \end{cases}$$

Then $\widetilde{f}_1 \in \mathcal{C}([a, c])$ and $\widetilde{f}_2 \in \mathcal{C}([c, b])$, and so we put $\int_a^b f := \int_a^c \widetilde{f}_1 + \int_c^b \widetilde{f}_2$.

For $f \in \mathcal{C}_{\mathrm{pw}}([a, b])$ with f discontinuous and/or undefined on $\{x_k\}_{k=1}^n \subseteq [a, b]$, we just do this repeatedly: break $[a, b]$ into n subintervals on which f is continuous and has left/right limits at the endpoints, create a continuous function by setting the values at the endpoints to be those left/right limits, and then defining $\int_a^b f$ as the sum of the (ordinary) integrals over those subintervals. We then recover the analogue of Theorem 2.1. We can even obtain a new meaning of integrating over a subinterval of $[a, b]$: for $a \le c < d \le b$, we have

$$\int_c^d f = \int_a^b f\chi_{[c,d]}, \qquad \chi_{[c,d]}(x) := \begin{cases} 1, & c \le x \le d \\ 0, & x < c \text{ or } c > d. \end{cases}$$

28.3 Theorem. *Every part of Theorem 2.1 remains true for $I = [a, b]$ when $\mathcal{C}([a, b])$ is replaced by $\mathcal{C}_{\mathrm{pw}}([a, b])$.*

28.4 Problem. Define

$$f\colon [-1,1] \to \mathbb{C}\colon x \mapsto \begin{cases} 0, & -1 \le x < 0 \\ 1, & 0 \le x \le 1 \end{cases} \qquad \text{and} \qquad F\colon [-1,1] \to \mathbb{C}\colon x \mapsto \int_{-1}^{x} f.$$

Show that $F \in \mathcal{C}([-1,1])$ but F is not differentiable at 0. Thus the fundamental theorem of calculus can fail when the integrand is not continuous!

We have defined piecewise continuity on a closed, bounded subinterval of $\mathbb{R}$, but the Fourier transform requires us to work on all of (well, most of) $\mathbb{R}$. We might like to say that $f \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ if $f \in \mathcal{C}_{\text{pw}}([a,b])$ for all $[a,b] \subseteq \mathbb{R}$, but this is a little awkward—what exactly is the domain of f? Here it helps to be more explicit (and annoying).

28.5 Definition. *Let $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ be a subset such that $\{x_k\}_{k=1}^{\infty} \cap [a,b]$ is finite for any closed, bounded subinterval $[a,b] \subseteq \mathbb{R}$; equivalently, there is $d > 0$ such that $|x_k - x_j| \ge d$ for all $j \ne k$. Suppose that $f\colon \mathbb{R} \setminus \{x_k\}_{k=1}^{\infty} \to \mathbb{C}$ is continuous and that the limits*

$$f(x^{\pm}) = \lim_{s \to x^{\pm}} f(x)$$

exist for all $x \in \mathbb{R}$. Then f is **PIECEWISE CONTINUOUS** *on $\mathbb{R}$. The set of all piecewise continuous functions on $\mathbb{R}$ is $\mathcal{C}_{\text{pw}}(\mathbb{R})$.*

As before, a piecewise continuous function on $\mathbb{R}$ does not have to be defined "on" $\mathbb{R}$, though certainly we have $\mathcal{C}(\mathbb{R}) \subseteq \mathcal{C}_{\text{pw}}(\mathbb{R})$. Unlike before, since $\mathbb{R}$ has no endpoints, we just need to know that the left and right limits at any point in $\mathbb{R}$ exist. Continuity forces their equality off $\{x_k\}_{k=1}^{\infty}$. Also, the condition that $\{x_k\}_{k=1}^{\infty} \cap [a,b]$ be finite for all closed, bounded subintervals $[a,b] \subseteq \mathbb{R}$ just ensures that the points at which $f \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ is discontinuous and/or undefined do not "crowd up" too much anywhere. In practice, the functions that we will typically meet will be discontinuous and/or undefined at only a "few" points in $\mathbb{R}$, probably finitely many, sometimes only at one or two points.

We now define improper integrals of piecewise continuous functions on $\mathbb{R}$ using the integral for piecewise continuous functions on subintervals of $\mathbb{R}$.

28.6 Definition. **(i)** *Let $f \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ and suppose that the limits*

$$\int_{-\infty}^{0} f := \lim_{a \to -\infty} \int_{a}^{0} f \qquad \text{and} \qquad \int_{0}^{\infty} f := \lim_{b \to \infty} \int_{0}^{b} f$$

exist. Then f is **INTEGRABLE** *on $\mathbb{R}$ and we put*

$$\int_{-\infty}^{\infty} f = \int_{-\infty}^{0} f + \int_{0}^{\infty} f.$$

(ii) *A function $f \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ is* **ABSOLUTELY INTEGRABLE** *if $|f|$ is integrable, and we denote by L^1, or sometimes $L^1(\mathbb{R})$, the set of all absolutely integrable functions. The L^1-*

(SEMI)NORM *of* $f \in L^1$ *is*

$$\|f\|_{L^1} := \int_{-\infty}^{\infty} |f|.$$

As much as possible, we should try to avoid working with the definition of the integral (sometimes this is *not* possible) to show that a function is integrable. Fortunately, the comparison test carries over.

28.7 Theorem (Comparison test). *Let* $f \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ *and* $g \in L^1$ *with* $|f(x)| \le |g(x)|$ *for* $x \in \mathbb{R} \setminus \{x_k\}_{k=1}^{\infty}$ *for some set* $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$. *Here we assume that* $\mathbb{R} \setminus \{x_k\}_{k=1}^{\infty}$ *is contained in the domains of both* f *and* g. *Then* $f \in L^1$ *and*

$$\left|\int_{-\infty}^{\infty} f\right| \le \int_{-\infty}^{\infty} |f| \le \int_{-\infty}^{\infty} |g|.$$

28.8 Problem. Prove that L^1 is a vector space in the sense that $\alpha f \in L^1$ and $f + g \in L^1$ for all f, $g \in L^1$ and $\alpha \in \mathbb{R}$. (For the sum, use the triangle inequality for real numbers and the comparison test; if $f \colon \mathbb{R} \setminus \{x_k\}_{k=1}^{\infty} \to \mathbb{C}$ and $g \colon \mathbb{R} \setminus \{y_k\}_{k=1}^{\infty} \to \mathbb{C}$ are continuous, interpret the domain of $f + g$ as $\mathbb{R} \setminus \{x_k, y_k\}_{k=1}^{\infty}$.)

28.9 Problem. Prove that $\|\cdot\|_{L^1}$ satisfies three of the four properties that one usually demands of a "norm" on a vector space:

$$\|f\|_{L^1} \ge 0, \qquad \|\alpha f\|_{L^1} = |\alpha| \, \|f\|_{L^1}, \qquad \text{and} \qquad \|f+g\|_{L^1} \le \|f\|_{L^1} + \|g\|_{L^1}$$

for all f, $g \in L^1$ and $\alpha \in \mathbb{R}$. However, find an example of $f \in L^1$ such that $f \ne 0$ but $\|f\|_{L^1} = 0$. For this reason, we might strictly prefer to call $\|\cdot\|_{L^1}$ a "seminorm" (for it to be a norm, we would want $\|f\|_{L^1} = 0$ to force $f = 0$).

At last we are ready to (re)define the Fourier transform.

28.10 Definition. *The* FOURIER TRANSFORM *of* $f \in L^1$ *at* $k \in \mathbb{R}$ *is*

$$\widehat{f}(k) = \mathfrak{F}[f](k) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} \, dx.$$

We first observe that while $f \in L^1$ need not be defined on all of $\mathbb{R}$, the Fourier transform definitely is; convergence of the Fourier integral follows from the comparison test and the identity $|e^{-ikx}| = 1$ for all k, $x \in \mathbb{R}$. We will soon see that $\widehat{f}$ can be a *much* nicer function than f.

Now we actually calculate a Fourier transform (this is the first time that we have done so directly from the definition!).

28.11 Example. Fix $r > 0$ and let

$$f(x) := \begin{cases} 1, & |x| \le r \\ 0, & |x| > r. \end{cases} \tag{28.1}$$

It is a worthwhile exercise in the definition of the integral to show $f \in L^1$; we leave this as a problem. It then follows from Problem 21.12 that

$$\widehat{f}(k) = \frac{1}{\sqrt{2\pi}} \lim_{R \to \infty} \int_{-R}^{R} f(x) e^{-ikx} \, dx,$$

so we calculate this integral. We may as well assume $R > r$, and then

$$\int_{-R}^{R} f(x) e^{-ikx} \, dx = \int_{-r}^{r} e^{-ikx} \, dx.$$

That is,

$$\lim_{R \to \infty} \int_{-R}^{R} f(x) e^{-ikx} \, dx = \int_{-r}^{r} e^{-ikx} \, dx,$$

and so

$$\widehat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} e^{-ikx} \, dx.$$

For $k = 0$, this gives

$$\widehat{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} 1 \, dx = \frac{2r}{\sqrt{2\pi}},$$

while for $k \ne 0$ we have

$$\begin{aligned} \widehat{f}(k) = \frac{1}{\sqrt{2\pi}} \frac{1}{-ik} \int_{-r}^{r} -ike^{-ikx} \, dx = -\frac{1}{ik\sqrt{2\pi}} \int_{-r}^{r} \partial_x [e^{-ikx}] \, dx = \frac{e^{ikr} - e^{-ikr}}{ik\sqrt{2\pi}} \\ = \frac{2r}{kr\sqrt{2\pi}} \frac{e^{ikr} - e^{-ikr}}{2i} = \frac{2r}{\sqrt{2\pi}} \frac{\sin(kr)}{kr}. \end{aligned}$$

Now put

$$\operatorname{sinc}(X) := \begin{cases} \sin(X)/X, & X \ne 0 \\ 1, & X = 0. \end{cases}$$

We have shown

$$\widehat{f}(k) = \frac{2r}{\sqrt{2\pi}} \operatorname{sinc}(kr),$$

and this is a rather nicer function than f, since we recall from calculus that sinc is infinitely differentiable.

28.12 Problem. Do that worthwhile exercise and show that $f \in L^1$, where f is defined in (28.1).

28.13 Problem. Let $f \in L^1$.

(i) Prove that if f is even, then $\widehat{f}$ is even, while if f is odd, then $\widehat{f}$ is odd.

(ii) Prove that $\overline{\widehat{f}(k)} = \widehat{f}(-k)$. Here $\overline{x+iy} = x - iy$ is the **COMPLEX CONJUGATE** of $x + iy \in \mathbb{C}$. [Hint: *use the fact that* $\int_a^b \overline{f} = \overline{\int_a^b f}$.]

(iii) Prove that if f is real-valued and even, then $\widehat{f}$ is also real-valued. How does this explain the result of Example 28.11?

The nice properties of $\widehat{f}$ as computed in Example 28.11 carry over well to Fourier transforms in general.

28.14 Theorem. *Let* $f \in L^1$.

(i) *The Fourier transform is linear: for* f, $g \in L^1$ *and* α, $\beta \in \mathbb{C}$,

$$\widehat{\alpha f + \beta g}(k) = \alpha \widehat{f}(k) + \beta \widehat{g}(k).$$

(ii) *The Fourier transform* $\widehat{f}$ *is bounded with the estimate*

$$|\widehat{f}(k)| \le \frac{\|f\|_{L^1}}{\sqrt{2\pi}}$$

for all $k \in \mathbb{R}$.

(iii) *The Fourier transform is continuous:* $\widehat{f} \in \mathcal{C}(\mathbb{R})$.

(iv) [Riemann–Lebesgue lemma] *The Fourier transform vanishes at* $\pm\infty$*:* $\lim_{k \to \pm\infty} \widehat{f}(k) = 0$.

Proof. Linearity is just the vector space property of L^1 and the linearity of the integral. The boundedness estimate is the triangle inequality:

$$|\widehat{f}(k)| = \frac{1}{\sqrt{2\pi}} \left| \int_{-\infty}^{\infty} f(x) e^{-ikx} \, dx \right| \le \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |f(x) e^{-ikx}| \, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |f(x)| \, dx = \frac{\|f\|_{L^1}}{\sqrt{2\pi}}.$$

The proofs of continuity and vanishing at $\pm\infty$ require more ϵ–δ-type analysis than we care to do here. We discuss one situation of continuity in the problem below, under more hypotheses, and we discuss vanishing at $\pm\infty$ later under other, and more, hypotheses.

Note that $f \in L^1$ implies $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$, by definition of L^1, but now $\widehat{f} \in \mathcal{C}(\mathbb{R})$: this is much better behavior for $\widehat{f}$! ■

28.15 Problem. **(i)** A natural way to try to prove continuity of $\widehat{f}$ is to rewrite

$$|\widehat{f}(k_2) - \widehat{f}(k_1)| = \frac{|k_2 - k_1|}{\sqrt{2\pi}} \left| \int_{-\infty}^{\infty} \left(\int_0^1 e^{-ix[(1-t)k_1 + tk_2]} \, dt \right) x f(x) \, dx \right|. \tag{28.2}$$

First show that (28.2) is true. The utility of (28.2) is that it exposes transparently a factor of $k_2 - k_1$; we might hope that if $k_2 - k_1$ is small, then we could use the boundedness of the improper integral to show that $|\widehat{f}(k_2) - \widehat{f}(k_1)|$ is small. Unfortunately, this integral may not converge unless $(Jf)(x) := xf(x)$ is absolutely integrable. Suppose, now, that $Jf \in L^1$ and use (28.2) to prove continuity of $\widehat{f}$.

(ii) Prove that if all we know is that $\widehat{f}$ is continuous and vanishes at $\pm\infty$, then $\widehat{f}$ is bounded (although this does not help us recover the bound in terms of $\|f\|_{L^1}$). *Do not use the definition of* $\widehat{f}$*, just the two properties of being continuous on* $\mathbb{R}$ *and vanishing at* $\pm\infty$.

Day 29: Friday, October 18.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Proposition 1 on p. 432 proves the all-important identity $\widehat{f'}(k) = ik\widehat{f}(k)$ using the decay order language from p. 432. See also Corollary 1 on that page. Theorem 1 on p. 433 gives another condition on $f \in L^1$ that guarantees $\widehat{f} \in L^1$ (among other nice things). See also the remark and example on p. 434. Pages 447–448 state and discuss the inversion theorem, and its proof begins on p. 498. Example 3 on pp. 436–437 applies the Fourier transform to $y' = ay + f(x)$.

We formalize the most useful property of the Fourier transform.

29.1 Theorem. *Let* $f \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ *with* $f' \in L^1(\mathbb{R})$*. Then* $\widehat{f'}(k) = ik\widehat{f}(k)$ *for all* $k \in \mathbb{R}$*.*

Proof. Since $f' \in L^1$, Problem 21.12 allows us to write

$$\widehat{f'}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(x)e^{-ikx} \, dx = \frac{1}{\sqrt{2\pi}} \lim_{R\to\infty} \int_{-R}^{R} f'(x)e^{-ikx} \, dx.$$

The goal is to get f, not f', to show up in the Fourier integral. We are now working with the integral of the product of a derivative (the $f'(x)$ factor) and a very well-behaved factor that is easy to differentiate and antidifferentiate (the e^{-ikx} factor), so we integrate by parts with $dv = f'(x)\,dx$ (as we do not necessarily have more derivatives on f) and $u = e^{-ikx}$ to find

$$\begin{aligned}\int_{-R}^{R} f'(x)e^{-ikx} \, dx &= f(x)e^{-ikx}\Big|_{x=-R}^{x=R} - \int_{-R}^{R} f(x)(-ik)e^{-ikx} \, dx \\ &= f(R)e^{-ikR} - f(-R)e^{ikR} + ik \int_{-R}^{R} f(x)e^{-ikx} \, dx.\end{aligned}$$

Now we see the Fourier integral for f start to emerge. It would be nice if those "boundary" terms at the start vanished, and they do: the hypotheses here imply $\lim_{R\to\pm\infty} f(R) = 0$, and so we have

$$\lim_{R\to\infty}\left(f(R)e^{-ikR} - f(-R)e^{ikR} + ik\int_{-R}^{R} f(x)e^{-ikx}\,dx\right) = 0 + ik\int_{-\infty}^{\infty} f(x)e^{-ikx}\,dx.$$

Incorporating our insistent factor of $1/\sqrt{2\pi}$, we are done. ■

29.2 Problem. Let $f \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ with $f' \in L^1(\mathbb{R})$. This problem shows $\lim_{x\to\infty} f(x) = 0$, and an analogous argument can treat the limit at $-\infty$.

(i) By considering the identity $f(x) = f(0) + \int_0^x f'(s)\,ds$, show that $L := \lim_{x\to\infty} f(x)$ exists.

(ii) Prove by contradiction that $L = 0$: if $L > 0$, argue that for some $M > 0$ and all $x \geq M$, we have $f(x) \geq L/2$. Why does this contradict $f \in L^1$? Make a similar argument if $L < 0$.

(iii) Give an example of $f \in L^1$ such that $\lim_{x\to\infty} f(x) \neq 0$. [Hint: *construct f so that the limit does not even exist. It will be hard to get a continuous f to do that, so allow f to be discontinuous, possibly in infinitely many places; f can be mostly* 0*, except where it is not* 0.]

29.3 Remark. *The previous problem and theorem have profound consequences. Here are the easy sound bites before we see those consequences. First, ODE become algebraic equations under the Fourier transform, while PDE become ODE in the "time" variable under the Fourier transform in the "spatial" variable:*

$$\widehat{ODE} = \textit{algebraic}, \qquad \widehat{PDE} = ODE.$$

Second, and unfortunately, it is not guaranteed that the candidate solution from the Fourier transform will encompass all solutions to a problem. The reason is that the hypotheses of Theorem 29.1 presume decay of the solution at (spatial) infinity.

As we previously observed in our study of the heat equation, we can iterate Theorem 29.1 and take the transform of any derivative.

29.4 Corollary. *Let $f \in \mathcal{C}^r(\mathbb{R})$ for some $r \geq 1$ and suppose $f^{(j)} \in L^1$ for $j = 0, \ldots, r$. Then $\widehat{f^{(r)}}(k) = (ik)^r \widehat{f}(k)$.*

This gives new insight into the decay of Fourier coefficients as $k \to \pm\infty$. For $k \neq 0$, the corollary says

$$\widehat{f}(k) = \frac{\widehat{f^{(r)}}(k}{(ik)^r} \quad \text{and so} \quad |\widehat{f}(k)| \leq \frac{\left\|f^{(r)}\right\|_{L^1}}{\sqrt{2\pi}}\left(\frac{1}{|k|^r}\right). \tag{29.1}$$

That is, if f is r-times continuously differentiable and all of its first r derivatives are integrable, then $\widehat{f}$ decays like k^{-r} at $\pm\infty$. This is fast!

Next we consider inverting the Fourier transform. Untheorem 22.2 gives us the (unfortunately false) hope that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)e^{ikx}\ dk \tag{29.2}$$

for $f \in L^1$. The first problem here is that maybe $\widehat{f} \notin L^1$. This turns out to be the case with the function f from Example 28.11, namely because sinc $\notin L^1$ (which follows from a picky argument involving exploding lower bounds and the harmonic series).

It is possible to add hypotheses to $f \in L^1$ that guarantee $\widehat{f} \in L^1$. Here is one example.

29.5 Problem. Let $f \in L^1 \cap \mathcal{C}^2(\mathbb{R})$ with f', $f'' \in L^1$. Show that $\widehat{f} \in L^1$. [Hint: *how fast does* $\widehat{f}$ *decay?*]

So, we might want to relax the inversion statement to involve the weaker symmetric limit in the integral: maybe

$$f(x) = \frac{1}{\sqrt{2\pi}} \text{P.V.} \int_{-\infty}^{\infty} \widehat{f}(k)e^{ikx}\ dk = \frac{1}{\sqrt{2\pi}} \lim_{R\to\infty} \int_{-R}^{R} \widehat{f}(k)e^{ikx}\ dk.$$

This too turns out to be false.

29.6 Problem. Explain why by considering

$$f(x) := \begin{cases} 1, & x = 0 \\ 0, & x \neq 0. \end{cases}$$

Explain how this means that Untheorem 22.3 is really just that—an untheorem, not a true theorem.

The true story of Fourier inversion involves "averages." For $f \in L^1$, the limits

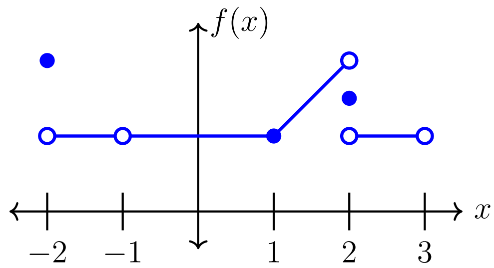
$$f(x^{\pm}) := \lim_{s \to x^{\pm}} f(s)$$

exist for all $x \in \mathbb{R}$, and, with slightly more hypotheses on f, the actual Fourier inversion formula is

$$\frac{f(x^+) + f(x^-)}{2} = \frac{1}{\sqrt{2\pi}} \lim_{R\to\infty} \int_{-R}^{R} \widehat{f}(k)e^{ikx}\ dk. \tag{29.3}$$

29.7 Problem. Suppose that $f \in L^1 \cap \mathcal{C}(\mathbb{R})$ with $\widehat{f} \in L^1$ and (29.3) is true for some $x \in \mathbb{R}$. Why does that imply about (29.2)?

What are those other hypotheses? The function f needs to be more than just piecewise continuous (like all functions in L^1)—it needs to be *piecewise continuously differentiable*. We first give the definition by a picture, and then we give the definition.



This is a pretty nice function on $[-2, 3]$. It has some jump discontinuities, and it is not defined at -1, and it is defined and continuous but not differentiable at 1, but it is definitely the case that

$$f(x^{\pm}) := \lim_{s \to x^{\pm}} f(s)$$

exist for all $x \in (-2, 3)$, and that

$$\lim_{s \to -2^+} f(s) \quad \text{and} \quad \lim_{s \to 3^-} f(s)$$

exist, and that the limits of the left and right difference quotients

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x^+)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(x+h) - f(x^-)}{h} \tag{29.4}$$

exist at all $x \in (-2, 3)$. Also, the limits of the appropriate difference quotients at the endpoints exist:

$$\lim_{h \to 0^+} \frac{f(-2+h) - f(-2^+)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(3+h) - f(3^-)}{h}.$$

Suppose that we take the unusual step of defining $f'(x)$ to be either of the difference quotients in (29.4) when they are both equal. Then

$$f'(x) = \begin{cases} 0, & x \neq -2,\ 1,\ 2,\ 3 \text{ and } x < 1 \text{ or } x > 2 \\ 1, & 1 < x < 2. \end{cases}$$

This function f' is piecewise continuous on $[-2, 3]$ in the sense of Definition 28.1!

Our current interest, however, is in functions defined on (most of) $\mathbb{R}$, so we adapt this example from $[-2, 3]$ to $\mathbb{R}$.

29.8 Definition. *Let $f \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ and suppose that*

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x^+)}{h} \quad \textit{and} \quad \lim_{h \to 0^-} \frac{f(x+h) - f(x^-)}{h}$$

exist for all $x \in \mathbb{R}$. (In the expression $f(x+h)$ we are assuming that $h \neq 0$ is so small relative to x that f is actually defined at $x+h$; this is possible because, by definition of $\mathcal{C}_{\text{pw}}(\mathbb{R})$, f is defined at all but countably many points in $\mathbb{R}$, and these points all lie at a certain minimum positive distance from each other.)

Suppose also that these limits are equal for $x \in \mathbb{R} \setminus \{x_k\}_{k=1}^{\infty}$, where $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ is a set such that $\{x_k\}_{k=1}^{\infty} \cap [a,b]$ is finite for any closed, bounded subinterval $[a,b] \subseteq \mathbb{R}$. For such x, let $f'(x)$ be either of the limits above. Suppose last that $f' \in \mathcal{C}_{\text{pw}}(\mathbb{R})$. Then we say that f is **PIECEWISE CONTINUOUSLY DIFFERENTIABLE** *on $\mathbb{R}$, and we denote the set of such functions by $\mathcal{C}^1_{\text{pw}}(\mathbb{R})$.*

As with $\mathcal{C}_{\text{pw}}(\mathbb{R})$, a function in $\mathcal{C}^1_{\text{pw}}(\mathbb{R})$ need not be defined on all of $\mathbb{R}$. Additionally, and unfortunately, a function in $\mathcal{C}^1_{\text{pw}}(\mathbb{R})$ need not be differentiable on all of its domain—thus the domains of f and f' as defined above may be different! It is probably best not to think too hard about this and to keep the picture above foremost in mind.

Now, ponderously, we can state the true Fourier inversion theorem.

29.9 Theorem (Fourier inversion). *Let $f \in L^1 \cap \mathcal{C}^1_{\text{pw}}(\mathbb{R})$. Then*

$$\frac{f(x^+) + f(x^-)}{2} = \frac{1}{\sqrt{2\pi}} \lim_{R \to \infty} \int_{-R}^{R} \widehat{f}(k) e^{ikx} \, dk.$$

As with Theorem 28.14, the proof of this theorem is more analysis-heavy than appropriate for our class. In practice, we will not use it all that much—remember that we want to use the Fourier transform to find candidate solutions for differential equations. If we assume that such solutions are nice enough that the ideal inversion formula (29.2) applies, then we have hope of checking it rigorously. (Anyway, even Theorem 29.9 is better than the full result for inverting the Laplace transform, which requires complex analysis and line integrals.)

Last, we give a name to the integral appearing in the ideal version (29.2) of Fourier inversion.

29.10 Definition. *Let $g \in L^1$. The* **INVERSE FOURIER TRANSFORM** *of g is*

$$\widecheck{g}(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k) e^{ikx} \, dk.$$

The inverse Fourier transform, therefore, is just the "reflection" $\widecheck{g}(x) = \widehat{g}(-x)$. Theorem 29.9 therefore says that if $f \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ with $\widehat{f} \in L^1$, then

$$f = \widecheck{\widehat{f}}.$$

At this point we might be disappointed and frustrated with Fourier inversion. The hypotheses are complicated and the results are awkward. One life lesson might be to move away from seeking *pointwise* equalities and more toward *averaging* (we already saw this with continuous dependence on initial conditions for the heat equation). Perhaps there is a "norm" $\|\cdot\|$ in which we might meaningfully measure function "size" and for which we would have

$$\|f - \widecheck{\widehat{f}}\| = 0$$

with fewer hypotheses on f? There is, but it requires more analysis and integration theory than we can develop here.

Instead, we return to applying the Fourier transform to differential equations. Let $a \in \mathbb{R}$ and $f \in L^1$. We will find a solution candidate for the ODE $y' = ay + f(x)$ via the Fourier transform. Of course we do not need the transform *at all* to solve this ODE—we have variation of parameters. The point here is to see more about how the Fourier transform interacts with differential equations and how some other, less obvious properties of the transform naturally arise.

As usual, we work backwards. Assume there exists $y \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ with $y' \in L^1$ such that $y' = ay + f(x)$. We take the transform of both sides of $y' = ay + f(x)$ to find

$$\widehat{y'}(k) = \widehat{ay + f}(k)$$

and then use the derivative property and linearity to have

$$ik\widehat{y}(k) = a\widehat{y}(k) + \widehat{f}(k).$$

This is the algebraic equation that $\widehat{y}(k)$ must solve—no more derivatives in here.

We rearrange this to

$$(ik - a)\widehat{y}(k) = \widehat{f}(k).$$

If $a \neq 0$, then $\text{Re}(ik - a) = -a \neq 0$, so $ik - a \neq 0$ for all $k \in \mathbb{C}$. Then we can solve for $\widehat{y}(k)$:

$$\widehat{y}(k) = \frac{\widehat{f}(k)}{ik - a}.$$

29.11 Problem. If $a = 0$, then the division strategy above fails. Show that if $y \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ with $y' \in L$ and $y' = f(x)$ (it is redundant here to specify $f \in L^1$), then $\widehat{f}(0) = 0$. This is a "solvability condition" on f: under the assumptions on y, we cannot solve $y' = f(x)$ for all $f \in L^1$.

At this point we might use the inversion theorem to suggest that

$$y(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{y}(k)e^{ikx}\ dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\widehat{f}(k)}{ik - a} e^{ikx}\ dk.$$

This is essentially what we did with the heat equation, and then we would study this integral intensely. However, the integral above looks *nothing* like any result from variation of parameters, which suggests that we can do better.

One idea is to look more carefully at the integrand. Maybe we can find $g \in L^1$ such that

$$\widehat{g}(k) = \frac{1}{ik - a},$$

and then the integral would be

$$\int_{-\infty}^{\infty} \widehat{f}(k)\widehat{g}(k)e^{ikx}\ dk = \sqrt{2\pi}\,\widecheck{\widehat{f}\widehat{g}}(x).$$

Maybe there is something more profound that we can say about $\widecheck{\widehat{f}\widehat{g}}(x)$. How does the Fourier transform interact with products?

29.12 Problem. Explain why you should expect $\widehat{fg}(k) \neq \widehat{f}(k)\widehat{g}(k)$ in general.

Day 30: Monday, October 21.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 435–436 discuss convolution.

We first take up the task of finding $g \in L^1$ such that

$$\widehat{g}(k) = \frac{1}{ik - a}.$$

Our approach is wholly nonrigorous: we fool around. This equation is equivalent to

$$ik\widehat{g}(k) - a\widehat{g}(k) = 1,$$

and the left side looks like a derivative relation:

$$ik\widehat{g}(k) - a\widehat{g}(k) = \widehat{g'}(k) - \widehat{ag}(k) = \widehat{g' - ag}(k).$$

So, we would like

$$\widehat{g' - ag}(k) = 1. \tag{30.1}$$

Now it looks like we are assuming that $g \in \mathcal{C}^1_{\text{pw}}(\mathbb{R})$ with $g' \in L^1$, but, remember, we are only fooling around. The identity (30.1) reads

$$\int_{-\infty}^{\infty} (g'(x) - ag(x))e^{-ikx}\, dx = \sqrt{2\pi}.$$

As they say, "if it moves, differentiate it." Differentiate both sides with respect to k and pass the derivative through the integral (why not? we are just fooling around) to find

$$\int_{-\infty}^{\infty} x(g'(x) - ag(x))e^{-ikx}\, dx = 0. \tag{30.2}$$

Put $q(x) := x(g'(x) - ag(x))$. Then (30.2) says

$$\widehat{q}(k) = 0$$

for all k, and, notwithstanding our disappointments with the inverse transform, we want this to imply $q(x) = 0$ for all x. Thus, maybe,

$$x(g'(x) - ag(x)) = 0$$

for all x, so for $x \neq 0$,

$$g'(x) = ag(x).$$

Then $g(x) = g(0)e^{ax}$, right? But, for $g(0) \neq 0$, such g explode at $\pm\infty$ (depending on the sign of a), and so $g \notin L^1$. Maybe, however, the elimination of $x = 0$ from consideration above gives us an idea. We do not need $g \in \mathcal{C}^1(\mathbb{R})$ to have $g \in L^1$. What if we "break" the exponential where it starts to explode?

For simplicity, assume from now on that $a > 0$. Put

$$\mathsf{E}_a^+(x) := \begin{cases} e^{ax}, & x < 0 \\ 0, & x \geq 0. \end{cases}$$

Then $\mathsf{E}_a^+ \in \mathcal{C}_{\text{pw}}^1(\mathbb{R})$ and in fact E_a^+ is differentiable everywhere except at $x = 0$ with

$$(\mathsf{E}_a^+)'(x) = \begin{cases} ae^{ax}, & x < 0 \\ 0, & x > 0 \end{cases} = a\mathsf{E}_a^+(x).$$

So, E_a^+ seems to be doing what we expect from our very dodgy formal manipulations above.

Now we check that E_a^+ actually does what we want.

30.1 Problem. Using only the definitions of L^1 and the Fourier transform, show that $\mathsf{E}_a^+ \in L^1$ and

$$\widehat{\mathsf{E}_a^+}(x) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{a - ik}\right).$$

It then follows from our previous work that if $y' = ay + f(x)$ with y, y', $f \in L^1$, then

$$\widehat{y}(k) = \frac{\widehat{f}(k)}{ik - a} = -\sqrt{2\pi}\widehat{f}(k)\widehat{\mathsf{E}_a^+}(k).$$

So, what is going on with the product of those transforms?

Start with f, $g \in L^1$. Then

$$2\pi\widehat{f}(k)\widehat{g}(k) = \left(\int_{-\infty}^{\infty} f(x)e^{-ikx}\, dx\right)\left(\int_{-\infty}^{\infty} g(y)e^{-iky}\, dy\right).$$

We have collected those insistent, persistent factors of $\sqrt{2\pi}$ on the left, and we have written different variables of integration for the different integrals—this is just good mathematical grammar. Since the integrals are numbers, we can move one inside the other by linearity:

$$2\pi\widehat{f}(k)\widehat{g}(k) = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(x)e^{-ikx}\, dx\right) g(y)e^{-iky}\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)g(y)e^{-ik(x+y)}\, dx\, dy.$$

We want to see a single Fourier transform pop out of all this. It would be nice if the exponential did not depend on both x and y, so we substitute $s = x + y$ with $x = s - y$ and $ds = dx$ in the inner integral to find

$$\int_{-\infty}^{\infty} f(x)g(y)e^{-ik(x+y)}\, dx = \int_{-\infty}^{\infty} f(s - y)g(y)e^{-iks}\, ds.$$

Then

$$2\pi\widehat{f}(k)\widehat{g}(k) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s-y)g(y)e^{-iks}\ ds\ dy.$$

If we could get that e^{-iks} out by itself, the double integral would look more like a Fourier integral. We can accomplish this by interchanging the order of integration—a perilous Fubination for a doubly improper integral:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s-y)g(y)e^{-iks}\ ds\ dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s-y)g(y)e^{-iks}\ dy\ ds.$$

30.2 Theorem. *[Fubini, highly specialized] This works. More precisely let f, $g \in L^1$ and let $a \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ with a bounded (there is $M > 0$ such that $|a(s)| \le M$ for all $s \in \mathbb{R}$). Then*

$$\int_a^b\int_c^d f(x-y)g(y)h(x)\ dx\ dy = \int_c^d\int_a^b f(x-y)g(y)h(x)\ dy\ dx$$

for $-\infty \le a \le b \le \infty$ and $-\infty \le c \le d \le \infty$.

We take the version of Fubini above for granted and will not prove it. Thus

$$2\pi\widehat{f}(k)\widehat{g}(k) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s-y)g(y)e^{-iks}\ dy\ ds = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(s-y)g(y)\ dy\right)e^{-iks}\ ds.$$

The inner integral on the right has a special name.

30.3 Theorem. *Let $f \in L^1$ and $g \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ with g bounded. Then the* **CONVOLUTION** *of f and g,*

$$(f * g)(x) := \int_{-\infty}^{\infty} f(x-y)g(y)\ dy,$$

converges for each $x \in \mathbb{R}$.

30.4 Problem. Prove this theorem with the comparison test. Then show that $f*g = g*f$ by substituting.

In our toy problem, we have assumed $f \in L^1$ from the start, and now we are working with $g = \mathsf{E}_a^+$, which is definitely bounded, so the boundedness hypothesis in the definition of convolution does not restrict us here. We have shown, more generally, that (after changing variables from s back to x)

$$2\pi\widehat{f}(k)\widehat{g}(k) = \int_{-\infty}^{\infty}(f*g)(x)e^{-ikx}\ dx,$$

but that seems to be presuming $f * g \in L^1$, no?

30.5 Theorem. *No. (Yes?) Let f, $g \in L^1$ and suppose that at least one of these functions is bounded. Then $f * g \in L^1$ and*

$$\|f * g\|_{L^1} \le \|f\|_{L^1} \|g\|_{L^1}.$$

30.6 Problem. Prove this theorem as follows. First, the hypotheses guarantee that $f * g$ is defined on $\mathbb{R}$. Next, fix $R > 0$. Use the triangle inequality and Fubini's theorem and fill in the details of the following estimates:

$$\int_{-R}^{R} |(f * g)(x)| \, dx \le \int_{-\infty}^{\infty} \int_{-R}^{R} |f(x-y)||g(y)| \, dx \, dy \le \|f\|_{L^1} \|g\|_{L^1}.$$

Then apply part (iii) of Problem 21.12.

Embiggened by this result, we have

$$\widehat{f}(k)\widehat{g}(k) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (f * g)(x) e^{-ikx} \, dx \right) = \frac{\widehat{f * g}(k)}{\sqrt{2\pi}}.$$

This is the answer to our question of how the Fourier transform interacts with products.

30.7 Theorem. *Let f, $g \in L^1$ with at least one of these functions bounded. Then*

$$\widehat{f}(k)\widehat{g}(k) = \frac{\widehat{f * g}(k)}{\sqrt{2\pi}}.$$

Now we use the Fourier transform, once and for all, and faster, on our toy problem $y' = ay + f(x)$. Once more, we are assuming that $y \in \mathcal{C}^1(\mathbb{R})$ actually solves this ODE (not a big deal, since we do know how to solve it!) with y, $y' \in L^1$. Then

$$\widehat{y}(k) = -\sqrt{2\pi}\widehat{f}(k)\widehat{\mathsf{E}_a^+}(k) = -\sqrt{2\pi} \left(\frac{1}{\sqrt{2\pi}} \right) \widehat{f * \mathsf{E}_a^+}(k) = -\widehat{f * \mathsf{E}_a^+}(k).$$

This strongly suggests that a solution candidate should be

$$y(x) = -(f * \mathsf{E}_a^+)(x).$$

We work on this convolution:

$$(f * \mathsf{E}_a^+)(x) = \int_{-\infty}^{\infty} f(x-y)\mathsf{E}_a^+(y) \, dy = \int_{-\infty}^{0} f(x-y)e^{ay} \, dy = -\int_{\infty}^{x} f(s)e^{a(x-s)} \, ds$$

after substituting $s = x - y$ with $y = x - s$ and $ds = -dy$. Thus

$$(f * \mathsf{E}_a^+)(x) = e^{ax} \int_{x}^{\infty} e^{-as} f(s) \, ds,$$

and so the solution to $y' = ay + f(x)$ should be

$$y(x) = -e^{ax} \int_x^{\infty} e^{-as} f(s)\, ds. \tag{30.3}$$

Actually checking this is not too hard with the product rule and FTC1, which still works for improper integrals.

More interesting might be the question of why the Fourier transform (laboriously) only gave us *one* solution to $y' = ay + f(x)$. We never specified an initial condition, so we might be expecting an arbitrary constant somewhere in the solution. Where is it? To use the Fourier transform, we presumed $y \in \mathcal{C}^1(\mathbb{R})$ with y, $y' \in L^1$. That forces $\lim_{x \to \pm\infty} y(x) = 0$. The long-ago Example 4.1 told us that the only solution to $y' = ay + f(x)$ with $\lim_{x \to \infty} y(x) = 0$ has exactly this form (30.3). *The Fourier transform will not necessarily yield all solutions to a differential equation because legitimate application of the transform presumes decay at spatial infinity.*

30.8 Problem. Redo the Fourier analysis of $y' = ay + f(x)$ assuming $a < 0$ and using the function

$$\mathsf{E}_a^-(x) := \begin{cases} 0, & x < 0 \\ e^{ax}, & x \geq 0. \end{cases}$$

Use the now-developed machinery of convolutions to streamline your response—no need for all of the asides as developed above. [Hint: *you will need a formula for* $\widehat{\mathsf{E}_a^-}(k)$ *when* $a < 0$. *Can you possibly relate this to the existing formula for* $\widehat{\mathsf{E}_{-a}^+}(k)$*?*]

Day 31: Wednesday, October 23.

Here is an abstract situation that arises in practice with the heat equation. We need one more function space.

31.1 Definition. *A function* $f \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ *is* **BOUNDED** *if there exists* $C > 0$ *such that* $|f(x)| \leq C$ *for all* x *in the domain of* f. *The set of all bounded functions is* L^{∞}.

31.2 Problem. Prove that L^{∞} is a vector space in the sense that $f + g \in L^{\infty}$ and $\alpha f \in L^{\infty}$ for all f, $g \in L^{\infty}$ and $\alpha \in \mathbb{C}$.

31.3 Problem. Give an example of $f \in L^1$ such that $f \notin L^{\infty}$. [Hint: *f can be mostly* 0, *except where it is not* 0.]

We would like to put a norm on L^{∞} by

$$\|f\|_{\infty} := \max_{x \in \mathbb{R}} |f(x)|.$$

The problem is that a (piecewise) continuous function may be bounded and yet never attain a maximum on $\mathbb{R}$; think of arctan or $f(x) = 1 - \operatorname{sech}(x) = 1 - 2/(e^x + e^{-x})$. The real norm on L^∞ comes from replacing max by sup, which requires just a little too much real analysis for this course.

Back to Fourier business. Let m, $f \in L^1$ with at least one of m, f in L^∞. Then the convolution $g := m * f$ is defined and $g \in L^1$. Moreover, $\widehat{g} = \widehat{m}\widehat{f}/\sqrt{2\pi}$. We know $\widehat{f} \in L^\infty$ by basic Fourier transform properties; suppose also that m is so nice that $\widehat{m} \in L^1$. Then by the comparison test $\widehat{m}\widehat{f} \in L^1$, and so $\widehat{g} \in L^1$. Suppose, somehow, that we know that $g \in \mathcal{C}^1(\mathbb{R})$. Then Fourier inversion shows $g = \check{\widehat{g}}$, equivalently, $g(x) = \widehat{\widehat{g}}(-x)$. Since $\widehat{g} \in L^1$, its transform $\widehat{\widehat{g}}$ vanishes at infinity. That is, we have shown

$$\lim_{x \to \pm\infty} (m * f)(x) = 0.$$

Here is how this applies to the heat equation. We know that if $f \in L^\infty$, then the only solution to

$$\begin{cases} u_t = u_{xx}, \ x \in \mathbb{R}, \ t > 0 \\ u(x, 0) = f(x), \ x \in \mathbb{R} \end{cases}$$

satisfies

$$u(x,t) = \int_{-\infty}^{\infty} H(x-y,t)f(y) \ dy, \qquad H(s,t) = \frac{e^{-s^2/4t}}{\sqrt{4\pi t}}, \ s \in \mathbb{R}, \ t > 0.$$

This looks like "convolution in space":

$$u(x,t) = (H(\cdot,t) * f)(x).$$

For $t > 0$ fixed, the function $H(\cdot,t)$ is superbly nice: we have $H(\cdot,t) \in L^1 \cap L^\infty$ and $\widehat{H}(\cdot,t) \in L^1$ since $\widehat{H}(\cdot,t)$ is "basically" another Gaussian like H. Also, differentiating under the integral (which we *still* need to check implies that $u(\cdot,t) \in \mathcal{C}^1(\mathbb{R})$. So, with $t > 0$ fixed and $m = H(\cdot,t)$, the work above tells us that $\lim_{x\to\pm\infty} u(x,t) = 0$. This resembles the result of Problem 24.6 with the limit now in space rather than time.

31.4 Problem. Let $f \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ with $\widehat{f} \in L^1$. Prove that $f \in L^\infty$ and find $C > 0$ such that $|f(x)| \le C$ for all $x \in \mathbb{R}$. [Hint: *Fourier inversion.*]

Day 32: Friday, October 25.

You took Exam 2.

Day 33: Monday, October 28.

We know well that

$$u(x,t) := (H(\cdot,t) * f)(x), \qquad H(s,t) := \frac{e^{-s^2/4t}}{\sqrt{4\pi t}}$$

solves the heat equation

$$\begin{cases} u_t = u_{xx}, & (x,t) \in \mathcal{D} \\ u(x,0) = f(x), \end{cases} \qquad \mathcal{D} := \{(x,t) \in \mathbb{R}^2 \mid x \in \mathbb{R},\ t > 0\}.$$

We have only required $f \in L^1 \cup L^\infty$ for the convolution $H(\cdot,t) * f$ to be defined for $t > 0$. Since we presume no differentiability of u at $t = 0$, we have not said anything about the requisite regularity of f. It turns out that is not too important.

33.1 Example. Let

$$f(x) := \begin{cases} 1, & |x| \le 1 \\ 0, & |x| > 1. \end{cases}$$

This f is decidedly discontinuous. But since $f(y) = 0$ for $|y| > 1$, we have

$$u(x,t) := (H(\cdot,t) * f)(x) = \int_{-1}^{1} H(x-y,t)\ dy = \frac{1}{\sqrt{4\pi t}} \int_{-1}^{1} e^{-(x-y)^2/4t}\ dy.$$

Now u is a very nice function; the integrand meets all of the hypotheses for Leibniz's rule for definite integrals, and, repeatedly differentiating under the integral, we find $u \in \mathcal{C}^\infty(\mathcal{D})$. That is, the solution u is vastly *smoother* than the initial data f.

The situation above is really a more general consequence of convolution. Suppose that $f \in L^1$ and $g \in L^\infty$, so $f * g$ is defined. Suppose further that $f \in \mathcal{C}^1(\mathbb{R})$ and $f' \in L^1$; then $f' * g$ is defined, and we would expect from differentiation under the integral that

$$(f*g)'(x) = \partial_x \left[\int_{-\infty}^{\infty} f(x-y)g(y)\ dy\right] = \int_{-\infty}^{\infty} \partial_x[f(x-y)g(y)]\ dy = \int_{-\infty}^{\infty} f'(x-y)g(y)\ dy$$
$$= (f' * g)(x).$$

Here is another reason why we would expect $(f*g)' = f' * g$. Suppose $f,\ g \in L^1$, $f \in \mathcal{C}^1(\mathbb{R})$, $f' \in L^1$, and $g \in L^\infty$. Then, again, both $f * g$ and $f' * g$ are defined, and both are in L^1 with

$$\widehat{f' * g}(k) = \sqrt{2\pi}\widehat{f'}(k)\widehat{g}(k) = \sqrt{2\pi} ik\widehat{f}(k)\widehat{g}(k) = ik\big(\sqrt{2\pi}\widehat{f}(k)\widehat{g}(k)\big) = ik\widehat{f*g}(k) = \widehat{(f*g)'}(k).$$

This is not a proof that $f*g$ is differentiable, but the putative equality $\widehat{f' * g} = \widehat{(f*g)'}$ does lead us to expect $(f*g)' = f' * g$.

It happens that we have omitted an argument in our study of convolutions: to have $f * g \in L^1$, we need $f * g \in \mathcal{C}_{\text{pw}}(\mathbb{R})$. We never proved this, but we can get better: if $f \in L^1$ and $g \in L^\infty$, then $f * g \in \mathcal{C}(\mathbb{R})$. That is, convolution of piecewise continuous functions is genuinely continuous.

33.2 Theorem. *If $f \in L^1$ and $g \in L^\infty$, then $f * g \in \mathcal{C}(\mathbb{R})$.*

Proof. Fix $x \in \mathbb{R}$ It suffices to show

$$\lim_{h\to 0}(f * g)(x+h) = (f * g)(x).$$

We consider the difference

$$\begin{aligned}(f * g)(x+h) - (f * g)(x) &= \int_{-\infty}^{\infty} f(x+h-y)g(y)\ dy - \int_{-\infty}^{\infty} f(x-y)g(y)\ dy \\ &= \int_{-\infty}^{\infty} \big[f(x+h-y) - f(x-y)\big]g(y)\ dy.\end{aligned}$$

Then we estimate

$$|(f*g)(x+h)-(f*g)(x)| \le \int_{-\infty}^{\infty} \big|\big[f(x+h-y)-f(x-y)\big]g(y)\big|\ dy \le C\int_{-\infty}^{\infty} \big|f(x-y+h)-f(x-y)\big|\ dy,$$

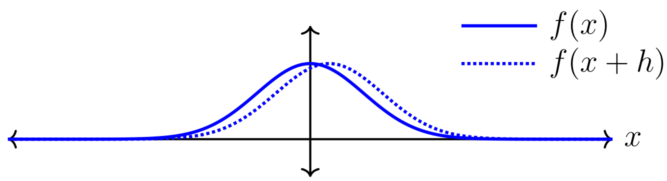
where $C > 0$ is such that $|g(y)| \le C$ for all $y \in \mathbb{R}$. Substitute $s = x - y$ above to find

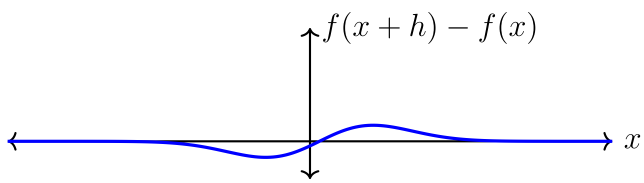
$$|(f*g)(x+h)-(f*g)(x)| \le C\int_{-\infty}^{\infty} |f(s+h)-f(s)|\ dy = C\left\|S^h f - f\right\|_{L^1}, \qquad (S^h f)(s) := f(s+h).$$

It turns out that

$$\lim_{h\to 0}\left\|S^h f - f\right\|_{L^1} = 0$$

for any $f \in L^1$. Here are some pictures to suggest why this is true.





The idea is that for h small, the graphs of f and $S^h f$ are so close that they have roughly the same area underneath. The closeness of the graphs of $S^h f$ and f "far away" from the origin in the first picture also indicates the proof strategy of showing $\lim_{h\to 0}\left\|S^h f - f\right\|_{L^1} = 0$: split the integral into three integrals over $(-\infty, -M]$, $[-M, M]$, and $[M, \infty)$, where $M > 0$ is large enough so that the integrals over $(-\infty, -M]$ and $[M, \infty)$ are very small, and then exploit the piecewise continuity of f on $[-M, M]$ to take h small enough that the integral over $[-M, M]$ is very small. ■

33.3 Problem. For $r > 0$, let

$$\chi(x,r) := \begin{cases} 1, & |x| \le r \\ 0, & |x| > r. \end{cases} \tag{33.1}$$

Let $f \in L^1$. Explain why $(f * \chi(\cdot, r))(x)/2r$ is the average value of f on $[x-r, x+r]$, and so we can view convolution as (up to a constant factor) an "averaging" operation, in addition to sometimes a "smoothing" one.

We do one more application of the Fourier transform to solving PDE and rederive D'Alembert's formula

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

for the wave equation

$$\begin{cases} u_{tt} = u_{xx}, & x, \ t \in \mathbb{R} \\ u(x,0) = f(x), & x \in \mathbb{R} \\ u_t(x,0) = g(x), & x \in \mathbb{R}. \end{cases}$$

Of course, the goal is not to see another derivation of D'Alembert's formula so much as to see the Fourier transform in action.

Taking the Fourier transform, we arrive at the ODE for $\widehat{u}(k, \cdot)$:

$$\begin{cases} \partial_t^2[\widehat{u}](k,t) + k^2\widehat{u}(k,t) = 0 \\ \widehat{u}(k,0) = \widehat{f}(k) \\ \partial_t[\widehat{u}](k,0) = \widehat{g}(k). \end{cases}$$

This is a more notationally complicated version of the problem

$$\begin{cases} y'' + k^2 y = 0 \\ y(0) = y_0 \\ y'(0) = y_1, \end{cases}$$

and the solution to that is

$$y(t) = y_0\cos(kt) + y_1 t\,\mathrm{sinc}(kt), \qquad \mathrm{sinc}(\tau) := \begin{cases} \sin(\tau)/\tau, & \tau \ne 0 \\ 1, & \tau = 0. \end{cases}$$

33.4 Problem. Check that. This is a little different from the similar-looking result in Problem 5.5, as here we are not necessarily assuming $k \ne 0$.

With that ODE result, we have

$$\widehat{u}(k,t) = \widehat{f}(k)\cos(kt) + \widehat{g}(k)t\,\mathrm{sinc}(kt),$$

and so Fourier inversion suggests the candidate solution

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{u}(k,t)e^{ikx}\,dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)\cos(kt)e^{ikx}\,dk + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k)t\,\mathrm{sinc}(kt)e^{ikx}\,dk.$$

We examine each term on the right separately.

First, we use the complex form of the cosine,

$$\cos(\tau) = \frac{e^{i\tau} + e^{-i\tau}}{2},$$

to express

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)\cos(kt)e^{ikx}\,dk = \frac{1}{2}\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)e^{ik(x+t)}\,dk\right) + \frac{1}{2}\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)e^{ik(x-t)}\,dk\right),$$

and then Fourier inversion gives

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)\cos(kt)e^{ikx}\,dk = \frac{f(x+t) + f(x-t)}{2}.$$

This is the first term in Duhamel's formula, so we are on the right track.

The second term in our candidate solution is

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k)t\,\mathrm{sinc}(kt)e^{ikx}\,dk = \frac{1}{2}\left[\int_{-\infty}^{\infty} \widehat{g}(k)\left(\frac{2t\,\mathrm{sinc}(kt)}{\sqrt{2\pi}}\right)e^{ikx}\,dk\right] = \frac{1}{2}\int_{-\infty}^{\infty} \widehat{g}(k)\widehat{\chi}(k,t)e^{ikx}\,dk,$$

where χ was defined in (33.1) and its Fourier transform was calculated in Example 28.11. That is,

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k)t\,\mathrm{sinc}(kt)e^{ikx}\,dk = \frac{\sqrt{2\pi}}{2}\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k)\widehat{\chi}(k,t)e^{ikx}\,dk\right) = \frac{\sqrt{2\pi}}{2}\widecheck{\widehat{g}\widehat{\chi}(\cdot,t)}(x).$$

Recall that for functions ϕ and ψ, we have the general relation

$$\widehat{\phi * \psi} = \sqrt{2\pi}\widehat{\phi}\widehat{\psi}$$

and so we expect

$$\phi * \psi = \widecheck{\widehat{\phi * \psi}} = \sqrt{2\pi}\widecheck{\widehat{\phi}\widehat{\psi}}.$$

Thus the second term in the candidate solution should be

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k)t\,\mathrm{sinc}(kt)e^{ikx}\,dk = \frac{(g * \chi(\cdot,t))(x)}{2}.$$

33.5 Problem. Use the definition of convolution and the definition of χ in (33.1) to show, as expected from D'Alembert's formula, that

$$\frac{(g * \chi(\cdot, t))(x)}{2} = \frac{1}{2}\int_{x-t}^{x+t} g(s)\ ds.$$

Day 34: Wednesday, October 30.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

The examples on pp. 683–684 suggest that we cannot naively differentiate under the integral all the time. Pages 684–687 discuss dominated convergence; this is wholly optional. Pages 687–689 discuss differentiating under the integral—note the remark at the top of p. 689 on piecewise continuity. Pages 339–340 provide a concise historical overview of Laplace's equation. Pages 342–343 discuss steady-state solutions to heat and wave.

The time has come to pay our dues and think seriously about differentiating under the integral. For a long time, we have blithely assumed that we can do this, and when we could, it proved things like solutions to the heat equation given by convolution, and smoothing of initial data. Why does this work?

Here is the broad situation. We have an interval $I \subseteq \mathbb{R}$, and we put

$$\mathcal{D} := \left\{(h, \tau) \in \mathbb{R}^2 \ \middle|\ h \in I,\ \tau \in \mathbb{R}\right\}.$$

And we have a function $q \colon \mathcal{D} \to \mathbb{C}$. The unusual choice of variables here will pay off shortly. For some $h_0 \in I$, we want to say

$$\lim_{h \to h_0} \int_{-\infty}^{\infty} q(h, \tau)\ d\tau = \int_{-\infty}^{\infty} \lim_{h \to h_0} q(h, \tau)\ d\tau. \tag{34.1}$$

What do we need first for this equality to make sense and second for it to be true?

Of course, the integrals must be defined and the limit on the right must exist. So, we want $q(h, \cdot) \in L^1$ for all $h \in I$, and we want the limit

$$Q(\tau) := \lim_{h \to h_0} q(h, \tau)$$

to exist for all $\tau \in \mathbb{R}$, and we want $Q \in L^1$. Is this enough to guarantee that the limit on the left exists and the equality (34.1) is true?

34.1 Problem. No. Let

$$q(h, \tau) := \begin{cases} 1, \ \dfrac{1}{|h|} \le \tau < \dfrac{2}{|h|}, \ h \neq 0, \ \tau \in \mathbb{R} \\ \\ 0, \ \text{otherwise}. \end{cases}$$

(i) To get a feel for q, first graph $q(1, \cdot)$, $q(1/2, \cdot)$, and $q(1/4, \cdot)$.

(ii) Show $\lim_{h \to 0} q(h, \tau) = 0$ for any $\tau \in \mathbb{R}$. [Hint: *what happens if* $|h|$ *is small enough that* $\tau < 1/|h|$*?*]

(iii) Compute $\int_{-\infty}^{\infty} q(h, \tau) \, d\tau$ and show that $\lim_{h \to 0} \int_{-\infty}^{\infty} q(h, \tau) \, d\tau = \infty$.

(iv) Explain why there exists no $M \in L^1$ such that $|q(h, \tau)| \le |M(\tau)|$ for all $(h, \tau) \in \mathbb{R}^2$. [Hint: *given* $\tau > 0$*, compute* $q(1/\tau, \tau)$ *and conclude* $|M(\tau)| \ge 1$ *when* $\tau > 0$*. Alternatively, use the previous part to get a lower bound on* $\|M\|_{L^1}$ *in terms of* $|h|$*; what happens as* $h \to 0$*?*]

34.2 Problem. And no. Let

$$q(h, \tau) := \begin{cases} 1, & \dfrac{1}{|h|} \le \tau < \dfrac{1}{|h|} + 1, \ h \ne 0, \ \tau \in \mathbb{R} \\ \\ 0, & \text{otherwise.} \end{cases}$$

(i) To get a feel for q, first graph $q(1, \cdot)$, $q(1/2, \cdot)$, and $q(1/4, \cdot)$.

(ii) Show $\lim_{h \to 0} q(h, \tau) = 0$ for any $\tau \in \mathbb{R}$. [Hint: *again, what happens if* $|h|$ *is small enough that* $\tau < 1/|h|$*?*]

(iii) Compute $\int_{-\infty}^{\infty} q(h, \tau) \, d\tau$ and show that $\lim_{h \to 0} \int_{-\infty}^{\infty} q(h, \tau) \, d\tau = 1$.

(iv) Explain why there exists no $M \in L^1$ such that $|q(h, \tau)| \le |M(\tau)|$ for all $(h, \tau) \in \mathbb{R}^2$. [Hint: *again, given* $\tau > 0$*, compute* $q(1/\tau, \tau)$ *and conclude* $|M(\tau)| \ge 1$ *when* $\tau > 0$*.*]

We need an extra hypothesis on q, which might feel vaguely reminiscent of the "trapping" or "dominating" behavior of the comparison test. The right thing to assume is the existence of $M \in L^1$ such that $|q(h, \tau)| \le |M(\tau)|$ for all $(h, \tau) \in \mathcal{D}$. (Indeed, assuming this makes the hypothesis $q(h, \cdot) \in L^1$ unnecessary by the comparison test!)

Here is the technical result.

34.3 Theorem (Dominated convergence). *Let* $I \subseteq \mathbb{R}$ *be an interval, let* $\mathcal{D} := \{(h, \tau) \in \mathbb{R}^2 \mid h \in I, \ \tau \in \mathbb{R}\}$*, let* $h_0 \in I$*, and let* $q \colon \mathcal{D} \to \mathbb{C}$ *be a function with the following properties.*

(i) $q(h, \cdot) \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ *for each* $h \in I$*.*

(ii) *The limit* $Q(\tau) := \lim_{h \to h_0} q(h, \tau)$ *exists for all* $\tau \in \mathbb{R}$*.*

(iii) $Q \in \mathcal{C}_{\text{pw}}(\mathbb{R})$*.*

(iv) *There is* $M \in L^1$ *such that* $|q(h, \tau)| \le |M(\tau)|$ *for all* $(h, \tau) \in \mathcal{D}$*.*

Then

$$\lim_{h \to h_0} \int_{-\infty}^{\infty} q(h, \tau)\, d\tau = \int_{-\infty}^{\infty} \lim_{h \to h_0} Q(\tau)\, d\tau.$$

34.4 Problem. **(i)** Let $f \in L^1$ and $g \in L^\infty$. Theorem 33.2 tells us that $f * g$ is continuous. Why does dominated convergence not prove this? [Hint: *f and g may only be piecewise continuous.*]

(ii) We previously claimed that with $H(s,t) = e^{-s^2/4t}/\sqrt{4\pi t}$ and $f \in L^1 \cup L^\infty$, we had $\lim_{t \to 0^+}(H(\cdot, t) * f)(x) = f(x)$. Why does dominated convergence not prove this? (By the way, Theorem 34.3 still works if h_0 is an endpoint of I and the limit $\lim_{h \to h_0}$ is replaced with $\lim_{h \to h_0^+}$ or $\lim_{h \to h_0^-}$, so say something more profound than "It's the wrong kind of limit.")

We now apply this to give precise hypotheses on when differentiating under an *improper* integral is valid. Now we take an interval $I \subseteq \mathbb{R}$ and, changing notation slightly, put $\mathcal{D} := \{(s, \tau) \in \mathbb{R}^2 \mid s \in I,\ \tau \in \mathbb{R}\}$. Let $f \colon \mathcal{D} \to \mathbb{C}$, $f = f(s, \tau)$, be a function; we want

$$\partial_s \left[\int_{-\infty}^{\infty} f(s, \tau)\, d\tau\right] = \int_{-\infty}^{\infty} f_s(s, \tau)\, d\tau.$$

For each side even to make sense, this demands that f_s exists on $\mathcal{D}$, that $f(s, \cdot) \in L^1$ for all $s \in I$, and that $f_s(s, \cdot) \in L^1$. This equality is really saying

$$\lim_{h \to 0} \int_{-\infty}^{\infty} \frac{f(s+h, \tau) - f(s, \tau)}{h}\, d\tau = \int_{-\infty}^{\infty} \lim_{h \to 0} \frac{f(s+h, \tau) - f(s, \tau)}{h}\, d\tau$$

for all $s \in I$ (where the hypothesis $f(s, \cdot) \in L^1$ for all $s \in I$ ensures that the integrals are defined). So, for $s \in I$ fixed, we want to interchange the limit as $h \to 0$ with

$$q(h, \tau) := \begin{cases} [f(s+h, \tau) - f(s, \tau)]/h, & h \neq 0,\ \tau \in \mathbb{R} \\ f_s(s, \tau), & h = 0,\ \tau \in \mathbb{R}. \end{cases}$$

We are thinking now of q defined for $h \approx 0$ (how close depending on s) and $\tau \in \mathbb{R}$.

We check the hypotheses of dominated convergence. Since $f(s, \cdot) \in L^1$ for all s, so also $f(s, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$, so $q(h, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$. Since f_s is defined on $\mathcal{D}$, the limit $f_s(s, \tau) = \lim_{h \to 0} q(h, \tau)$ exists for all τ, and since $f_s(s, \cdot) \in L^1$, we also have $f_s(s, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$. All that is missing is a dominating function, and we claim that we just need to dominate f_s.

That is, suppose there exists $M \in L^1$ such that $|f_s(s, \tau)| \le |M(\tau)|$ for all $(s, \tau) \in \mathcal{D}$. We claim that $|q(h, \tau)| \le |M(\tau)|$ for $h \approx 0$ and $\tau \in \mathbb{R}$, and we really only need to show this for $h \neq 0$. So, fix $\sigma \in I$, $\tau \in \mathbb{R}$, and $h \neq 0$ small relative to σ. Then

$$q(h, \tau) = \frac{f(\sigma + h, \tau) - f(\sigma, \tau)}{h} = \frac{1}{h} \int_{\sigma}^{\sigma + h} f_s(s, \tau)\, ds,$$

and so we estimate (if $h > 0$; if $h < 0$ we just flip the limits of integration)

$$|q(h, \tau)| \le \frac{1}{h} \int_\sigma^{\sigma+h} |f_s(s, \tau)| \, ds \le \frac{1}{h} \int_\sigma^{\sigma+h} |M(\tau)| \, ds = \frac{1}{h}(\sigma + h - \sigma)|M(\tau)| = |M(\tau)|.$$

We have therefore checked all of the hypotheses of dominated convergence.

Here is our streamlined result.

34.5 Theorem (Differentiating under the integral for improper integrals). *Let $I \subseteq \mathbb{R}$ be an interval, let $\mathcal{D} := \{(s, \tau) \in \mathbb{R}^2 \mid s \in I, \ \tau \in \mathbb{R}\}$, and let $f : \mathcal{D} \to \mathbb{C}$ be a function with the following properties.*

(i) *$f(s, \cdot) \in L^1$ for all $s \in I$.*

(ii) *f_s exists on $\mathcal{D}$.*

(iii) *$f_s(s, \cdot) \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ for all $s \in I$.*

(iv) *There exists $M \in L^1$ such that $|f_s(s, \tau)| \le |M(\tau)|$ for all $(s, \tau) \in \mathcal{D}$.*

Then

$$\partial_s \left[\int_{-\infty}^\infty f(s, \tau) \, d\tau \right] = \int_{-\infty}^\infty f_s(s, \tau) \, d\tau.$$

We apply this to check over differentiating a convolution. Recall that for $f * g$ to be defined, we need to assume either $f \in L^1$ and $g \in L^\infty$, or $f \in L^\infty$ and $g \in L^1$. If $f \in \mathcal{C}^1(\mathbb{R})$, we expect $(f * g)' = f' * g$, so we need to assume either $f' \in L^1$ and $g \in L^\infty$ or $f' \in L^\infty$ and $g \in L^1$. Put $F(x, g) := f(x - y)g(y)$, so $(f * g)(x) = \int_{-\infty}^\infty F(x, y) \, dy$. To differentiate under the integral, we need to dominate $|F_x(x, y)| \le |M(y)|$ for some $M \in L^1$. Since $F_x(x, y) = f'(x - y)g(y)$, we can take $M = Cg$ if we assume $g \in L^1$ and $f' \in L^\infty$ with $|f'(s)| \le C$ for all $s \in \mathbb{R}$.

34.6 Problem. This may feel a little less general than our initial possibility that $f * g$ is defined when $f \in L^1$ and $g \in L^\infty$. The pesky problem is that f' is evaluated at $x - y$; if $g \in L^\infty$, there is really no way to guarantee $|f'(x - y)| \le |M(y)|$ for all $x, y \in \mathbb{R}$ with $M \in L^1$. In fact, show that if $|f'(x-y)| \le |M(y)|$ for all $x, y \in \mathbb{R}$, then $|f'(s)| \le |M(t)|$ for all $s, t \in \mathbb{R}$. This would be a *very* strong domination condition on f', one that is probably too hard to check in practice.

Here is the technical result on differentiating convolution.

34.7 Theorem. *Let $f \in L^\infty \cap \mathcal{C}^1(\mathbb{R})$ with $f' \in L^\infty$, and let $g \in L^1$. Then $f * g \in \mathcal{C}^1(\mathbb{R})$ with $(f * g)' = f' * g$.*

We might call this the "smoothing" property of convolution: g can be a very "bad" function—possibly nowhere differentiable (such functions do exist!) but still piecewise continuous—

and yet, if f is differentiable, then so is $f * g$.

34.8 Problem. The discussion above proves almost all of Theorem 34.7. Under the hypotheses of that theorem, explain why $f' * g \in \mathcal{C}(\mathbb{R})$. [Hint: *Theorem 33.2.*]

34.9 Problem. Let $\mathcal{D} := \{(s,t) \in \mathbb{R}^2 \mid s \in \mathbb{R},\ t > 0\}$. With $H(s,t) = e^{-s^2/4t}/\sqrt{4\pi t}$ defined on $\mathcal{D}$, set $u(x,t) := (H(\cdot,t) * f)(x)$ with $f \in L^1$. Prove that $u \in \mathcal{C}^\infty(\mathcal{D})$ by checking that the hypotheses of Theorem 34.7 are met for the x-derivatives and by using Theorem 34.5 for the t-derivatives. [Hint: *to make things easier for the t-derivatives, ignore the factor of $1/\sqrt{4\pi t}$, since that can factor out of the integral; to get a good bound on the integrand, prove differentiability for $t \in [1/n, \infty)$ with $n \geq 1$.*]

34.10 Problem. What additional hypothesis on $f \in L^1$ guarantees $\widehat{f} \in \mathcal{C}^1(\mathbb{R})$?

We now proceed to take up the last of the major linear PDE of the course: Laplace's equation. Recall that transport, wave, and heat read

$$u_t + u_x = 0, \qquad u_{tt} - u_{xx} = 0, \qquad \text{and} \qquad u_t - u_{xx} = 0.$$

Transport and wave are, in a certain sense, symmetric in x and t; interchanging the order of the variables ($u(x,t)$ or $u(t,x)$) really results in an equivalent problem. That is certainly not the case in heat, due to the imbalanced derivatives. But wave and heat are alike in that one variable is "privileged" over the other: one variable comes with a negative sign, and so there is a moral distinction between the time and the space variables.

Laplace's equation is unlike these preceding three problems. Written incorrectly, it is

$$u_{xx} + u_{tt} = 0,$$

so there is still a transport/wave-type symmetry in the variables, but neither is "privileged" over the other with a negative sign. Written correctly (as no one uses time for a variable in Laplace's equation), it is

$$u_{xx} + u_{yy} = 0,$$

and we think of both x and y as spatial coordinates in $\mathbb{R}^2$. Thus Laplace's equation will expose us to some more geometric and topological tools than its predecessor.

Here is one reason to care about Laplace's equation. The two-dimensional generalizations of the wave and heat equations (think of a surface vibrating or a stovetop being heated) are

$$u_{tt} = u_{xx} + u_{yy} \qquad \text{and} \qquad u_t = u_{xx} + u_{yy},$$

respectively. Perhaps one is interested in **STEADY-STATE SOLUTIONS** u, which are independent of time: $u_t = 0$ and thus $u_{tt} = 0$. Such solutions depend only on spatial behavior, and so the 2D steady-state wave and heat equations both reduce to Laplace's equation: $u_{xx} + u_{yy} = 0$.

We call the operator

$$\Delta u := u_{xx} + u_{yy}$$

the **Laplacian**, and if $\Delta u = 0$, then we say that u is **harmonic**. While we will find some concrete solutions to Laplace's equation, meaningful solutions often depend in a complicated way on the geometry of the underlying domain of the desired solution, and we will not have as neat formulas as we do for transport and wave, or even heat. Instead, our interest will be in learning about *properties* of harmonic functions, and especially how those properties require us to learn more about the geometry and topology of $\mathbb{R}^2$.

Day 35: Friday, November 1.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 473–474 solve Laplace's equation in the upper half-plane via the Fourier transform. Problem 7 on p. 394 discusses some of the challenges in differentiating the candidate solution. Example 1 on p. 341 presents the radial solution to Laplace's equation for $n = 2$; Example 1 on pp. 24–25 does it for $n = 3$; and the remarks on p. 26 generalize this to n dimensions.

As a first example, we solve a version of Laplace's equation with the Fourier transform. We do this really for more practice with Fourier transforms and some ODE concepts than for any deep insight into Laplace's equation.

35.1 Example. Consider the problem

$$\begin{cases} \Delta u = 0, \ x \in \mathbb{R}, \ y > 0 \\ u(x,0) = f(x), \ x \in \mathbb{R}, \ y = 0. \end{cases} \tag{35.1}$$

This asks us to solve Laplace's equation on the **upper half-plane** $\{(x,y) \in \mathbb{R}^2 \mid y > 0\}$ with "boundary" data specified at $y = 0$ via f. Since the spatial variable x can be any number in $\mathbb{R}$, we take the Fourier transform in x and obtain an ODE in y for the transform:

$$\begin{cases} \partial_y^2[\widehat{u}](k,y) - k^2\widehat{u}(k,y) = 0 \\ \widehat{u}(k,0) = \widehat{f}(k). \end{cases}$$

As usual, this might look more familiar as an ODE if we recast it in different notation:

$$\begin{cases} z''(y) - k^2 z(y) = 0 \\ z(0) = \widehat{f}(k). \end{cases}$$

We will only solve this for $k \neq 0$, as when we formally invert the transform to get a candidate solution, the Fourier integral will not care about the value of $\widehat{u}(k,y)$ at $k = 0$. At the very least, then, for $k \neq 0$ we can say $z(y) = c_1(k)e^{ky} + c_2(k)e^{-ky}$ for some constants $c_(k)$ and $c_2(k)$; here we are thinking of k as a parameter in the problem. The challenge is that we have only one initial condition, but this is a second-order ODE. Nothing is specifying $z'(0)$.

However, we do have some extra structure underlying our assumptions. Namely, we expect $\widehat{u}(\cdot, y) \in L^\infty$: for each $y > 0$, there should be $M(y) > 0$ such that if $k \in \mathbb{R}$, then $|\widehat{u}(k, y)| \le M(y)$. The problem is that the term $c_1(k)e^{ky}$ can be unbounded for $k > 0$, and likewise the term $c_2(k)e^{-ky}$ can be unbounded for $k < 0$, unless we have further control on $c_1(k)$ and $c_2(k)$. (We might try $c_1(k) = e^{-ky}$, but c_1 is not allowed to depend on y.)

The simplest solution is brute force: we choose c_1 and c_2 to "turn off" the problematic exponentials at the problematic values of k. That is, we take $c_1(k) = 0$ for $k > 0$ and $c_2(k) = 0$ for $k < 0$. Then

$$z(y) = c_1(k)e^{ky} + c_2(k)e^{-ky} = \begin{cases} c_2(k)e^{-ky}, & k > 0 \\ c_1(k)e^{ky}, & k < 0. \end{cases}$$

We can write this more compactly as

$$z(y) = \begin{cases} c_2(k)e^{-|k|y}, & k > 0 \\ c_1(k)e^{-|k|y}, & k < 0 \end{cases} = c(k)e^{-|k|y}, \ k \ne 0, \qquad c(k) := \begin{cases} c_2(k), & k > 0 \\ c_1(k), & k < 0. \end{cases}$$

Then we need $\widehat{f}(k) = z(0) = c(k)$.

Returning to the actual problem, we have deduced that if u solves Laplace's equation in the upper half-plane, then we should have $\widehat{u}(k, y) = \widehat{f}(k)e^{-|k|y}$. It would be nice to recognize the second factor as a transform, for the purposes of convolution, and, gloriously, we can: if

$$g(x, y) := \frac{1}{\sqrt{2\pi}} \frac{2y}{x^2 + y^2},$$

then $\widehat{g}(k, y) = e^{-|k|y}$. We discuss this more below. Thus

$$\widehat{u}(k, y) = \widehat{g}(k, y)\widehat{f}(k) = \sqrt{2\pi}\,\widehat{g(\cdot, y) * f}(k),$$

and so our candidate solution is

$$u(x, y) = \sqrt{2\pi}(g(\cdot, y) * f)(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{(x - s)^2 + y^2} f(s) \, ds,$$

once the dust settles from some algebra.

With the formal work done, we can show that u as defined above actually solves (35.1). Most of this is just careful work with differentiating under the integral, as we show below. It can also be shown that $\lim_{y \to 0^+} u(x, y) = f(x)$, which is morally similar to the result (24.1) for the heat equation (and technically just as involved).

35.2 Problem. Let $f \in L^1$. For $x \in \mathbb{R}$ and $y > 0$, put

$$u(x, y) := (L(\cdot, y) * f)(x), \qquad L(x, y) := \frac{y}{x^2 + y^2}.$$

(i) Use Theorem 34.7 to check that $u(\cdot, y)$ is twice differentiable for all $y > 0$.

(ii) Fix $x \in \mathbb{R}$. Let $n \geq 1$ be an integer and let $\mathcal{D}_n := \{(y, s) \in \mathbb{R}^2 \mid y \geq 1/n, \ s \in \mathbb{R}\}$. Use Theorem 34.5 to show that $u(x, \cdot)$ is twice differentiable on $\mathcal{D}_n$. (Note that s is now the variable of integration, so in that theorem switch the roles of s and τ.)

(iii) Actually calculate u_{xx} and u_{yy} and show $u_{xx} + u_{yy} = 0$.

Now we start to study more deeply the geometry of Laplace's equation.

35.3 Example. Since the variables x and y are "symmetric" in $u_{xx} + u_{yy} = 0$, we might look for a solution that is also "symmetric" in x and y; in particular, such a solution should satisfy $u(x, y) = u(y, x)$. Possibly the most symmetric structure in $\mathbb{R}^2$ is a circle centered at the origin, and so we look for solutions that are the same over circles. That is, $u(x, y)$ should only depend on the value of $\sqrt{x^2 + y^2}$, which is the distance from (x, y) to the origin. Thus we posit

$$u(x, y) = f(r(x, y)), \qquad r(x, y) := \sqrt{x^2 + y^2},$$

where $f \in \mathcal{C}^2((0, \infty))$, and we need to determine f. Since r is not differentiable at $(0, 0)$, we do not expect this u to be defined on or differentiable on all of $\mathbb{R}^2$.

We first compute

$$r_x(x, y) = \frac{1}{2\sqrt{x^2 + y^2}} 2x = x[r(x, y)]^{-1}$$

and

$$r_{xx}(x, y) = [r(x, y)]^{-1} - x[r(x, y)]^{-2} r_x(x, y) = [r(x, y)]^{-1} - x^2[r(x, y)]^{-3}.$$

Then

$$u_x(x, y) = f'(r(x, y)) r_x(x, y)$$

and so

$$\begin{aligned} u_{xx}(x, y) &= f''(r(x, y))[r_x(x, y)]^2 + f'(r(x, y)) r_{xx}(x, y) \\ &= f''(r(x, y)) x^2 [r(x, y)]^{-2} + f'(r(x, y)) \big([r(x, y)]^{-1} - x^2[r(x, y)]^{-3}\big). \end{aligned}$$

Of course, the y-partials are the same, with the factors of x replaced by y. Thus $u_{xx} + u_{yy} = 0$ if and only if

$$f''(r)(x^2 + y^2) r^{-2} + f'(r)\big(2r^{-1} - (x^2 + y^2) r^{-3}\big) = 0,$$

where we have abbreviated $r = r(x, y)$. Since $r^2 = x^2 + y^2$, this becomes

$$f''(r) + f'(r)(2r^{-1} - r^{-1}) = 0,$$

and so

$$f''(r) + r^{-1} f'(r) = 0.$$

Recall that we only want $f \in \mathcal{C}^2((0,\infty))$. We are saying that f must satisfy

$$f''(\sqrt{x^2+y^2}) + f'(\sqrt{x^2+y^2})/\sqrt{x^2+y^2} = 0$$

for all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$; taking $y = 0$ and $x > 0$ means

$$f''(x) + x^{-1}f'(x) = 0.$$

This is an ODE for f, and the analysis in this paragraph is morally the same as finding the ODE for a traveling wave profile (recall Example 13.1).

Rather than privilege x over y, we overwork our variables and take r to be the independent variable of f; thus we want to solve

$$f''(r) + r^{-1}f'(r) = 0.$$

This is a variable-coefficient second-order linear ODE—not the friendliest of beasts. However, if we put $g(r) = f'(r)$, then it becomes

$$g'(r) = -r^{-1}g(r),$$

which is a separable ODE; it solution (fixing the initial value at $r = 1$, not $r = 0$, since neither f nor g should be defined at $r = 0$) is

$$g(r) = g(1)\exp\left(\int_1^r -\rho^{-1}\,d\rho\right) = g(1)e^{-\ln(r)} = g(1)e^{\ln(r^{-1})} = g(1)r^{-1},$$

and so f must satisfy

$$f'(r) = g(1)r^{-1},$$

thus

$$f(r) = g(1)\ln(r) + f(1).$$

In more evocative notation,

$$u(x,y) := c_1\ln(\sqrt{x^2+y^2}) + c_2$$

solves Laplace's equation on $\mathbb{R}^2 \setminus \{(0,0)\}$ for any c_1, $c_2 \in \mathbb{R}$.

35.4 Problem. Find all functions $f \in \mathcal{C}^2((0,\infty))$ such that

$$u(x,y,z) := f(\sqrt{x^2+y^2+z^2})$$

solves $u_{xx} + u_{yy} + u_{zz} = 0$ on $\mathbb{R}^3 \setminus \{(0,0,0)\}$. What changes in the analysis from the preceding example does the inclusion of the variable z require?

35.5 Problem. Prove the following ROTATIONAL INVARIANCE of the Laplacian. Suppose that u is harmonic on $\mathbb{R}^2$. Let a, b, $\theta \in \mathbb{R}$ and define

$$U(X,Y) := u(\cos(\theta)X + \sin(\theta)Y + a, -\sin(\theta)X + \cos(\theta)Y + b).$$

This change of coordinates may be easier to visualize as

$$\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}.$$

This is a rotation by the angle θ coupled with a shift by the vector (a,b). Prove that $\Delta U = 0$. This is interpreted as the invariance of Δ under rotations and affine translations.

Day 36: Monday, November 4.

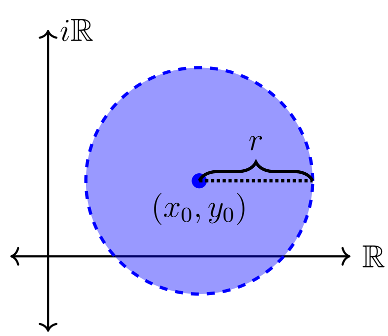
Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Page 385 discusses elementary topology of $\mathbb{R}^2$. Theorem 1 on pp. 386–387 presents the (weak) maximum principle with a proof different from ours in class (and the one in the notes here is a bit of a refinement of that proof from class, anyway)—the key difference is that the book proves it by contradiction. Theorem 4 on p. 391 is the (strong) maximum principle, and again the book's proof is different—it uses a supremum, whereas our proof used a covering lemma. Both tools require more knowledge of analysis beyond our course. Read Example 3 on p. 356 and Example 3 on p. 388.

Because of the two-dimensional geometry underling $\Delta u := u_{xx} + u_{yy} = 0$, there are somewhat broader domains appropriate for u with Laplace's equation than were available for transport, wave, and heat. We gradually introduce some elementary topology of $\mathbb{R}^2$ to manage those domains.

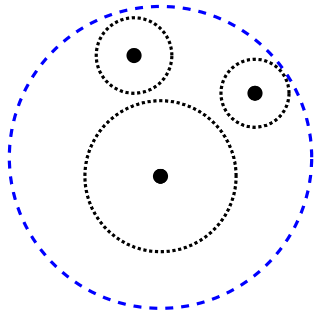
36.1 Definition. *The* OPEN BALL *of radius $r > 0$ centered at $(x_0, y_0) \in \mathbb{R}^2$ is*

$$\mathcal{B}((x_0,y_0);r) := \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; (x-x_0)^2 + (y-y_0)^2 < r^2\right\}.$$



Open balls are immensely useful for controlling the geometry of certain subsets of $\mathbb{R}^2$. In single-variable calculus, most of the interesting results occurred for functions defined on open or closed intervals. Here is the two-dimensional analogue of open intervals.

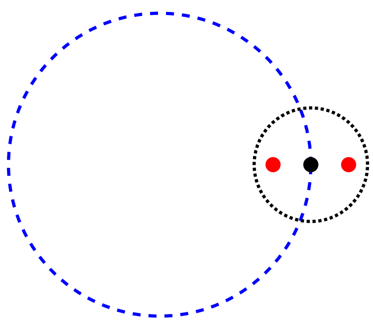
36.2 Definition. *A set $\mathcal{D} \subseteq \mathbb{R}^2$ is* **OPEN** *if for each $(x, y) \in \mathcal{D}$, there is $r > 0$ such that $\mathcal{B}((x, y); r) \subseteq \mathcal{D}$.*



In single-variable calculus we also need to keep track of the endpoints of open intervals: the points a and b function as the "boundary" of (a, b) in the sense that perturbing even slightly from a or b lands one either in (a, b) or in $\mathbb{R} \setminus (a, b)$.

36.3 Definition. *Let $\mathcal{D} \subseteq \mathbb{R}^2$. The* **BOUNDARY** *of $\mathcal{D}$ is the set of all points $(x, y) \in \mathbb{R}^2$ such that for every $r > 0$, both*

$$\mathcal{B}((x, y); r) \cap \mathcal{D} \neq \varnothing \quad \text{and} \quad \mathcal{B}((x, y); r) \cap (\mathbb{R}^2 \setminus \mathcal{D}) \neq \varnothing.$$



We denote the boundary of $\mathcal{D}$ by $\partial\mathcal{D}$.

The interval (a, b) is open, but when we include the boundary points and turn it into $[a, b]$, it becomes closed.

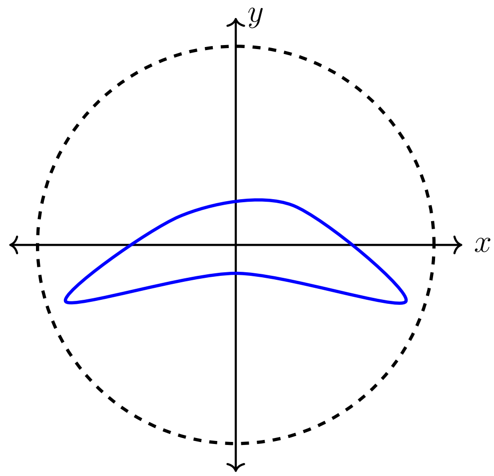
36.4 Definition. *The* **CLOSURE** *of $\mathcal{D} \subseteq \mathbb{R}^2$ is $\overline{\mathcal{D}} := \mathcal{D} \cup \partial\mathcal{D}$.*

The closure is a key component of the extreme value theorem, which we previously used for a very anodyne subset of $\mathbb{R}^2$ in the maximum principle for the heat equation. We review it more carefully here since Laplace's equation can involve much more complicated geometries.

First, recall that both kinds of intervals (a, b) and $[a, b]$ are (at least for a, $b \in \mathbb{R}$)

bounded—nothing can be all that large in those intervals.

36.5 Definition. *A set $\mathcal{D} \subseteq \mathbb{R}^2$ is* **BOUNDED** *if there exists $R > 0$ such that $\mathcal{D} \subseteq \mathcal{B}((0,0); R)$.*



Here are some comforting and useful results.

36.6 Problem. Let $\mathcal{D} \subseteq \mathbb{R}^2$.

(i) Prove that if $\mathcal{D}$ is open, then $\mathcal{D} \cap \partial\mathcal{D} = \varnothing$, and consequently $\mathcal{D} \neq \overline{\mathcal{D}}$.

(ii) Prove that $\mathcal{D}$ is bounded if and only if $\overline{\mathcal{D}}$ is bounded.

These results are comforting because we should expect them from our one-dimensional intuition. The first is particularly useful because it ensures that we can specify the behavior of a function defined on $\overline{\mathcal{D}}$ separately on $\mathcal{D}$ and $\partial\mathcal{D}$ and not have any conflicting overlap; in particular, we can pose the so-called **DIRICHLET PROBLEM** for Laplace's equation as

$$\begin{cases} \Delta u = 0 \text{ on } \mathcal{D} \\ u(x,y) = f(x,y) \text{ on } \partial\mathcal{D}, \end{cases}$$

for $\mathcal{D} \subseteq \mathbb{R}^2$ open, with the idea that $u \in \mathcal{C}(\overline{\mathcal{D}}) \cap \mathcal{C}^2(\mathcal{D})$, i.e., we require differentiability only on the "interior" $\mathcal{D}$.

36.7 Remark. *The notation $\mathcal{C}(\overline{\mathcal{D}}) \cap \mathcal{C}^2(\mathcal{D})$ is customary but awful. A function in $\mathcal{C}(\overline{\mathcal{D}})$ has domain $\overline{\mathcal{D}}$, whereas a function in $\mathcal{C}^2(\mathcal{D})$ has domain $\mathcal{D}$. For $\mathcal{D}$ open, we have $\mathcal{D} \neq \overline{\mathcal{D}}$. Since two functions with different domains cannot be equal, we really should have $\mathcal{C}(\overline{\mathcal{D}}) \cap \mathcal{C}^2(\mathcal{D}) = \varnothing$. What we mean by $u \in \mathcal{C}(\overline{\mathcal{D}}) \cap \mathcal{C}^2(\mathcal{D})$ is that $u \in \mathcal{C}(\overline{\mathcal{D}})$ and $u\big|_{\mathcal{D}} \in \mathcal{C}^2(\mathcal{D})$, where $u\big|_{\mathcal{D}}$ is the* **RESTRICTION** *$u\big|_{\mathcal{D}} \colon \mathcal{D} \to \mathbb{R} \colon (x,y) \mapsto u(x,y)$. Best not to think too hard about all this.*

Now we have all the machinery necessary for our version of the extreme value theorem.

36.8 Theorem (Extreme value). *Let $\mathcal{D} \subseteq \mathbb{R}$ be bounded and let $u \in \mathcal{C}(\overline{\mathcal{D}})$. Then there exist $(x_{\min}, y_{\min})$, $(x_{\max}, y_{\max}) \in \mathcal{D}$ such that*

$$u(x_{\min}, y_{\min}) = \min_{(x,y)\in\overline{\mathcal{D}}} u(x,y) \quad \text{and} \quad u(x_{\max}, y_{\max}) = \max_{(x,y)\in\overline{\mathcal{D}}} u(x,y). \tag{36.1}$$

This theorem does not tell us *where* the extreme values occur in $\overline{\mathcal{D}}$—perhaps in $\mathcal{D}$, perhaps in $\partial\mathcal{D}$. We can say much more for a harmonic function.

36.9 Theorem (Weak maximum principle). *Let $\mathcal{D} \subseteq \mathbb{R}$ be open and bounded and let $u \in \mathcal{C}(\overline{\mathcal{D}})$ be harmonic in $\mathcal{D}$. Then u attains its extreme values on $\partial\mathcal{D}$: there exist $(x_{\min}, y_{\min})$, $(x_{\max}, y_{\max}) \in \partial\mathcal{D}$ such that (36.1) hold.*

Proof. The important thing to prove here is not the existence of the extreme values (that is the extreme value theorem) but rather their *location*: on $\partial\mathcal{D}$. We often think of the boundary as a "lower-dimensional" set than $\mathcal{D}$; whereas $\mathcal{D}$ is genuinely two-dimensional when $\mathcal{D}$ is open (such a $\mathcal{D}$ contains lots of open balls), often the boundary is parametrized as the image of a curve that depends on only one variable—a very one-dimensional set. Thus the boundary may be "easier" to work with than $\mathcal{D}$.

Here is how we would like the proof to work, although it will not work this way. Suppose that u attains its maximum at $(x_0, y_0) \in \mathcal{D}$; since $\mathcal{D}$ is open, this means $(x_0, y_0) \notin \partial\mathcal{D}$. Then $u_{xx}(x_0, y_0) \le 0$ and $u_{yy}(x_0, y_0) \le 0$ by the second derivative test. If either of these inequalities is strict, we have $\Delta u(x_0, y_0) < 0$. This contradicts the assumption that $\Delta u(x_0, y_0) = 0$. The problem is that we cannot guarantee that either inequality is strict.

We therefore modify u into a function whose second derivatives are more tractable. Specifically, we first summon up a function $w \in \mathcal{C}(\overline{\mathcal{D}})$ such that $w(x,y) \ge 0$ on $\mathcal{D}$ and $\Delta w(x,y) > 0$ on $\mathcal{D}$. Then we put $v(x,y) := u(x,y) + \epsilon w(x,y)$ for some $\epsilon > 0$, which we will select later. We compute $\Delta v(x,y) = \epsilon \Delta w(x,y) > 0$ on $\mathcal{D}$. The extreme value theorem gives $(x_1, y_1) \in \overline{\mathcal{D}}$ such that v attains its maximum at (x_1, y_1). If $(x_1, y_1) \in \mathcal{D}$, the second derivative test gives $\Delta v(x_1, y_1) \le 0$, a contradiction. Thus $(x_1, y_1) \in \partial\mathcal{D}$.

Now we use the nonnegativity of w to compare

$$u(x,y) \le u(x,y) + \epsilon w(x,y) = v(x,y) \le v(x_1,y_1) = u(x_1,y_1) + \epsilon w(x_1,y_1).$$

for any $(x,y) \in \overline{\mathcal{D}}$. And last we employ ϵ. Since $w \in \mathcal{C}(\overline{\mathcal{D}})$, the extreme value theorem guarantees the existence of $M \ge 0$ such that $w(x,y) \le M$ for all $(x,y) \in \overline{\mathcal{D}}$. Thus $v(x_1,y_1) \le u(x_1,y_1) + \epsilon M$, and so we have shown

$$u(x,y) \le u(x_1,y_1) + \epsilon w(x_1,y_1) \le u(x_1,y_1) + \epsilon M$$

for all $(x,y) \in \overline{\mathcal{D}}$ and all $\epsilon > 0$. Taking the limit as $\epsilon \to 0^+$, we obtain

$$u(x,y) = \lim_{\epsilon\to 0^+} u(x,y) \le \lim_{\epsilon\to 0^+} u(x_1,y_1) + \epsilon M = u(x_1,y_1)$$

for all $(x,y) \in \overline{\mathcal{D}}$. That is,

$$u(x_1,y_1) = \max_{(x,y)\in\overline{\mathcal{D}}} u(x,y) \quad \text{and} \quad (x_1,y_1) \in \partial\mathcal{D},$$

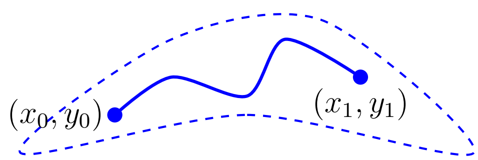
as desired. ■

36.10 Problem. Reread the proof of the maximum principle for the heat equation (Theorem 25.1) and compare it to the one above for Laplace's equation. In what sense does the result for the heat equation provide more information for a more restricted geometry? How do derivative tests from calculus appear in each proof? What are the roles of the functions v and the parameters ϵ in each proof? How are they similar and/or different?

36.11 Problem. Prove the statement in the weak maximum principle about the minimum of u. [Hint: *what is $-u$ doing?*]

The weak maximum principle is "weak" because while it *guarantees* the attainment of the maximum on the boundary, it does not *preclude* the attainment of the maximum on the "interior." After all, a function in general can attain its maximum in many places (think of sine and cosine). If we add some more geometric restrictions, this cannot happen for a harmonic function.

36.12 Definition. *A set $\mathcal{D} \subseteq \mathbb{R}^2$ is* **CONNECTED** *if for each (x_0, y_0), $(x_1, y_1) \in \mathcal{D}$, there is a continuous function $\gamma \colon [0,1] \to \mathcal{D}$ such that $\gamma(0) = (x_0, y_0)$ and $\gamma(1) = (x_1, y_1)$.*



36.13 Theorem. *Let $\mathcal{D} \subseteq \mathbb{R}^2$ be open and connected. If u is harmonic on $\mathcal{D}$, and if there is $(x_0, y_0) \in \mathcal{D}$ such that $u(x_0, y_0) = \max_{(x,y)\in\mathcal{D}} u(x, y)$ or $u(x_0, y_0) = \min_{(x,y)\in\mathcal{D}} u(x, y)$, then u is constant on $\mathcal{D}$.*

Proof. First, note what this theorem is *not* saying: $\mathcal{D}$ is not necessarily bounded, and u is not necessarily continuous on $\overline{\mathcal{D}}$. Thus the extreme value theorem does not come into play: the hypotheses simply presume the attainment of the maximum in the "interior" $\mathcal{D}$. To prove this theorem, we need a claim, which we will justify later: if $\mathcal{B}((x_\star, y_\star); r) \subseteq \mathcal{D}$ for some $(x_\star, y_\star) \in \mathcal{D}$ and $r > 0$, and if $u(x_\star, y_\star) = \max_{(x,y)\in\mathcal{B}((x_\star,y_\star);r)} u(x, y)$, then u is constant on $\mathcal{B}((x_\star, y_\star); r)$.

Fix $(x_1, y_1) \in \mathcal{D}$. Our goal is to show $u(x_0, y_0) = u(x_1, y_1)$; that will certainly prove that u is constant on $\mathcal{D}$. We first deploy connectedness: let $\gamma \colon [0,1] \to \mathcal{D}$ be continuous with $\gamma(0) = (x_0, y_0)$ and $\gamma(1) = (x_1, y_1)$. Since $\mathcal{D}$ is open, there is $r > 0$ such that $\mathcal{B}((x_0, y_0); r) \subseteq \mathcal{D}$. The claim implies that u is constant on $\mathcal{B}((x_0, y_0); r)$, for certainly $u(x_0, y_0)$ is the maximum value of u on this ball. If $(x_1, y_1) \in \mathcal{B}((x_0, y_0); r)$, then we are done.

Otherwise, we deploy a totally nontrivial topological trick. We can "cover" the image of γ in $\mathcal{D}$ with a finite number of overlapping open balls, starting with $\mathcal{B}((x_0, y_0); r)$ above, such that the center of the kth ball is contained in the $(k-1)$st ball, with (x_1, y_1) as the center of the last ball. Our ability to perform this covering hinges on a technical compactness argument, which we will not pursue further here.

This means that the center of the second ball, which we call (X_1, Y_1), is contained in $\mathcal{B}((x_0, y_0); r)$, and so $u(X_1, Y_1) = u(x_0, y_0)$. Then u attains its maximum on the second ball at the center, and so u is constant on that ball, and specifically u equals $u(x_0, y_0)$ on the second ball. Thus $u(X_1, Y_1) = u(x_0, y_0)$. If (x_1, y_1) is in the second ball, stop. Otherwise, proceed to the third ball: its center (X_2, Y_2) is in the second ball, so $u(X_2, Y_2) = u(x_0, y_0)$, and so u attains its maximum on the third ball at the center, and so u is constant on the third ball. Turn the crank… ■

36.14 Problem. Draw a picture illustrating the situation in the proof above, assuming that it takes four balls to cover the image of γ in $\mathcal{D}$. (Be sure to include both $\mathcal{D}$ and the image of γ in your picture.) Remember that the center of the first ball is (x_0, y_0), and the center of the second ball lies in the first ball, and the center of the third ball lies in the second ball, and the center of the fourth ball (which is (x_1, y_1)) lies in the third ball.

36.15 Problem. Give an example of $\mathcal{D} \subseteq \mathbb{R}^2$ open and u harmonic on $\mathcal{D}$ such that u attains its maximum in $\mathcal{D}$ and is not constant on $\mathcal{D}$. [Hint: *such a $\mathcal{D}$ cannot be connected; try to define u "piecewise" on different "components" of $\mathcal{D}$.*]

36.16 Problem. Prove the "strong minimum" principle part of Theorem 36.13. [Hint: *as usual, think about $-u$.*]

36.17 Remark. *There is a strong maximum principle for the heat equation that says that if the maximum is achieved at an "interior" point of the rectangle in Theorem 25.1 (in addition to the requisite existence on the parabolic boundary), then the solution is constant on that rectangle. This is hard to prove because it depends on a more complicated version of the claim from the first paragraph of the proof of Theorem 36.13. (It is fair to view the "covering" topological trick in the last paragraph above as also quite hard—and it is—but the situation with heat is simply harder.)*

Day 37: Wednesday, November 6.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Theorem 2 on p. 387 uses the weak maximum principle to prove uniqueness for the Dirichlet problem; Theorem 3 on that page is continuous dependence on boundary data. Read Example 1 on p. 386 Pages 366–368 derive Laplace's equation in polar coordinates and discuss periodicity issues in θ; see Proposition 1 on p. 366 and equation (3) on p. 368.

A nice contrapositive of the strong maximum principle is that any nonconstant function that is harmonic on an open, connected set must achieve its extreme values on the boundary. A nice immediate consequence of the maximum principle(s) for Laplace's equation is uniqueness for the Dirichlet problem.

37.1 Theorem. *Let $\mathcal{D} \subseteq \mathbb{R}^2$ be open and either bounded or connected. There exists at most one solution of*

$$\begin{cases} \Delta u = 0 \ \text{on } \mathcal{D} \\ u(x, y) = f(x, y) \ \text{on } \partial\mathcal{D}. \end{cases}$$

Proof. As usual, it suffices to show that the only solution to

$$\begin{cases} \Delta u = 0 \text{ on } \mathcal{D} \\ u(x, y) = 0 \text{ on } \partial\mathcal{D}. \end{cases}$$

is $u = 0$. The maximum principle (weak if $\mathcal{D}$ is bounded but maybe not connected, strong if $\mathcal{D}$ is connected but maybe not bounded) guarantees

$$\max_{(x,y)\in\overline{\mathcal{D}}} u(x, y) = \max_{(x,y)\in\partial\mathcal{D}} u(x, y) = 0,$$

and the corresponding "minimum" principle says

$$\min_{(x,y)\in\overline{\mathcal{D}}} u(x, y) = \min_{(x,y)\in\partial\mathcal{D}} u(x, y) = 0$$

as well. Thus $0 \le u(x, y) \le 0$ for all $(x, y) \in \overline{\mathcal{D}}$. ■

37.2 Problem. Prove the following "continuous dependence on boundary conditions" result for the Dirichlet problem. Suppose that $\mathcal{D} \subseteq \mathbb{R}^2$ is open and bounded. Let f_1, $f_2 \in \mathcal{C}(\partial\mathcal{D})$ and suppose there is $\epsilon > 0$ such that $|f_1(x, y) - f_2(x, y)| < \epsilon$ for all $(x, y) \in \partial\mathcal{D}$. Let u and v solve

$$\begin{cases} \Delta u = 0 \text{ on } \mathcal{D} \\ u(x, y) = f_1(x, y) \text{ on } \partial\mathcal{D} \end{cases} \quad \text{and} \quad \begin{cases} \Delta v = 0 \text{ on } \mathcal{D} \\ v(x, y) = f_2(x, y) \text{ on } \partial\mathcal{D}. \end{cases}$$

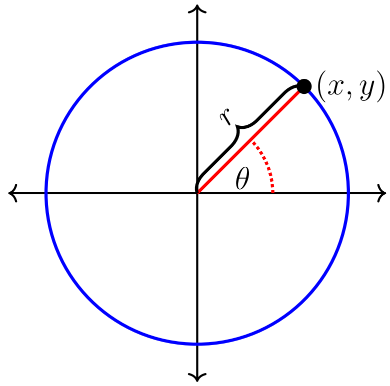
Prove that $|u(x, y) - v(x, y)| < \epsilon$ for all $(x, y) \in \overline{\mathcal{D}}$.

Now we need to fill in some gaps in our work on the strong maximum principle (recall that we left unproved a claim within the proof). This will lead to some related results, and manipulations of integrals, that are all valuable by themselves. Specifically, we will want to study Laplace's equation on balls, and polar coordinates are ideal for balls.

Recall that for any $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, there exist $r > 0$ and $\theta \in \mathbb{R}$ such that

$$x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta).$$

The value r is uniquely determined by $r = \sqrt{x^2 + y^2}$, whereas θ can be replaced by $\theta + 2\pi k$ for any $k \in \mathbb{Z}$; the choice of θ is unique if it is restricted to belong to $(-\pi, \pi]$.



We are interested in the following situation. Suppose that $u \colon \mathcal{B}((0,0); R) \to \mathbb{R}$ is harmonic for some $R > 0$. We can "rectangularize" the domain of u by putting

$$\mathcal{R} := \big\{(r, \theta) \in \mathbb{R}^2 \;\big|\; 0 \le r \le R,\ -\pi \le \theta \le \pi\big\}$$

and then defining

$$U \colon \mathcal{R} \to \mathbb{R} \colon (r, \theta) \mapsto u(r\cos(\theta), r\sin(\theta)).$$

The upshot here is that $\mathcal{R}$ is a "simpler" region than $\mathcal{B}((0,0); R)$, especially from the point of view of integration. So, what does U do? Since this is a course on PDE, we should find a PDE that U solves. And since $u_{xx} + u_{yy} = 0$, a natural starting point is to compute U_{rr} and $U_{\theta\theta}$. This is mostly a thankless calculation with the multivariable chain rule and what we find is

$$U_{rr} + \frac{1}{r^2}U_{\theta\theta} + \frac{1}{r}U_r = 0,\ 0 < r \le R,\ -\pi \le \theta \le \pi. \tag{37.1}$$

This looks quite a bit worse than Laplace's equation: there is the first-order partial U_r in there, and this is actually a *variable*-coefficient PDE (and we have only ever studied *constant*-coefficient PDE).

Here is that calculation, in which we suppress the inputs to u and its partials:

$$U_r(r, \theta) = u_x\cos(\theta) + u_y\sin(\theta), \tag{37.2}$$

$$\begin{aligned} U_{rr}(r, \theta) &= \big[u_x x\cos(\theta) + u_{xy}\sin(\theta)\big]\cos(\theta) + \big[u_{yx}\cos(\theta) + u_{yy}\sin(\theta)\big]\sin(\theta) \\ &= u_{xx}\cos^2(\theta) + 2u_{xy}\sin(\theta)\cos(\theta) + u_{yy}\sin^2(\theta), \end{aligned}$$

$$U_\theta(r, \theta) = -u_x r\sin(\theta) + u_y r\cos(\theta),$$

and

$$\begin{aligned} U_{\theta\theta}(r, \theta) = &-\big[-u_{xx}r\sin(\theta) + u_{xy}r\cos(\theta)\big]r\sin(\theta) \\ &- u_x r\cos(\theta) \\ &+ \big[-u_{yx}r\sin(\theta) + u_{yy}r\cos(\theta)\big]r\cos(\theta) \end{aligned}$$

$$
\begin{aligned}
&- u_y r\sin(\theta) \\
&= u_{xx}r^2\sin^2(\theta) - 2u_{xy}r^2\sin(\theta)\cos(\theta) + u_{yy}r^2\cos^2(\theta) \\
&- r\big[u_x r\cos(\theta) + ru_y\sin(\theta)\big] \\
&= u_{xx}r^2\sin^2(\theta) - 2u_{xy}r^2\sin(\theta)\cos(\theta) + u_{yy}r^2\cos^2(\theta) - rU_r(r,\theta).
\end{aligned}
$$

If we try to add U_{rr} and $U_{\theta\theta}$ (the natural thing to do, since that is ΔU), the terms that we would like to combine do not combine well because of the factor of r^2 throughout $U_{\theta\theta}$. Assuming $r \neq 0$ and dividing through by r^2 gives

$$U_{rr} + U_{\theta\theta} = u_{xx} + u_{yy} - \frac{1}{r}U_r.$$

Thus

$$U_{rr}(r,\theta) + \frac{1}{r^2}U_{\theta\theta}(r,\theta) + \frac{1}{r}U_r(r,\theta) = \Delta u(r\cos(\theta), r\sin(\theta)), \tag{37.3}$$

so when u is harmonic we end up with (37.1).

37.3 Remark. *Our strategy with Laplace's equation in polar coordinates will be to assume that $\Delta u = 0$ and then obtain that (37.1) holds. However, one could work backwards: solve (37.1) and then define $u(x,y) = U(\mathrm{P}(x,y), \Theta(x,y))$, where $\mathrm{P}(x,y) = \sqrt{x^2+y^2}$ and $\Theta(x,y)$ is chosen to be continuously differentiable and to satisfy $x = \mathrm{P}(x,y)\cos(\Theta(x,y))$ and $y = \mathrm{P}(x,y)\sin(\Theta(x,y))$. For this to work, $U(r,\cdot)$ needs to be 2π-periodic.*

Day 38: Friday, November 8.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

The mean value theorem is proved at the top of p. 390; this relies on Lemma 1 from p. 389. See the remark on p. 390 for a vector calculus interpretation of that lemma in terms of Green's theorem and line integrals.

We now build the tools to prove that one unproved claim in the proof of the strong maximum principle. Our chief tool will be the following mean value principle, which is quite valuable in and of itself. To appreciate the formula below, suppose that $u\colon \mathcal{B}((x_0,y_0);R) \to \mathbb{R}$ is continuous for some $(x_0,y_0) \in \mathbb{R}^2$ and $R > 0$. Let $0 < r < R$. Then u is defined on the circle of radius r centered at (x_0,y_0); any point on this circle has the form $(x_0 + r\cos(\theta), y_0 + r\sin(\theta))$ for some $\theta \in [-\pi,\pi]$. The restriction of u to this circle is therefore the map $\theta \mapsto u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))$, and the average value of this map is

$$\frac{1}{2\pi}u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\ d\theta.$$

The mean value principle for harmonic functions states that $u(x_0,y_0)$ equals this integral: the value of u at the center of a ball is equal to the average value of u on any "concentric" circle within this ball. This is a remarkable property of "stability" and "averaging" of harmonic functions—the average value over circles is "stable" in that it does not change as the radius of the circle changes.

38.1 Theorem (Mean value principle for harmonic functions). *Let u be harmonic on $\mathcal{B}((x_0, y_0); R)$. Then*

$$u(x_0, y_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\ d\theta,\ \ 0 \le r < R.$$

Proof. We give the proof only for $x_0 = y_0 = 0$. Define

$$\phi\colon [0, R) \to \mathbb{R}\colon r \mapsto \frac{1}{2\pi} \int_{-\pi}^{\pi} u(r\cos(\theta), r\sin(\theta))\ d\theta.$$

Then

$$\phi(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(0, 0)\ d\theta = u(0, 0)$$

and so if we can show that ϕ is constant, then $\phi(r) = \phi(0) = u(0, 0)$ for all r, and that is our desired result.

We do this by computing $\phi'(r)$ for $r > 0$ and showing $\phi'(r) = 0$. We differentiate under the integral to find

$$\begin{aligned}\phi'(r) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_r\big[u(r\cos(\theta), r\sin(\theta))\big]\ d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \big[u_x(r\cos(\theta), r\sin(\theta))\cos(\theta) + u_y(r\cos(\theta), r\sin(\theta))\sin(\theta)\big]\ d\theta.\end{aligned}$$

Now here is the advantage of polar coordinates: we recall from (37.2) that with $U(r, \theta) := u(r\cos(\theta), r\sin(\theta))$ and u harmonic, we have

$$U_r(r, \theta) = u_x(r\cos(\theta), r\sin(\theta))\cos(\theta) + u_y(r\cos(\theta), r\sin(\theta))\sin(\theta).$$

Thus

$$\phi'(r) = \frac{1}{2\pi} \int_{-\pi}^{\pi} U_r(r, \theta)\ d\theta.$$

This may look no better, but we will compute below that, for $r > 0$,

$$\int_{-\pi}^{\pi} U_r(r, \theta)\ d\theta = \frac{1}{r} \iint_{x^2+y^2 \le r^2} \Delta u(x, y)\ dx\ dy,$$

where by $\iint_{x^2+y^r \le r^2}$ we mean the double integral over the (closed) ball $\big\{(x, y) \in \mathbb{R}^2 \ \big|\ x^2 + y^2 \le r^2\big\}$. Since $\Delta u = 0$, this gives the desired identity $\phi'(r) = 0$. ■

Now here is that last calculation.

38.2 Lemma. *Let $u \in \mathcal{C}^2(\mathcal{B}((0,0);R))$. As usual, for $0 \le r \le R$ and $\theta \in \mathbb{R}$, let $U(r,\theta) := u(r\cos(\theta), r\sin(\theta))$. Then*

$$\iint_{x^2+y^2 \le \rho^2} \Delta u(x,y)\ dx\ dy = \rho \int_{-\pi}^{\pi} U_r(\rho,\theta)$$

for any $\rho \in (0,R)$.

Proof. We switch to polar coordinates. Recall that for $f \in \mathcal{C}(\mathcal{B}(0,0);R)$ we have

$$\iint_{x^2+y^2 \le \rho^2} f(x,y)\ dx\ dy = \int_0^{2\pi} \int_0^{\rho} f(r\cos(\theta), r\sin(\theta)) r\ dr\ d\theta.$$

Thus

$$\begin{aligned}\mathcal{I}(\rho) := \iint_{x^2+y^2 \le \rho^2} \Delta u(x,y)\ dx\ dy &= \int_0^{2\pi} \int_0^{\rho} \Delta u(r\cos(\theta), r\sin(\theta)) r\ dr\ d\theta \\ &= \int_0^{2\pi} \int_0^{\rho} \left(U_{rr}(r,\theta) + \frac{1}{r^2} U_{\theta\theta}(r,\theta) + \frac{1}{r} U_r(r,\theta)\right) r\ dr\ d\theta. \qquad (38.1)\end{aligned}$$

In (37.3) we calculated

$$\Delta u(r\cos(\theta), r\sin(\theta)) = U_{rr}(r,\theta) + \frac{1}{r^2} U_{\theta\theta}(r,\theta) + \frac{1}{r} U_r(r,\theta) \qquad (38.2)$$

for $r \ne 0$. Since $\Delta u \in \mathcal{C}(\mathcal{B}((0,0);R))$, the limit as $r \to 0$ of the right side is defined, and so the integral over $[0,\rho]$ with respect to r in (38.1) is not really improper; just think of the integrand as being 0 at $r = 0$, thanks to that extra factor of r. (Morally, this is like thinking of $\mathrm{sinc}(x)$ as $\sin(x)/x$ even at $x = 0$.)

Then

$$\mathcal{I}(\rho) = \underbrace{\int_0^{2\pi} \int_0^{\rho} \big[U_r(r,\theta) + rU_{rr}(r,\theta)\big]\ dr\ d\theta}_{\mathcal{I}_1(\rho)} + \underbrace{\int_0^{2\pi} \int_0^{\rho} \frac{U_{\theta\theta}(r,\theta)}{r}\ dr\ d\theta}_{\mathcal{I}_2(\rho)}.$$

We actually do need to justify why splitting up this integral is valid: does separating that $U_{\theta\theta}(r,\theta)/r$ term from the rest cause any problems? No: from (37.3) we have

$$\frac{U_{\theta\theta}(r,\theta)}{r} = r\Delta u(r\cos(\theta), r\sin(\theta)) - rU_{rr}(r,\theta) - U_r(r,\theta),$$

and so the limit as $r \to 0$ of the integrand in $\mathcal{I}_2(\rho)$ exists. Thus, again, the integral over $[0,\rho]$ in $\mathcal{I}_2(\rho)$ is not improper, and so we think of the integrand as taking that limiting value at $r = 0$.

Now we can work on the two integrals $\mathcal{I}_1(\rho)$ and $\mathcal{I}_2(\rho)$. We recognize the product rule in the first:

$$U_r(r,\theta) + rU_{rr}(r,\theta) = (1 \cdot U_r(r,\theta)) + rU_{rr}(r,\theta) = \partial_r[rU_r(r,\theta)],$$

and so

$$\mathcal{I}_1(\rho) = \int_0^{2\pi} \int_0^{\rho} \partial_r[rU_r(r,\theta)] \; dr \; d\theta = \int_0^{2\pi} \rho U_r(\rho,\theta) \; d\theta = \rho \int_{-\pi}^{\pi} U_r(\rho,\theta) \; d\theta$$

by the 2π-periodicity of $U_r(\rho,\cdot)$.

This is exactly the value that we want for all of $\mathcal{I}(\rho)$, so we hope that $\mathcal{I}_2(\rho) = 0$. To compute that, we interchange the order of integration (permissible by the standard version of Fubini's theorem for double integrals over rectangles, not the ticklish version from Theorem 30.2, because, by the remarks above, the integrand in $\mathcal{I}_2(\rho)$ is continuous):

$$\mathcal{I}_2(\rho) = \int_0^{\rho} \int_0^{2\pi} \frac{U_{\theta\theta}(r,\theta)}{r} \; d\theta \; dr = \int_0^{\rho} \frac{U_\theta(r,2\pi) - U_\theta(r,0)}{r} \; dr.$$

By definition of U, we have $U(r,\theta+2\pi) = U(r,\theta)$ for all r and θ, and so $U_\theta(r,\cdot)$ is also 2π-periodic for each r. That is, $U_\theta(r,2\pi) - U_\theta(r,0) = 0$ for all r, and so $\mathcal{I}_2(\rho) = 0$, as desired. ■

38.3 Problem. Prove the mean value principle for a general center (x_0,y_0). [Hint: *apply the version proved above to* $v(x,y) := u(x+x_0,y+y_0)$ *for* $(x,y) \in \mathcal{B}((0,0);R)$.]

38.4 Problem. We have largely been studying Laplace's equation in two dimensions: $u_{xx} + u_{yy} = 0$. In one dimension, Laplace's equation is $y'' = 0$, where $y = y(x)$. The analogue of open balls (and, more generally, open sets) in one dimension is open intervals (a,b).

(i) Assume here that $-\infty < a < b < \infty$. Prove that all solutions to $y'' = 0$ on (a,b) satisfy the strong maximum principle in the sense that if $y'' = 0$ on (a,b) and y is continuous on $[a,b]$, and if there is $x_0 \in (a,b)$ such that $y(x_0) = \max_{a \le x \le b} y(x)$, then y is constant on $[a,b]$.

(ii) Let $x_0 \in \mathbb{R}$ and $R > 0$. Prove that all solutions to $y'' = 0$ on $(x_0 - R, x_0 + R)$ satisfy the mean value equation

$$y(x_0) = \frac{1}{2r} \int_{x_0-r}^{x_0+r} y(x) \; dx$$

for $r \in (0,R)$.

It turns out that the mean value principle also characterizes harmonic functions in the following sense.

38.5 Corollary (Converse to the mean value principle). *Let* $\mathcal{D} \subseteq \mathbb{R}^2$ *be open and let* $u \in \mathcal{C}^2(\mathcal{D})$. *Suppose that for any* $(x_0,y_0) \in \mathcal{D}$ *and any* $R > 0$ *such that* $\mathcal{B}((x_0,y_0);R) \subseteq \mathcal{D}$, *u satisfies the mean value identity*

$$u(x_0,y_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(x_0 + r\cos(\theta), y_0 + r\sin(\theta)) \; d\theta$$

for $0 < r < R$. *Then u is harmonic on* $\mathcal{D}$.

Proof. Suppose not. Then there is $(x_0, y_0) \in \mathcal{D}$ such that $\Delta u(x_0, y_0) \neq 0$, and so either $\Delta u(x_0, y_0) > 0$ or $\Delta u(x_0, y_0) < 0$. Suppose the former; then by continuity and the openness of $\mathcal{D}$, there is $R > 0$ such that $\Delta u(x, y) > 0$ for $(x, y) \in \mathcal{B}((x_0, y_0); R) \subseteq \mathcal{D}$.

Now define

$$\phi\colon [0, R) \to \mathbb{R}\colon r \mapsto \frac{1}{2\pi}\int_{-\pi}^{\pi} u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\ d\theta.$$

Then ϕ is constant: $\phi(r) = u(x_0, y_0)$ for all r. But, as in the proof of the mean value principle, we may compute

$$\phi'(r) = \frac{1}{2\pi r}\iint_{(x-x_0)^2+(y-y_0)^2 \le r^2} \Delta u(x, y)\ dx\ dy > 0,$$

which is impossible if ϕ is constant. ■

38.6 Problem. Let u be harmonic on $\mathcal{B}((0, 0); R)$. Prove that

$$u(0, 0) = \frac{1}{\pi r^2}\iint_{x^2+y^2 \le r^2} u(x, y)\ dx\ dy$$

for any $r \in (0, R)$. Since πr^2 is the area of the circle of radius r, we can interpret the double integral on the right as the average value of u over $\mathcal{B}((0, 0); r)$, and so this formula is another version of the mean value principle in a more two-dimensional sense. [Hint: *first, obtain from the ordinary mean value principle*

$$\frac{1}{2\pi}\int_0^{2\pi} u(0, 0)\rho\ d\theta = \frac{1}{2\pi}\int_0^{2\pi} u(\rho\cos(\theta), \rho\sin(\theta))\rho\ d\theta.$$

Integrate both sides from $\rho = 0$ to $\rho = r$, simplify the left side, interchange the order of integration on the right, and think about polar coordinates.]

Day 39: Monday, November 11.

Now we can fill in a lingering gap from our proof of the strong maximum principle. The following lemma is essentially the strong maximum principle for balls; recall that we "chained this result along" using connectedness and one topological trick to get the strong maximum principle for arbitrary open and connected sets.

39.1 Lemma. *Let u be harmonic on $\mathcal{B}((x_0, y_0); R)$. If*

$$u(x_0, y_0) = \max_{(x,y)\in\mathcal{B}((x_0,y_0);R)} u(x, y),$$

then u is constant on $\mathcal{B}((x_0, y_0); R)$.

Proof. Fix $0 < r < R$. The mean value principle gives

$$u(x_0, y_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\ d\theta,$$

and gently rewriting the left side gives

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} u(x_0, y_0)\ d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\ d\theta.$$

Subtracting, we find

$$\int_{-\pi}^{\pi} \left[u(x_0, y_0) - u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\right]\ d\theta = 0.$$

The integrand here is nonnegative since

$$u(x_0, y_0) = \max_{(x,y)\in\mathcal{B}((x_0,y_0);R)} u(x, y) \geq u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))$$

for all $\theta \in [-\pi, \pi]$ and $r \in (0, R)$. Since the integral is 0, the integrand must be identically 0. Thus

$$u(x_0, y_0) = u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))$$

for any $r \in (0, R)$ and $\theta \in [-\pi, \pi]$. Any point $(x, y) \in \mathcal{B}((x_0, y_0); R)$ can be written as $u(x, y) = u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))$ for some such r and θ, and so u is indeed constant. ■

We prove one last "averaging" result for the Laplacian—not even for harmonic functions—just to emphasize the role of average value and integrals in connection with Δ. Informally, $\Delta u(x_0, y_0)$ is twice the average value of all of the second directional derivatives of u at (x_0, y_0).

39.2 Theorem. *Let $u \in \mathcal{C}^2(\mathcal{B}(x_0, y_0); R)$ for some $R > 0$ and put, as usual, $U(r, \theta) := u(x_0 + r\cos(\theta), y_0 + r\sin(\theta))$, so $U_r(0, \theta)$ is the directional derivative of u in the direction of $(\cos(\theta), \sin(\theta))$ through (x_0, y_0). Then*

$$\frac{\Delta u(x_0, y_0)}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} U_{rr}(0, \theta)\ d\theta,$$

where the expression on the right is the average value of all of the second directional derivatives of u at (x_0, y_0).

Proof. We might be tempted to use the polar coordinates formula

$$U_{rr}(r, \theta) = \Delta u(x_0 + r\cos(\theta), y_0 + r\sin(\theta)) - \frac{1}{r}U_r(r, \theta) - U_{\theta\theta}(r, \theta),$$

but this is only valid for $r \neq 0$. Instead, we recall the formula for U_r from (37.2) and use that to calculate

$$U_{rr}(r, \theta) = u_{xx}(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\cos^2(\theta) + 2u_{xy}(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\sin(\theta)\cos(\theta)$$

$$+ u_{yy}(x_0 + r\cos(\theta), y_0 + r\sin(\theta))\sin^2(\theta)$$

and so

$$U_{rr}(0,\theta) = u_{xx}(x_0,y_0)\cos^2(\theta) + u_{xy}(x_0,y_0)\sin(2\theta) + u_{yy}(x_0,y_0)\sin^2(\theta).$$

We integrate over $[-\pi, \pi]$ with respect to θ and use the identities

$$\int_{-\pi}^{\pi} \cos^2(\theta)\ d\theta = \int_{-\pi}^{\pi} \sin^2(\theta)\ d\theta = \pi \quad \text{and} \quad \int_{-\pi}^{\pi} \sin(2\theta)\ d\theta = 0$$

to obtain

$$\int_{-\pi}^{\pi} U_{rr}(0,\theta)\ d\theta = \pi \Delta u(x_0,y_0),$$

from which the desired identity follows. ■

The very last feature of Laplace's equation that we will consider here is its connection to an *optimization* problem. In calculus, we learn well how to find the extreme values of real-valued functions defined on subsets of $\mathbb{R}^n$ (typically for $n = 1$ and $n = 2$). Remarkably, solutions to Laplace's equation minimize certain functions—whose domains are functions!

We will show this for a very restricted domain: the unit square. This will allow us to avoid invoking some technical, and maybe distracting, results from vector calculus and focus just on the PDE and integral manipulations. Let

$$\mathcal{R} := \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; 0 < x < 1,\ 0 < y < 1\right\}.$$

We will study solutions to the Dirichlet problem

$$\begin{cases} \Delta u = 0 \text{ in } \mathcal{R} \\ u = f \text{ on } \partial\mathcal{R}. \end{cases} \tag{39.1}$$

39.3 Problem. Why are solutions unique?

We first prove a highly useful "integration by parts" identity.

39.4 Problem. To motivate the following, let $f \in \mathcal{C}^2([a,b])$ and let $g \in \mathcal{C}^1([a,b])$ with $g(a) = g(b) = 0$. Show that

$$\int_a^b f''g = -\int_a^b f'g'.$$

39.5 Lemma. *Let $u \in \mathcal{C}^2(\mathcal{R}) \cap \mathcal{C}(\overline{\mathcal{R}})$ and $v \in \mathcal{C}^1(\mathcal{R}) \cap \mathcal{C}(\overline{\mathcal{R}})$ with $v = 0$ on $\partial\mathcal{R}$. Then*

$$\iint_{\overline{\mathcal{R}}} (\Delta u)v = -\iint_{\overline{\mathcal{R}}} (u_x v_x + u_y v_y).$$

Proof. We have

$$\iint_{\overline{\mathcal{R}}} (\Delta u)v = \int_0^1 \int_0^1 (u_{xx}(x,y) + u_{yy}(x,y))v(x,y)\ dx\ dy.$$

We focus on just the integral

$$\int_0^1 u_{xx}(x,y)v(x,y)\ dx.$$

Integrating by parts gives

$$\int_0^1 u_{xx}(x,y)v(x,y)\ dx = u_x(x,y)v(x,y)\big|_{x=0}^{x=1} - \int_0^1 u_x(x,y)v_x(x,y)\ dx.$$

Since $v = 0$ on $\partial\mathcal{R}$, we have in particular $v(1,y) = v(0,y) = 0$ for all y. Thus

$$\int_0^1 u_{xx}(x,y)v(x,y)\ dx = -\int_0^1 u_x(x,y)v_x(x,y)\ dx$$

and so

$$\iint_{\overline{\mathcal{R}}} u_{xx}(x,y)v(x,y)\ dx\ dy = -\iint_{\overline{\mathcal{R}}} u_x(x,y)v_x(x,y)\ dx\ dy.$$

A similar calculation, along with an interchange in the order of integration, shows

$$\iint_{\overline{\mathcal{R}}} u_{yy}(x,y)v(x,y)\ dx\ dy = -\iint_{\overline{\mathcal{R}}} u_y(x,y)v_y(x,y)\ dx\ dy, \tag{39.2}$$

which leads to the desired identity. ■

39.6 Problem. Do that similar calculation with the interchange of integrals to obtain (39.2).

39.7 Problem. Suppose that we have this integration by parts identity for a more general $\mathcal{D} \subseteq \mathbb{R}$:

$$\iint_{\overline{\mathcal{D}}} (\Delta u)v = -\iint_{\overline{\mathcal{D}}} (u_x v_x + u_y v_y) \tag{39.3}$$

for $u \in \mathcal{C}^2(\mathcal{D}) \cap \mathcal{C}(\overline{\mathcal{D}})$ and $v \in \mathcal{C}^1(\mathcal{D} \cap \mathcal{C}(\overline{\mathcal{D}})$ with $v = 0$ on $\partial\mathcal{D}$. Use this to prove uniqueness of solutions to the problem

$$\begin{cases} \Delta u = g \text{ in } \mathcal{D} \\ u = f \text{ on } \partial\mathcal{D}. \end{cases}$$

[Hint: *start with two solutions, subtract, and see what boundary conditions the difference meets. In* (39.3) *use* $u = v$.]

We apply this integration by parts result in the following quite possibly non-obvious way. Let

$$\mathcal{V} := \left\{ w \in \mathcal{C}^2(\mathcal{R}) \cap \mathcal{C}(\overline{\mathcal{R}}) \ \middle|\ w = f \text{ on } \partial\mathcal{R} \right\}.$$

Define
$$\mathcal{E}\colon \mathcal{V} \to \mathbb{R}\colon w \mapsto \iint_{\overline{\mathcal{R}}} (w_x^2 + w_y^2).$$
We will show that any solution u to the Dirichlet problem (39.1) is the minimizer of $\mathcal{E}$ in $\mathcal{V}$ in the sense that
$$\mathcal{E}[u] \le \mathcal{E}[w] \text{ for all } w \in \mathcal{V},$$
and, conversely, any minimizer of $\mathcal{E}$ in $\mathcal{V}$ solves the Dirichlet problem.

Day 40: Wednesday, November 13.

Proving that a solution to the Dirichlet problem minimizes $\mathcal{E}$ is not all that difficult, with one classical trick. Let u solve (39.1) and $w \in \mathcal{V}$; we want to show $\mathcal{E}[u] \le \mathcal{E}[w]$. We can make u show up in $\mathcal{E}[w]$ by adding 0:
$$\mathcal{E}[w] = \mathcal{E}[u + (w - u)].$$
Now we work on $\mathcal{E}[u+v]$ with $v := w - u$ and find
$$\mathcal{E}[u+v] = \mathcal{E}[u] + \mathcal{E}[v] + 2\iint_{\overline{\mathcal{R}}} (u_x v_x + u_y v_y).$$

40.1 Problem. Check that.

The double integral on the right looks like the result of integration by parts, which would allow us to bring $\Delta u = 0$ into the calculation. What matters here is that since $v = w - u$ with w, $u \in \mathcal{V}$, we have $v = f - f = 0$ on $\partial\mathcal{R}$. So, we *can* integrate by parts to find
$$\iint_{\overline{\mathcal{R}}} (u_x v_x + u_y v_y) = \iint_{\overline{\mathcal{R}}} (\Delta u) v = 0,$$
and so
$$\mathcal{E}[w] = \mathcal{E}[u+v] = \mathcal{E}[u] + \mathcal{E}[v] \ge \mathcal{E}[u]$$
since $\mathcal{E}[v] \ge 0$ for any v. This is the desired inequality.

Now suppose that $u \in \mathcal{V}$ minimizes $\mathcal{E}$. By definition of $\mathcal{V}$, we have $u = f$ on $\partial\mathcal{R}$, so we just need to show $\Delta u = 0$. This is the tricky part, and the first trick is to introduce a new function space.

For $r \ge 0$, let
$$\mathcal{C}_0^r(\mathcal{R}) := \left\{ v \in \mathcal{C}^r(\mathcal{R}) \cap \mathcal{C}(\overline{\mathcal{R}}) \;\middle|\; v = 0 \text{ on } \partial\mathcal{R} \right\}.$$

40.2 Problem. How did $\mathcal{C}_0^r(\mathcal{R})$ show up in Lemma 39.5?

Let $v \in \mathcal{C}_0^1(\mathcal{R})$ and $s \in \mathbb{R}$. If we know $u + sv \in \mathcal{V}$, then since u minimizes $\mathcal{E}$ we must have $\mathcal{E}[u] \le \mathcal{E}[u + sv]$.

40.3 Problem. Explain why $u + sv \in \mathcal{V}$.

We expand, similar to some work above,

$$\mathcal{E}[u + sv] = \mathcal{E}[u] + s^2\mathcal{E}[v] + 2s \iint_{\overline{\mathcal{R}}} (u_x v_x + u_y v_y).$$

40.4 Problem. Check that.

Integrating by parts, we find

$$\mathcal{E}[u] \le \mathcal{E}[u] + s^2\mathcal{E}[v] + 2s \iint_{\overline{\mathcal{R}}} (u_x v_x + u_y v_y) = \mathcal{E}[u] + s^2\mathcal{E}[v] - 2s \iint_{\overline{\mathcal{R}}} (\Delta u)v,$$

and so

$$s \iint_{\overline{\mathcal{R}}} (\Delta u)v \le \frac{s^2}{2}\mathcal{E}[v].$$

This is true for all $s \in \mathbb{R}$. When $s > 0$, we just have

$$\iint_{\overline{\mathcal{R}}} (\Delta u)v \le \frac{s}{2}\mathcal{E}[v]$$

and so

$$\iint_{\overline{\mathcal{R}}} (\Delta u)v \le \lim_{s \to 0^+} \frac{s}{2}\mathcal{E}[v] = 0.$$

40.5 Problem. By considering $s < 0$, show as well that

$$0 \le \iint_{\overline{\mathcal{R}}} (\Delta u)v.$$

Thus

$$\iint_{\overline{\mathcal{R}}} (\Delta u)v = 0.$$

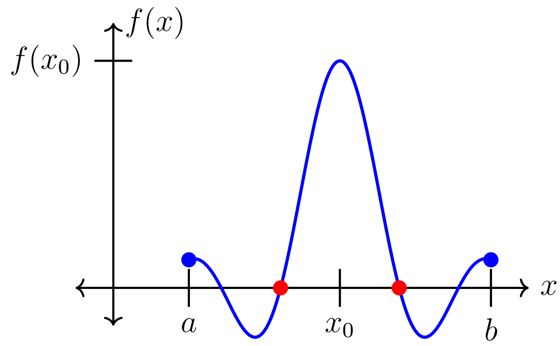
How does this help get us to $\Delta u = 0$?

We need a second trick, and this is an instance of our frequent claim that *integrals are tools for extracting data about functions*. Here the function that matters is Δu, and the tool is multiplying by v and integrating over $\overline{\mathcal{R}}$. The tool returns the data 0 all the time. It turns out that this is enough to conclude that $\Delta u = 0$ on $\mathcal{R}$, but it is easier to see why in one dimension.

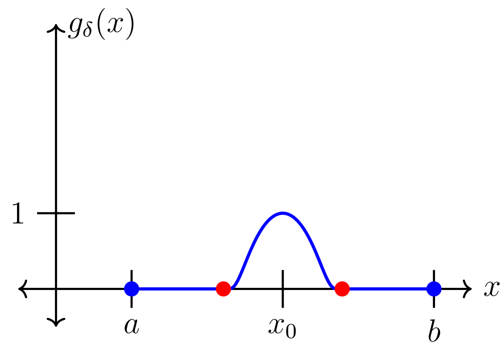
40.6 Theorem (Fundamental lemma of the calculus of variations). *Let $f \in \mathcal{C}([a, b])$ and suppose that $\int_a^b fg = 0$ for all $g \in \mathcal{C}^r([a, b])$ with $g(a) = g(b) = 0$ for some $r \ge 0$. Then $f = 0$.*

Proof. Suppose not. Then there is $x_0 \in [a, b]$ such that $f(x_0) \ne 0$. We assume $x_0 \in (a, b)$ and $f(x_0) > 0$; the work for $f(x_0) < 0$ is very similar, and the cases $x_0 = a$ or $x_0 = b$

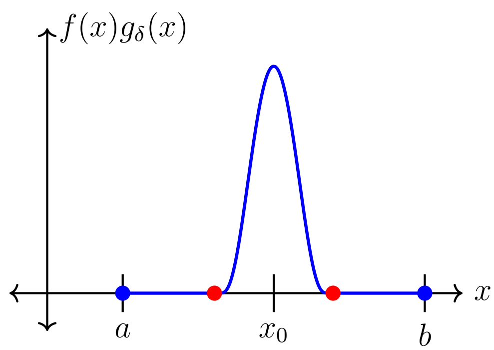
are a problem below. Continuity implies the existence of $\delta > 0$ such that $f(x) > 0$ for $x \in (x_0 - \delta, x_0 + \delta) \subseteq (a, b)$. This should feel very familiar, as we have made this argument multiple times before. Here is a picture.



We claim there is $g_\delta \in \mathcal{C}^\infty([a, b])$ such that $g_\delta(x) = 0$ for $a \leq x \leq x_0 - \delta$ and $x_0 + \delta \leq x \leq b$ but $g_\delta(x) > 0$ for $x_0 - \delta \leq x \leq x_0 + \delta$. Here is another picture.



And here is a picture of how the product fg_δ behaves.



Assuming the existence of this g_δ, we have

$$0 < \int_{x_0-\delta}^{x_0+\delta} f(x)g_\delta(x)\, dx = \int_a^b f(x)g_\delta(x)\, dx = 0,$$

where the last equality follows from the hypothesis. This is a contradiction. ■

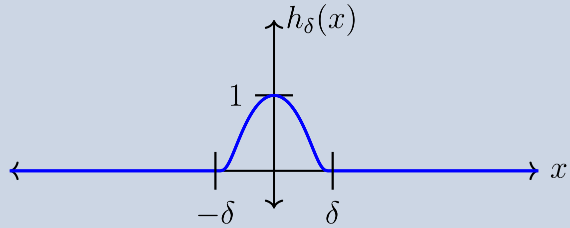
40.7 Problem. In the proof above, why did it suffice to assume $x_0 \in (a,b)$? What happens if $f(a) \neq 0$ but $f(x) = 0$ for $x \in (a,b]$?

Now we show the existence of that g_δ.

40.8 Lemma. *Let $\delta > 0$. Define*

$$h_\delta \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \begin{cases} e^{1/\delta^2} \exp\left(-\dfrac{1}{\delta^2 - x^2}\right), & |x| < \delta \\ 0, & |x| \geq \delta. \end{cases}$$

Then $h_\delta \in \mathcal{C}^\infty(\mathbb{R})$.



Proof. The challenge, of course, is ensuring differentiability at $x = \pm\delta$, as h_δ is infinitely differentiable elsewhere by properties of piecewise functions. At $x = \pm\delta$, the proof is mostly a technical induction argument that hinges on the very rapid vanishing of $e^{-1/x}$ as $x \to \infty$. ■

The factor of e^{1/δ^2} is mostly for convenience, to normalize $h_\delta(0) = 1$. We put

$$g_\delta \colon [a,b] \to \mathbb{R} \colon x \mapsto h_\delta(x - x_0)$$

to obtain the desired behavior.

Everything that we have said in one dimension carries over to two dimensions: if $w \in \mathcal{C}(\overline{\mathcal{D}})$ satisfies $\iint_{\overline{\mathcal{D}}} wv = 0$ for all $v \in \mathcal{C}^r(\mathcal{D}) \cap \mathcal{C}(\overline{\mathcal{D}})$ with $v = 0$ on $\partial\mathcal{D}$, then $w = 0$. The proof is basically the same as the above, except the function h_δ should be replaced by

$$(x,y) \mapsto \begin{cases} e^{1/\delta^2} \exp\left(-\dfrac{1}{\delta^2 - \big(x^2+y^2\big)}\right), & x^2+y^2 < \delta^2 \\ 0, & x^2+y^2 \geq \delta^2. \end{cases}$$

Day 41: Friday, November 15.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Pages 50–52 outline separation of variables with examples for heat- and wave-type equations. Pages 126–127 do separation of variables for the heat equation $u_t = u_{xx}$. Pages 127–130 treat the boundary condition $u(0,t) = u(\pi,t) = 0$ and pp. 130–133 do the periodic boundary condition $u(-\pi,t) = u(\pi,t)$, $u_x(-\pi,t) = u_x(\pi,t)$. See the remarks on p. 133 about the validity of Fourier series expansions. More examples of treating boundary conditions for the heat equation with separable solutions appear on pp. 157–160. Pages 285–289 cover separation of variables for the wave equation. All of these examples are worth reading for practice with the algebraic nuances of boundary conditions.

Throughout this course, we have accrued a handful of precious solution formulas—the glorious solution for the transport equation, D'Alembert's formula for the infinite wave equation (and its adaptations into semi-infinite and bounded problems), the convolution formula for the infinite rod heat equation, and two specialized solutions for Laplace's equation (the upper half-plane problem and the radial solution). We have sometimes used the explicit nature of these solution formulas to understand more about solutions to those PDE, but often we did not presume any particular solution formula, and we used the structure of the PDE and some calculus/analysis techniques (old and new) to learn more.

There is a classical solution method that we have not yet discussed: separation of variables. (This has some moral similarities to, but overall is quite distinct from, separable ODE.) This method yields concrete solutions to certain problems posed on bounded spatial domains (namely, the boundary value problems for heat and wave, which we previously studied without relying too much, if at all, on formulas), and it raises many interesting questions in analysis via the generation of Fourier series. In PDE courses of times past, separation of variables was often the major, possibly only, technique discussed—we live with more evolved sensibilities now. Our appreciation of this method will be as much for those analytic questions that it inspires as for the formulas that we get.

We develop this for the heat equation $u_t = u_{xx}$. If we guess that u has the "product" form

$$u(x,t) = X(x)T(t) \tag{41.1}$$

with $X \in \mathcal{C}^2(\mathbb{R})$ and $T \in \mathcal{C}^1(\mathbb{R})$, then

$$u_t(x,t) = X(x)T'(t) \qquad \text{and} \qquad u_{xx}(x,t) = X''(x)T(t).$$

Thus $u(x,t) = X(x)T(t)$ solves the heat equation $u_t = u_{xx}$ if and only if

$$X(x)T'(t) = X''(x)T(t). \tag{41.2}$$

If $X(x) = 0$ for all x, or $T(t) = 0$, then $u(x,t) = 0$ for all x and t. This is boring and trivial, so we assume that X and T are not identically zero. At the values of x and t such that $X(x) \neq 0$ and $T(t) \neq 0$, we divide (41.2) by $X(x)T(t)$ to find

$$\frac{X''(x)}{X(x)} = \frac{T'(t)}{T(t)}. \tag{41.3}$$

This is interesting. In slightly more compact notation, we have functions f, $g\colon \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = g(t)$$

for all x, $t \in \mathbb{R}$. Both the quantifier "for all" and the different variables x and t are important here! We can take $t = 0$ to conclude $f(x) = g(0)$ for all x, and so f is constant. And we can take $x = 0$ to conclude $f(0) = g(t)$ for all t, and so g is constant. Even better, f and g are the same constant:

$$f(x) = f(0) = g(0) = g(t)$$

for all x and t.

Applying this to (41.3), we find $\lambda \in \mathbb{R}$ such that

$$\frac{X''(x)}{X(x)} = \lambda \quad \text{and} \quad \frac{T'(t)}{T(t)} = \lambda \tag{41.4}$$

for all x, $t \in \mathbb{R}$ (or, at least those at which X and T are nonzero). We can view (41.1), (41.3), and (41.4) as three different layers of separated variables in that we break apart the dependence on and behavior of the x- and t-variables.

The equations in (41.4) are really

$$T'(t) = \lambda T(t) \quad \text{and} \quad X''(x) - \lambda X(x) = 0, \tag{41.5}$$

and we have spent our lives in ODE learning how to solve those. First, though, we may wonder if we lost any legitimacy in dividing by X and T—did a scurrilous division by zero mess up our calculations?

41.1 Problem. No. Convince yourself that if X and T solve (41.5), then $u(x,t) := X(x)T(t)$ does indeed solve $u_t = u_{xx}$.

The equation for T is easy to solve:

$$T(t) = T(0)e^{\lambda t}.$$

The equation for X deserves treatment by cases depending on the sign of λ, which can be any real number. For $\lambda > 0$, it reads

$$X'' - (\sqrt{\lambda})^2 X = 0,$$

and all solutions here are

$$X(x) = c_1 e^{\sqrt{\lambda}x} + c_2 e^{-\sqrt{\lambda}x}.$$

For $\lambda = 0$, the X-equation is just

$$X'' = 0,$$

and solutions are

$$X(x) = c_1 x + c_2.$$

And for $\lambda < 0$, if we write $\lambda = -|\lambda|$, then the X-equation is

$$X'' + |\lambda| X = 0,$$

thus

$$X(x) = c_1 \cos(\sqrt{|\lambda|}x) + c_2 \sin(\sqrt{|\lambda|}x).$$

All together, we have three types of product solutions for the heat equation. To avoid the square root, we will replace λ with λ^2 throughout. We make an important algebraic observation about the first case that will help later. Here, for $\lambda > 0$, we have

$$u(x,t) = \big(c_1 e^{\lambda x} + c_2 e^{-\lambda x}\big) T(0) e^{\lambda^2 t} = \big(T(0)c_1 e^{\lambda x} + T(0) c_2 e^{-\lambda x}\big) e^{\lambda^2 t}. \tag{41.6}$$

The values $T(0)$, c_1, and c_2 can be arbitrary real numbers, so their products are arbitrary too. Thus we can compress the solution to

$$u(x,t) = \big(c_1 e^{\lambda x} + c_2 e^{-\lambda x}\big) e^{\lambda^2 t}. \tag{41.7}$$

More precisely, every solution of the form (41.6) is of the form (41.7). Conversely, let C_1, $C_2 \in \mathbb{R}$ and take $T(0) = 1$, $c_1 = C_1$, and $c_2 = C_2$ to write (41.7) in the original form (41.6).

Similarly, for the other two cases we have product solutions of the form

$$u(x,t) = c_1 x + c_2 \tag{41.8}$$

and

$$u(x,t) = \big(c_1 \cos(\lambda x) + c_2 \sin(\lambda x)\big) e^{-\lambda^2 t}, \tag{41.9}$$

where here we may assume $\lambda > 0$. (Although $\lambda < 0$ still works—why?)

41.2 Problem. Which, if any, of these solutions could be obtained from the Fourier transform?

More generally, we could make a **PRODUCT ANSATZ** for a function u of n variables by writing u as a product of n functions, each of which depends on one, and only one, of the variables of u. For example, to solve the two-dimensional heat equation

$$u_t = u_{xx} + u_{yy} (= \Delta u),$$

we would guess $u(x,y,t) = X(x)Y(t)T(t)$, and then find three ODE, each governing one of X, Y, and T.

41.3 Problem. Find all product solutions $u(x,t) = X(x)T(t)$ for the transport equation $u_t + u_x = 0$. Is every solution to the transport equation a product solution?

However, just because we can do something does not mean that we should. Which, if any, of these product solutions for the heat equation are *relevant*? Separation of variables is particularly useful for boundary value problems on finite spatial domains, and so we consider

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le \pi, \ t \in \mathbb{R} \\ u(0,t) = u(\pi,t) = 0, \ t \in \mathbb{R}. \end{cases} \tag{41.10}$$

This is the heat equation for a finite rod of length π (chosen for convenience) with the ends kept at the constant temperature 0. We have not bothered to specify initial conditions yet. Which, if any, of these product solutions can meet the boundary conditions?

If we try (41.7) with $\lambda > 0$, then we need

$$0 = u(0,t) = (c_1 + c_2)e^{\lambda^2 t} \quad \text{and} \quad 0 = u(\pi,t) = (c_1 e^{\lambda\pi} + c_2 e^{-\lambda\pi})e^{\lambda^2 t}.$$

If we divide through by $e^{\lambda^2 t}$, this becomes the linear system

$$\begin{cases} c_1 + c_2 = 0 \\ c_1 e^{\lambda\pi} + c_2 e^{-\lambda\pi} = 0. \end{cases}$$

Here the unknowns are c_1 and c_2 (and maybe λ). Perhaps it is easiest to view this as a matrix-vector problem:

$$\begin{bmatrix} 1 & 1 \\ e^{\lambda\pi} & e^{-\lambda\pi} \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The determinant of this matrix is $e^{-\lambda\pi} - e^{\lambda\pi}$, which is nonzero.

41.4 Problem. Why?

Thus the only solution to this linear system is $c_1 = c_2 = 0$, and that means $u(x,t) = 0$. That is too boring, too *trivial*. We claim that nothing interesting happens if we try to use a solution of the form (41.8) to meet the boundary conditions.

41.5 Problem. What exactly happens?

And so we are down to solutions of the form (41.9). We want

$$0 = u(0,t) = c_1 e^{-\lambda^2 t} \quad \text{and} \quad 0 = u(\pi,t) = \big(c_1 \cos(\lambda\pi) + c_2 \sin(\lambda\pi)\big)e^{-\lambda^2 t}.$$

The first boundary condition immediately implies $c_1 = 0$, so the second reduces to

$$c_2 \sin(\lambda\pi) = 0.$$

We want to avoid $c_2 = 0$ so we avoid another trivial solution. Here is where λ finally comes into play: we can select λ so that $\sin(\lambda\pi) = 0$. Specifically, since $\sin(\tau) = 0$ if and only if $\tau = k\pi$ for some $k \in \mathbb{Z}$, we can take $\lambda = k \in \mathbb{Z}$ and conclude that

$$u(x,t) = c_2 \sin(kx)e^{-k^2 t} \tag{41.11}$$

solves (41.10). Strictly speaking, we should take $k \geq 1$, since we are assuming $\lambda > 0$, but we can check that the function above is a solution for any $k \in \mathbb{Z}$.

Unfortunately, while this is a nontrivial (not identically zero) solution, the only initial condition that it can meet is very boring:

$$f(x) = u(x,0) = c_2 \sin(kx).$$

We handled much more arbitrary initial conditions in our previous treatments of the finite rod heat equation. At the very least we could use linearity of the heat equation and "superposition" to show that a finite linear combination of functions of the form (41.11) solves (41.10) with a slightly more complicated initial condition. Specifically,

$$u(x,t) := \sum_{k=1}^{n} b_k \sin(kx) e^{-k^2 t}$$

solves

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le \pi, \ t \in \mathbb{R} \\ u(0,t) = u(\pi,t) = 0, \ t \in \mathbb{R} \\ u(x,0) = f(x), \ 0 \le x \le \pi \end{cases}$$

with

$$f(x) = \sum_{k=1}^{n} b_k \sin(kx).$$

41.6 Problem. Check that. Why would summing over nonpositive integers be redundant, e.g., taking $u(x,t) = \sum_{k=-1977}^{1138} b_k \sin(kx) e^{-k^2 t}$?

This initial condition is still very specific. We can check that if the problem above has a solution for some initial condition f, then $f \in \mathcal{C}^2([0,\pi])$ with $f(0) = f(\pi) = 0$. But surely not all such functions are sums of sines!

This motivates a profound idea: what if we can write an initial condition f as a *series* of sines:

$$f(x) = \sum_{k=1}^{\infty} b_k \sin(kx)?$$

Can we then obtain a solution in the form

$$u(x,t) = \sum_{k=1}^{\infty} b_k \sin(kx) e^{-k^2 t}?$$

We will review nuances of series presently; for now, the deeper questions are if we can represent f as such a series, and if u as so defined actually gives a differentiable function. (By the way, labeling the coefficients as b_k here is just to keep us in line with some notation that will appear momentarily.)

It turns out that it is easier to answer this question if we work with series of sines *and* cosines. Here is the motivation for that. Consider a finite rod of length 2π, with endpoints labeled at $\pm\pi$, that is bent into a circle with the ends at $\pm\pi$ joined together. If heat flows continuously around this rod, we should have $u(-\pi,t) = u(\pi,t)$ and $u_x(-\pi,t) = u_x(\pi,t)$. Alternatively, we could just think that the problem

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t \in \mathbb{R} \\ u(-\pi,t) = u(\pi,t), \ u_x(-\pi,t) = u_x(\pi,t), \ t \in \mathbb{R} \end{cases}$$

is worth solving.

41.7 Problem. What happens if we try to use product solutions of the forms (41.7) or (41.8) to solve this problem?

We get to the point and try solutions in the form (41.9). We first differentiate:

$$u_x(x,t) = \big(-c_1\lambda\sin(\lambda x) + c_2\lambda\cos(\lambda x)\big)e^{-\lambda^2 t}.$$

Then we want

$$\big(c_1\cos(-\lambda\pi) + c_2\sin(-\lambda\pi)\big)e^{-\lambda^2 t} = \big(c_1\cos(\lambda\pi) + c_2\sin(\lambda\pi)\big)e^{-\lambda^2 t}$$

and

$$\big(-c_1\lambda\sin(-\lambda\pi) + c_2\lambda\cos(-\lambda\pi)\big)e^{-\lambda^2 t} = \big(-c_1\lambda\sin(\lambda\pi) + c_2\lambda\cos(\lambda\pi)\big)e^{-\lambda^2 t}.$$

Simplifying, this leads to the system

$$2c_2\sin(\lambda\pi) = 0 \quad \text{and} \quad 2c_1\lambda\sin(\lambda\pi) = 0.$$

41.8 Problem. Check that.

Here we are assuming $\lambda > 0$, so to avoid a trivial solution and obtain maximum flexibility with both c_1 and c_2, we choose λ so that $\sin(\lambda\pi) = 0$. Thus, again, $\lambda = k \in \mathbb{Z}$, and we have a solution of the form

$$u(x,t) = \big(c_1\cos(kx) + c_2\sin(kx)\big)e^{-k^2 t}.$$

By linearity,

$$u(x,t) = \sum_{k=0}^{n} \big(a_k\cos(kx) + b_k\sin(kx)\big)e^{-k^2 t}$$

solves

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t \in \mathbb{R} \\ u(-\pi,t) = u(\pi,t), \ u_x(-\pi,t) = u_x(\pi,t), \ t \in \mathbb{R} \\ u(x,0) = f(x), \ -\pi \le x \le \pi \end{cases}$$

with

$$f(x) = \sum_{k=0}^{n} \big(a_k\cos(kx) + b_k\sin(kx)\big).$$

As before, we might conjecture that if we can write the initial data f as

$$f(x) = \sum_{k=0}^{\infty} \big(a_k\cos(kx) + b_k\sin(kx)\big),$$

then putting

$$u(x,t) = \sum_{k=0}^{\infty} \big(a_k\cos(kx) + b_k\sin(kx)\big)e^{-k^2 t}$$

solves the boundary value problem. Does it? Can we write f in this way? What are a_k and b_k? These are the fundamental questions that lead us to Fourier series.

41.9 Problem. Adding is sometimes easier than multiplying. What if we looked for "sum solutions" $u(x,t) = X(x) + T(t)$? Find all such solutions to the heat equation $u_t = u_{xx}$. Does this result make the function w from (26.1) seem less mysterious?

Day 42: Monday, November 18.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Sections 4.1 and 4.2 provide a huge amount of detail on Fourier series. All of this is worth reading, although you are not responsible for the many proofs (which will nonetheless clear up many lingering questions from class). The book defines Fourier series over symmetric intervals $[-L, L]$ for $L > 0$; we are taking $L = \pi$ for convenience. Trigonometric polynomials and the integral inner product are defined on pp. 189–190. Theorem 1 on p. 192 obtains the coefficients of trigonometric polynomials in terms of inner products. Fourier series are defined on pp. 193 and computed in several examples on pp. 193–198 and pp. 201–202. Theorem 2 on p. 198, Theorem 1 on p. 217, and Theorem 2 on p. 221 state "ideal" results for Fourier series convergence that relate to the natural assumptions of our heat boundary value problem. Again, you are not responsible for knowing the proofs of any of these theorems, but they are there if you are interested. See Remark 1 on p. 194 and the remark on p. 204 for hints of the delicacy of convergence otherwise.

We first briefly review some essential aspects of series convergence.

42.1 Definition. *Let (z_k) be a sequence in $\mathbb{C}$, i.e., a function $f \colon \mathbb{N} \to \mathbb{C}$ with $f(k) = z_k$. The term "series" and the symbol $\sum_{k=1}^{\infty} z_k$ have two meanings.*

(i) *We denote by $\sum_{k=1}^{\infty} z_k$ the sequence of partial sums $\big(\sum_{k=1}^{n} z_k\big)$. That is, $\sum_{k=1}^{\infty} z_k$ is the function $f \colon \mathbb{N} \to \mathbb{C}$ such that $f(n) = \sum_{k=1}^{n} z_k$.*

(ii) *If the limit $\lim_{n\to\infty} \sum_{k=1}^{n} z_k$ of this sequence of partial sums exists, then we also denote it by $\sum_{k=1}^{\infty} z_k$. That is, $\sum_{k=1}^{\infty} z_k$ is the number $\sum_{k=1}^{\infty} z_k := \lim_{n\to\infty} \sum_{k=1}^{n} z_k$.*

So, we can always give a rigorous meaning to the symbol $\sum_{k=1}^{\infty} z_k$ by interpreting it as a sequence of partial sums, and sometimes we are lucky enough to have convergence and think of this symbol as being a finite number. We will not make too much of a fuss about where the sum starts, i.e., at $k = 1$ or at some other value of k.

42.2 Example. Let $r \in \mathbb{C}$ with $|r| < 1$. Then for any integer $n \geq 1$ we have

$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r},$$

and so

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}.$$

This is the **GEOMETRIC SERIES**. But we could also think of $\sum_{k=0}^{\infty} r^k$ as the sequence of partial sums

$$\sum_{k=0}^{\infty} r^k = \left(\sum_{k=0}^{n} r^k\right) = \left(\frac{1-r^{n+1}}{1-r}\right).$$

The most important method of determining series convergence is the same as for improper integrals: the comparison test.

42.3 Theorem. *Let (z_k) and (w_k) be sequences with $|z_k| \le |w_k|$. Suppose that $\sum_{k=0}^{\infty} |w_k|$ converges. Then the series $\sum_{k=0}^{\infty} w_k$, $\sum_{k=0}^{\infty} |z_k|$, and $\sum_{k=0}^{\infty} z_k$ all converge, and we have the triangle inequality:*

$$\left|\sum_{k=0}^{\infty} z_k\right| \le \sum_{k=0}^{\infty} |z_k| \le \sum_{k=0}^{\infty} |w_k|.$$

Our goal in studying the periodic heat BVP

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t \in \mathbb{R} \\ u(-\pi, t) = u(\pi, t), \ u_x(-\pi, t) = u_x(\pi, t), \ t \in \mathbb{R} \\ u(x, 0) = f(x), \ -\pi \le x \le \pi \end{cases}$$

is to be able to write

$$f(x) = \sum_{k=0}^{\infty} \big(a_k \cos(kx) + b_k \sin(kx)\big) \quad \text{and then} \quad u(x,t) = \sum_{k=0}^{\infty} \big(a_k \cos(kx) + b_k \sin(kx)\big) e^{-k^2 t}.$$

We first think about such expansions for f and then consider if such u converge and are differentiable.

Some additional notation and calculations will help. For f, $g \in \mathcal{C}_{\text{pw}}([-\pi, \pi])$, as defined some time ago in Definition 28.1, put

$$\langle f, g \rangle := \int_{-\pi}^{\pi} fg.$$

It is then possible to calculate the following:

$$\langle \cos(k\cdot), \cos(j\cdot) \rangle = \begin{cases} 2\pi, \ k = j = 0 \\ \pi, \ k = j \ge 1 \\ 0, \ k \ne j, \end{cases}$$

$$\langle \sin(k\cdot), \sin(j\cdot) \rangle = \begin{cases} \pi, \ k = j \ge 1 \\ 0, \ k \ne j, \end{cases}$$

and

$$\langle \cos(k\cdot), \sin(j\cdot) \rangle = \begin{cases} \pi, & k = j \\ 0, & k \neq j. \end{cases}$$

Perhaps the simplest initial temperature distribution f for the problem at hand is

$$f(x) = \sum_{k=0}^{n} \big(a_k \cos(kx) + b_k \sin(kx)\big), \tag{42.1}$$

which we call a **TRIGONOMETRIC POLYNOMIAL**. Linearity of the integral gives

$$\langle f, \cos(j\cdot) \rangle = \sum_{k=0}^{n} \big(a_k \langle \cos(k\cdot), \cos(j\cdot) \rangle + b_k \langle \sin(k\cdot), \cos(j\cdot) \rangle \big) = \begin{cases} 2\pi a_0, & k = j = 0 \\ \pi a_j, & k = j \geq 1 \\ 0, & k \neq j. \end{cases}$$

Similarly,

$$\langle f, \sin(j\cdot) \rangle = \begin{cases} \pi b_j, & k = j \geq 1 \\ 0, & k \neq j. \end{cases}$$

The mismatch of 2π and π in the formula for $\langle f, \cos(j\cdot) \rangle$ and the fact that $\langle f, \sin(j\cdot) \rangle$ only returns b_j for $j \geq 1$ suggests that we revise our expression of a trigonometric polynomial from (42.1) to

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \big(a_k \cos(kx) + b_k \sin(kx)\big).$$

For f in this form, we now have the more consistent formulas

$$a_k = \frac{\langle f, \cos(k\cdot) \rangle}{\pi} \quad \text{and} \quad b_k = \frac{\langle f, \sin(k\cdot) \rangle}{\pi}.$$

This suggests (but does not *demand*) how we might define the coefficients a_k and b_k above in the ideal trigonometric expansion of an arbitrary $f \in \mathcal{C}_{\text{pw}}([-\pi, \pi])$.

42.4 Definition. *Let $f \in \mathcal{C}_{\text{pw}}([-\pi, \pi])$. The* **FOURIER COEFFICIENTS** *of f are*

$$\mathsf{a}_k[f] := \frac{\langle f, \cos(k\cdot) \rangle}{\pi} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) \, dx$$

and

$$\mathsf{b}_k[f] := \frac{\langle f, \sin(k\cdot) \rangle}{\pi} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) \, dx.$$

The **FORMAL FOURIER SERIES** *of f is*

$$\mathsf{FS}[f](x) := \frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{\infty} \big(\mathsf{a}_k[f] \cos(kx) + \mathsf{b}_k[f] \sin(kx)\big).$$

Actually computing the coefficients $\mathsf{a}_k[f]$ and $\mathsf{b}_k[f]$ is mostly an exercise in integration, often integration by parts. As with Fourier transforms, there is usually little insight to be gained in computing explicitly the Fourier coefficients, except maybe for some very special functions; later we will see that *estimating* Fourier coefficients can be more illuminating, and useful.

For $f \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$ and $x \in [-\pi,\pi]$, we can always interpret $\mathsf{FS}[f](x)$ as a sequence of partial sums:

$$\mathsf{FS}[f](x) = \left(\frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{n} \big(\mathsf{a}_k[f]\cos(kx) + \mathsf{b}_k[f]\sin(kx)\big)\right).$$

Ideally we would like to have $\mathsf{FS}[f](x) = f(x)$, i.e.,

$$f(x) = \lim_{n\to\infty} \frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{n} \big(\mathsf{a}_k[f]\cos(kx) + \mathsf{b}_k[f]\sin(kx)\big).$$

Determining if $\mathsf{FS}[f](x)$ converges, and to what (whether $f(x)$ or something else), has been a major thrust of modern analysis. At the very least, if $\mathsf{FS}[f] = f$, then f must be 2π-periodic, i.e., $f(-\pi) = f(\pi)$.

42.5 Problem. Why?

We might wonder if $\mathsf{a}_k[f]$ and $\mathsf{b}_k[f]$ are the "best" coefficients to use if we want to write f as an infinite sum of sines and cosines. For a variety of reasons, they are—not least because under suitable hypotheses on f we will indeed have $\mathsf{FS}[f](x) = f(x)$. Here is a formal indication of why these coefficients are the right ones.

42.6 Problem. Let $f \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$. Suppose that there are sequences (a_k) and (b_k) such that

$$f(x) = \sum_{k=0}^{\infty} \big(a_k\cos(kx) + b_k\sin(kx)\big)$$

for each $x \in [-\pi,\pi]$, and suppose as well that we can interchange integration and summation in the sense that

$$\langle f, g\rangle = \sum_{k=0}^{\infty} \big(a_k\langle\cos(k\cdot), g\rangle + b_k\langle\sin(k\cdot), g\rangle\big)$$

for any $g \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$. Prove that

$$a_0 = 2\mathsf{a}_0[f], \qquad a_k = \mathsf{a}_k[f],\ k \ge 1, \qquad \text{and} \qquad b_k = \mathsf{b}_k[f],\ k \ge 1.$$

And here is a precise indication of why the Fourier coefficients are the right ones.

42.7 Theorem. *Let $f \in \mathcal{C}^1([-\pi,\pi)$ with $f(-\pi) = f(\pi)$ and $f'(-\pi) = f(\pi)$. Then $\mathsf{FS}[f](x) = f(x)$ for all $x \in [-\pi,\pi]$.*

The hypothesis that f be continuously differentiable cannot be relaxed: there exist $f \in \mathcal{C}([-\pi,\pi])$ such that $\mathsf{FS}[f](x)$ diverges at one or more values of x! The periodicity hypotheses $f(-\pi) = f(\pi)$ and $f'(-\pi) = f(\pi)$ are not unnatural if we want f to solve a heat BVP.

42.8 Problem. Suppose that u solves

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t \in \mathbb{R} \\ u(-\pi,t) = u(\pi,t), \ u_x(-\pi,t) = u_x(\pi,t), \ t \in \mathbb{R} \\ u(x,0) = f(x). \end{cases}$$

Prove that $f \in \mathcal{C}^2([-\pi,\pi])$ with

$$f(-\pi) = f(\pi) \quad \text{and} \quad f'(-\pi) = f'(\pi).$$

Here are some useful estimates on the Fourier coefficients, which we state for $\mathsf{a}_k[f]$ only, though they are also true for $\mathsf{b}_k[f]$.

42.9 Problem. **(i)** Let $f \in \mathcal{C}([-\pi,\pi])$ and put $\|f\|_\infty := \max_{-\pi \le x \le \pi} |f(x)|$. Prove that

$$|\mathsf{a}_k[f]| \le 2\pi \|f\|_\infty .$$

(ii) Let $f \in \mathcal{C}^1([-\pi,\pi])$ with $f(-\pi) = f(\pi)$. Prove the existence of $C > 0$ (depending on f) such that for $k \ge 1$,

$$|\mathsf{a}_k[f]| \le \frac{C}{k}.$$

[Hint: *integrate by parts.*]

The second estimate above can be improved with more regularity and periodicity hypotheses on f. For example, if $f \in \mathcal{C}^2([-\pi,\pi])$ with $f(-\pi) = f(\pi)$ and $f'(-\pi) = f'(\pi)$, then

$$|\mathsf{a}_k[f]| \le \frac{C}{k^2} \quad \text{and} \quad |\mathsf{b}_k[f]| \le \frac{C}{k^2}$$

for $k \ge 1$.

42.10 Problem. Explain how this resembles the estimates (29.1) for the Fourier transform. (Maybe try (re)proving those estimates for the case $r = 2$.) When things boil down to integration by parts, how do the periodicity conditions here for Fourier coefficients play the same role as the hypothesis $f^{(j)} \in L^1$, $1 \le j \le r$, for the Fourier transform?

Our hope is now that if $f \in \mathcal{C}^2([-\pi,\pi])$ with $f(-\pi) = f(\pi)$ and $f'(-\pi) = f'(\pi)$, then

$$u(x,t) := \frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{\infty} \big(\mathsf{a}_k[f]\cos(kx) + \mathsf{b}_k[f]\sin(kx)\big)e^{-k^2t}$$

is defined and sufficiently differentiable to solve

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t \in \mathbb{R} \\ u(-\pi,t) = u(\pi,t), \ u_x(-\pi,t) = u_x(\pi,t), \ t \in \mathbb{R} \\ u(x,0) = f(x). \end{cases}$$

We will proceed assuming $\mathsf{a}_0[f] = \mathsf{b}_k[f] = 0$ for $k \ge 0$ and writing just a_k instead of $\mathsf{a}_k[f]$; generalizing to the full version of u above will then pose no real conceptual difficulties.

So, we are considering

$$u(x,t) = \sum_{k=1}^{\infty} a_k \cos(kx)e^{-k^2t}.$$

Does this series converge? The hypotheses on f imply $|a_k| \le C/k^2$, so each term of the series is bounded by

$$|a_k \cos(kx)e^{-k^2t}| \le \frac{C}{k^2}, \tag{42.2}$$

and $\sum_{k=1}^{\infty} k^{-2}$ converges by properties of p-series. The comparison test then implies that u converges.

The estimate (42.2) tacitly assumed $t \ge 0$ to obtain $0 \le e^{-k^2t} \le 1$. This suggests that we will not be able to solve our heat problem for all $t \in \mathbb{R}$ as initially posed (and as was feasible for initial temperature distribution f given by a trigonometric polynomial) but at best for $t \ge 0$. That is, we may have to content ourselves with solving

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t \ge 0 \\ u(-\pi,t) = u(\pi,t), \ u_x(-\pi,t) = u_x(\pi,t), \ t \ge 0 \\ u(x,0) = f(x), \ -\pi \le x \le \pi. \end{cases}$$

We can actually get convergence of $u(x,t)$ for $t > 0$ if the sequence (a_k) is merely bounded, i.e., if there is $C > 0$ such that $|a_k| \le C$ for all k. In this case, we estimate

$$|a_k \cos(kx)e^{-k^2t}| \le Ce^{-k^2t} = Ce^{-kt} = C(e^{-t})^k.$$

since $-k^2t < -kt$ for $k \ge 1$ and $t > 0$. When $t > 0$ we have $0 < e^{-t} < 1$, so the series $\sum_{k=0}^{\infty}(e^{-t})^k$ is a convergent geometric series. Thus $u(x,t)$ converges for all $x \in [-\pi,\pi]$ and $t > 0$. This too suggests that we may have to content ourselves with solving the heat problem for $t > 0$, which was the situation, of course, with the infinite rod heat equation.

As for the derivatives, we might hope that we can "differentiate under the sum" to find

$$u_t(x,t) = \partial_t\left[\sum_{k=1}^{\infty} a_k \cos(kx)e^{-k^2t}\right] = \sum_{k=1}^{\infty} \partial_t\big[a_k\cos(kx)e^{-k^2t}\big] = \sum_{k=1}^{\infty} -k^2a_k\cos(kx)e^{-k^2t}$$

and

$$u_{xx}(x,t) = \partial_x^2 \left[\sum_{k=1}^{\infty} a_k \cos(kx) e^{-k^2 t} \right] = \sum_{k=1}^{\infty} \partial_x^2 \left[a_k \cos(kx) e^{-k^2 t} \right] = \sum_{k=1}^{\infty} -k^2 a_k \cos(kx) e^{-k^2 t}.$$

If these interchanges of sum and derivative are valid, then we certainly have $u_t = u_{xx}$. So (why) does it work?

Day 43: Wednesday, November 20.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Term-by-term differentiation of a series is stated and proved on p. 689 as a corollary of differentiation under the (improper) integral. The remarks on p. 133 discuss how one might approximate an initial temperature distribution by a partial sum of its Fourier series and then use that trigonometric polynomial to build an approximate solution to the related PDE. Example 4 on p. 248 illustrates where the full Fourier series might fail (because the PDE does not really make sense). A nice definition of "formal solution" and "formal" appears on p. 249; this definition of "formal" applies to many situations beyond PDE. See also the remarks on "formal solutions" on p. 252. Pages 482–485 offer different perspectives on solutions to PDE via Fourier series by relating them to solutions obtained via the Fourier transform. Example 1 on p. 142 offers the comforting result that solutions to this periodic heat BVP are unique; the proof offers good practice with energy integrals.

We work on differentiating the function

$$u(x,t) := \sum_{k=1}^{\infty} a_k \cos(kx) e^{-k^2 t}, \tag{43.1}$$

assuming that the sequence (a_k) decays as $|a_k| \le C/k^2$. We will do just the t-partial, as the theory is the same for the x-partials. The trick is to recognize this series as an improper integral and then differentiate under the integral.

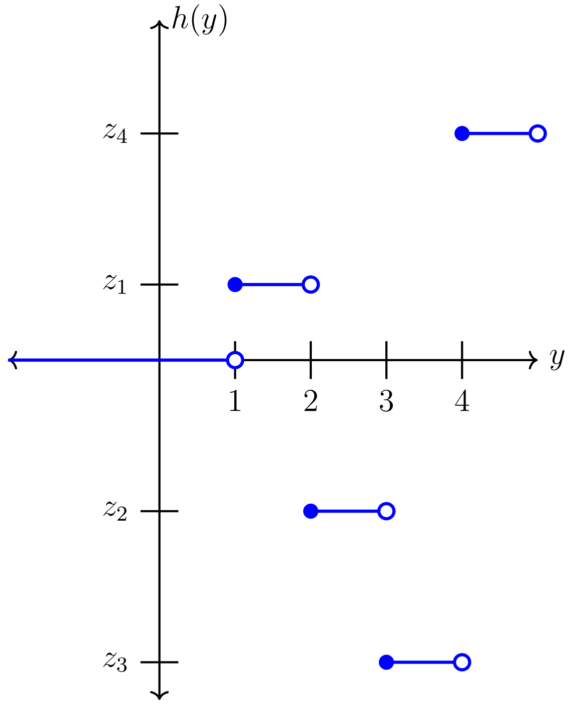
43.1 Lemma. *Let (z_k) be a sequence such that $\sum_{k=1}^{\infty} z_k$ converges. Define*

$$h \colon \mathbb{R} \to \mathbb{C} \colon y \mapsto \begin{cases} 0, & y < 1 \\ z_k, & k \le y < k+1, k \ge 1 \text{ is an integer.} \end{cases}$$

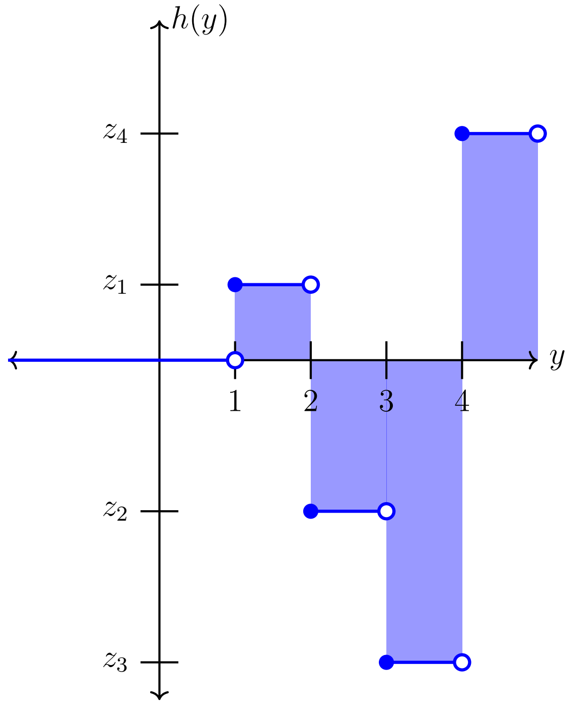
Then

$$\sum_{k=1}^{\infty} z_k = \int_{-\infty}^{\infty} h.$$

Proof. The crux of the proof are the following pictures. First, here is a sample graph of h, which indicates $h \in \mathcal{C}_{\text{pw}}(\mathbb{R})$, which in turn follows directly from the piecewise formula for h.



Second, the (net) area under the graph of h over any interval $[k, k+1]$ for integers $k \geq 1$ is z_k.



Adding everything, we obtain $\int_{-\infty}^{\infty} h = \sum_{k=1}^{\infty} z_k$. ∎

We can therefore rewrite u from (43.1) as

$$u(x,t) = \int_{-\infty}^{\infty} h(x,t,y)\ dy,$$

where

$$h(x,t,y) := \begin{cases} 0, \ y < 1 \\ a_k \cos(kx)e^{-k^2t}, \ k \le y < k+1, \ k \ge 1 \text{ is an integer.} \end{cases}$$

To differentiate h with respect to t, we want to find a "dominating" function $M = M(x,y)$ with $M(x,\cdot) \in L^1$ such that

$$|h_t(x,t,y)| \le |M(x,y)|.$$

We have

$$h_t(x,t,y) = \begin{cases} 0, \ y < 1 \\ -k^2 a_k \cos(kx)e^{-k^2t}, \ k \le y < k+1, \ k \ge 1 \text{ is an integer,} \end{cases}$$

and so we might try to take

$$M(x,y) := \begin{cases} 0, \ y < 1 \\ k^2|a_k|e^{-k^2t}, \ k \le y < k+1, \ k \ge 1 \text{ is an integer.} \end{cases}$$

This is not wholly successful, since M still depends on t. The right idea is to abandon hope of differentiability at $t = 0$ and to focus on $t > 0$. In particular, we will no longer require the boundary condition $u_x(-\pi,0) = u_x(\pi,0)$.

Let $n \ge 1$ be an integer and consider $t \ge 1/n$. By running through all integers n, we can cover all $t \in (0,\infty)$. Then with $t \ge 1/n$, we obtain

$$|h_t(x,t,y)| \le \begin{cases} 0, \ y < 1 \\ k^2|a_k|e^{-k^2/n}, \ k \le y < k+1, \ k \ge 1 \text{ is an integer.} \end{cases}$$

We are still assuming $|a_k| \le C/k^2$, so putting

$$M(x,y) := \begin{cases} 0, \ y < 1 \\ Ce^{-k^2/n}, \ k \le y < k+1, \ k \ge 1 \text{ is an integer} \end{cases}$$

seems like a good idea. Since $\sum_{k=1}^{\infty} e^{-k^2/n}$ converges with $n \ge 1$ fixed (by comparison to the geometric series), it follows that $M(x,\cdot) \in L^1$. This is enough to permit us to differentiate under the integral.

Here are the summaries of all of our work. First we state a result for the periodic finite heat equation, and we include a "continuity as $t \to 0^+$" condition similar to the one for the infinite heat equation.

43.2 Theorem. *Let $f \in \mathcal{C}^2([-\pi,\pi])$ with $f(-\pi) = f(\pi)$ and $f'(-\pi) = f'(\pi)$. The function*

$$u(x,t) := \frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{\infty} \big(\mathsf{a}_k[f]\cos(kx) + \mathsf{b}_k[f]\sin(kx)\big)e^{-k^2t}$$

is defined for $-\pi \le x \le \pi$ and $t \ge 0$ and solves

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t > 0 \\ u(-\pi,t) = u(\pi,t), \ u_x(-\pi,t) = u_x(\pi,t), \ t \ge 0 \\ u(x,0) = f(x), \ -\pi \le x \le \pi \\ \lim_{(s,t)\to(x,0^+)} u(s,t) = f(x), \ -\pi \le x \le \pi. \end{cases}$$

43.3 Problem. Explain why the boundary condition $u_x(-\pi,t) = u_x(\pi,t)$ holds for $t > 0$.

Next, we summarize our "differentiation under the sum" result.

43.4 Theorem. *Let $I \subseteq \mathbb{R}$ be an interval. For each integer $k \ge 0$, let $f_k \in \mathcal{C}^1(I)$ such that the following hold.*

(i) *$f(s) := \sum_{k=0}^{\infty} f_k(s)$ converges for each $s \in I$.*

(ii) *For each $k \ge 0$, there is $M_k \ge 0$ such that $|f_k'(s)| \le M_k$ for all $s \in I$.*

(iii) *$\sum_{k=0}^{\infty} M_k$ converges.*

Then $f \in \mathcal{C}^1(I)$ and

$$f'(s) = \sum_{k=0}^{\infty} f_k'(s).$$

We could develop a very similar result for the problem

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le \pi, \ t > 0 \\ u(0,t) = u(\pi,t) = 0, \ t \ge 0 \\ u(x,0) = f(x), \ 0 \le x \le \pi, \end{cases}$$

assuming that we had a valid expansion $f(x) = \sum_{k=1}^{\infty} b_k \sin(kx)$. We will not do this, as it would not teach us anything particularly new about PDE or analysis.

Rather, the more interesting mathematical question here is not about u but about f. What can we say about $\mathsf{FS}[f](x)$ for more general $f \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$? As we have noted, continuity alone is not enough to guarantee any kind of convergence. However, with slightly stronger assumptions we have the following analogue of Fourier inversion (Theorem 29.9). We need a variant on the function space $\mathcal{C}_{\text{pw}}^1(\mathbb{R})$ from Definition 29.8. Recall that we use the notation

$$f(x^{\pm}) := \lim_{s \to x^{+} \pm} f(s)$$

for a function f and a point x.

43.5 Definition. *Let a, $b \in \mathbb{R}$ with $a < b$. Let $f \in \mathcal{C}_{\mathrm{pw}}([a,b])$ and suppose that*

$$\lim_{h \to 0^{+}} \frac{f(x+h) - f(x^{+})}{h} \qquad \text{and} \qquad \lim_{h \to 0^{-}} \frac{f(x+h) - f(x^{-})}{h}$$

exist for all $x \in (a,b)$ and that they are equal for all but finitely many points in (a,b). Suppose also that the limits

$$\lim_{h \to 0^{+}} \frac{f(a+h) - f(a^{+})}{h} \qquad \text{and} \qquad \lim_{h \to 0^{-}} \frac{f(b+h) - f(b^{-})}{h}$$

exist. (In the expressions $f(x+h)$, $f(a+h)$, and $f(b+h)$ we are assuming that $h \neq 0$ is so small that f is actually defined at these points; this is possible, since f is undefined for at most finitely many points in $[a,b]$ by definition of $\mathcal{C}_{\mathrm{pw}}([a,b])$.) Then we say that f is PIECEWISE CONTINUOUSLY DIFFERENTIABLE *on $[a,b]$, and we denote the set of all such functions by $\mathcal{C}^{1}_{\mathrm{pw}}([a,b])$.*

Here is the best that we can say about the convergence of Fourier series in general.

43.6 Theorem. *Let $f \in \mathcal{C}^{1}_{\mathrm{pw}}([-\pi,\pi])$. Then*

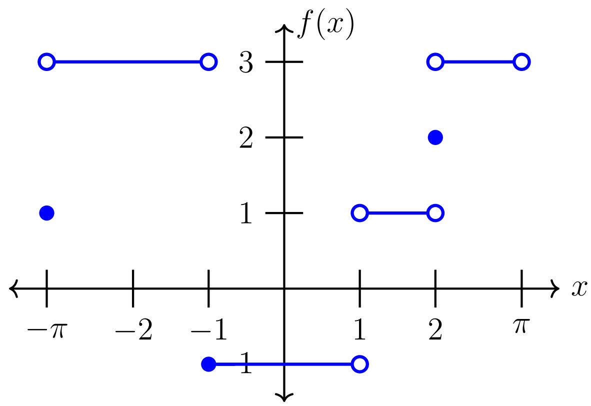
$$\mathsf{FS}[f](x) = \begin{cases} \dfrac{f(x^{+}) + f(x^{-})}{2}, & x \in (-\pi,\pi) \\[2ex] \dfrac{f(-\pi^{+}) + f(\pi^{-})}{2}, & x = \pm\pi. \end{cases}$$

Except at the endpoints $\pm\pi$, this theorem is effectively the analogue of the Fourier inversion result in Theorem 29.9. Now is a good time to point out that the Fourier coefficients (or "modes") $\mathsf{a}_k[f]$ and $\mathsf{b}_k[f]$ are the analogues of the Fourier coefficients $\widehat{f}(k)$, while the Fourier series is the analogue of the inverse Fourier transform.

43.7 Example. We define $f \in \mathcal{C}_{\text{pw}}([-\pi, \pi])$ by

$$f(x) = \begin{cases} 1, & x = -\pi \\ 3, & -\pi < x < -1 \\ -1, & -1 \le x < 1 \\ 1, & 1 \le x < 2 \\ 2, & x = 2 \\ 3, & 2 < x < \pi. \end{cases}$$

Note that f is not even defined at $x = 1$ and $x = \pi$. Here is the graph.



By continuity, we have $\mathsf{FS}[f](x)$ for $-\pi < x < -1$, $-1 < x < 1$, $1 < x < 2$, and $2 < x < \pi$. At $x = \pm\pi$, we have

$$\mathsf{FS}[f](x) = \frac{f(-\pi^+) + f(\pi^-)}{2} = \frac{3+3}{2} = 3.$$

Otherwise, we have

$$\mathsf{FS}[f](-1) = \frac{f(-1^-) + f(-1^+)}{2} = \frac{3 + (-1)}{2} = 1,$$

$$\mathsf{FS}[f](1) = \frac{f(1^-) + f(1^+)}{2} = \frac{-1+1}{2} = 0,$$

and

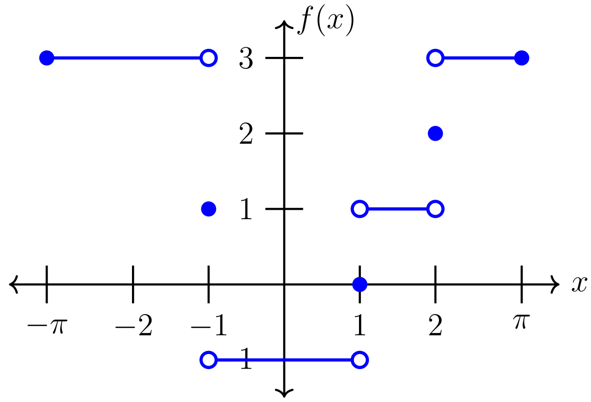
$$\mathsf{FS}[f](2) = \frac{f(2^-) + f(2^+)}{2} = \frac{1+3}{2} = 2.$$

That is,

$$\mathsf{FS}[f](x) = \begin{cases} 3, & -\pi \le x < -1 \\ 1, & x = -1 \\ -1, & -1 < x < 1 \\ 0, & x = 1 \\ 1, & 1 < x < 2 \\ 2, & x = 2 \\ 3, & 2 \le x \le \pi. \end{cases}$$

Among other things, we might note that $\mathsf{FS}[f]$ is defined on $[-\pi, \pi]$, even though f is undefined at some points in that interval. This is just like how a function's Fourier transform can be defined on $(-\infty, \infty)$ even if the function is undefined at some real numbers. *Integrals can eat and/or forgive bad behaviors at single points.*

Here is the graph of $\mathsf{FS}[f]$.



Day 44: Friday, November 22.

Material from *Basic Partial Differential Equations* by Bleecker & Csordas

Bessel's inequality and the Riemann–Lebesgue lemma are discussed on pp. 209–212 (in particular, see the illuminating Figure 2 on p. 212). The proof of Bessel's inequality on pp. 209–210 includes the "best approximation" result.

We do one example of an honest-to-goodness Fourier series calculation, in part to illustrate the annoyances underlying such calculations, but really to see a "smoothing" effect akin to that of the infinite heat equation.

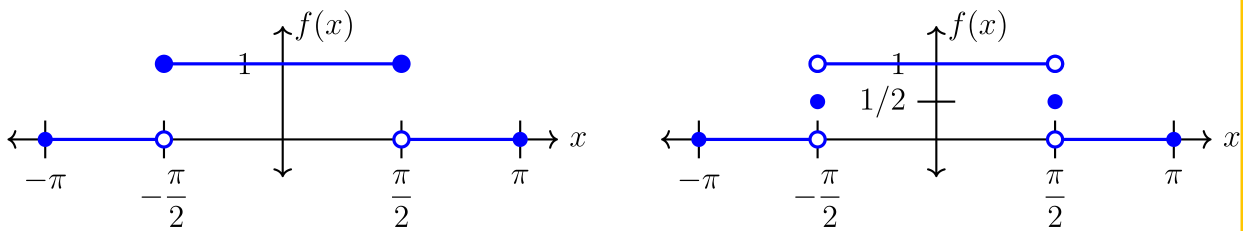
44.1 Example. We want to solve

$$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t > 0 \\ u(-\pi, t) = u(\pi, t), \ u_x(-\pi, t) = u_x(\pi, t) \ t > 0 \\ u(x, 0) = f(x), \ -\pi \le x \le \pi, \end{cases}$$

where

$$f \colon [-\pi, \pi] \to \mathbb{R} \colon x \mapsto \begin{cases} 1, & |x| \le \pi/2 \\ 0, & |x| > \pi/2. \end{cases}$$

Here are the graphs of f and $\mathsf{FS}[f]$.



We have the following convergence of $\mathsf{FS}[f]$:

$$\mathsf{FS}[f](x) = \begin{cases} 0, & -\pi \leq x < -\pi/2 \\ 1/2, & x = -\pi/2 \\ 1, & -\pi/2 < x < \pi/2 \\ 1/2, & x = \pi/2 \\ 0, & \pi/2 < x \leq \pi. \end{cases}$$

We begin by computing the Fourier coefficients of f. We have

$$\mathsf{a}_k[f] = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\ dx = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} \cos(kx)\ dx.$$

At $k = 0$, this reduces to $\mathsf{a}_0[f] = 1$. Otherwise, for $k \geq 1$, we have

$$\mathsf{a}_k[f] = \frac{\sin(kx)}{k\pi}\bigg|_{x=-\pi/2}^{x=\pi/2} = \frac{1}{k\pi}\left[\sin\left(\frac{k\pi}{2}\right) - \sin\left(-\frac{k\pi}{2}\right)\right] = \frac{2}{k\pi}\sin\left(\frac{k\pi}{2}\right).$$

If k is even, then $k/2 \in \mathbb{Z}$, and so $\sin(k\pi/2) = 0$. If k is odd, then $k = 2j+1$ for some $j \in \mathbb{Z}$, and thinking about the unit circle returns

$$\sin\left(\frac{(2j+1)\pi}{2}\right) = (-1)^j.$$

Thus

$$\mathsf{a}_{2j}[f] = 0,\ j \geq 1, \quad \text{and} \quad \mathsf{a}_{2j+1}[f] = \frac{2(-1)^j}{(2j+1)\pi}.$$

Last, since f is even and $\sin(k\cdot)$ is odd, we have

$$\mathsf{b}_k[f] = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\ dx = 0.$$

All together,

$$\mathsf{FS}[f](x) = \frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{\infty}\big(\mathsf{a}_k[f]\cos(kx) + \mathsf{b}_k[f]\sin(kx)\big) = \frac{1}{2} + \sum_{j=0}^{\infty}\frac{2(-1)^j}{(2j+1)\pi}.$$

We therefore expect that our solution u is

$$u(x,t) = \frac{1}{2} + \sum_{j=0}^{\infty} \frac{2(-1)^j}{(2j+1)\pi} \cos((2j+1)x)e^{-(2j+1)^2 t}.$$

We can check that u is sufficiently differentiable for $t > 0$ using Theorem 43.4. This mostly amounts to considering, more abstractly, a function v of the form $v(x,t) = \sum_{k=0}^{\infty} a_k \cos(kx)e^{-k^2 t}$ where there is $C > 0$ such that $|a_k| \le C$ for all k and checking the convergence of $\sum_{k=0}^{\infty} k^2 e^{-k^2 t}$, which holds for $t > 0$.

Here is the point: although the initial data f has discontinuities, the solution u is sufficiently continuously differentiable for all $t > 0$. *Once again, the heat equation smooths.*

44.2 Problem. Prove that $\sum_{k=0}^{\infty} k^2 e^{-k^2 t}$ converges for all $t > 0$.

Now we return to our earlier question about the convergence of $\mathsf{FS}[f]$ for $f \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$. For each $x \in [-\pi,\pi]$, the Fourier series $\mathsf{FS}[f](x)$ makes sense as a sequence of partial sums, but without further hypotheses on f, it need not converge to $f(x)$, or even converge at all. However, we can say something profound about convergence "on average."

44.3 Theorem. *Let $f \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$. Then*

$$\lim_{n\to\infty} \left\| f - \left(\frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{n} \big(\mathsf{a}_k[f]\cos(k\cdot) + \mathsf{b}_k[f]\sin(k\cdot)\big) \right) \right\| = 0. \tag{44.1}$$

That is, "on average" the nth partial sum of the Fourier series for f becomes "very close" to f when n is "large." We now have three results on Fourier series convergence: this, Theorem 43.6, and Theorem 42.7. We have gradually weakened our hypotheses on f, and, in some sense, our results have weakened. Theorem 44.3 no longer guarantees that $\mathsf{FS}[f](x)$ converges to $f(x)$ or even converges as a series at all.

44.4 Problem. (i) Define

$$g_n \colon [-\pi,\pi] \to \mathbb{R} \colon x \mapsto \left(\frac{x}{\pi}\right)^n.$$

Compute

$$\lim_{n\to\infty} \|g_n\| \qquad \text{and} \qquad \lim_{n\to\infty} g_n(\pi).$$

(ii) Define

$$g_n \colon [-\pi,\pi] \to \mathbb{R} \colon x \mapsto \begin{cases} 0, & -\pi \le x < 0 \\ \sqrt{n}, & 0 \le x < 1/n \\ 0, & 1/n \le x \le \pi. \end{cases}$$

Compute

$$\lim_{n\to\infty} \|g_n\| \quad \text{and} \quad \lim_{n\to\infty} g_n(0).$$

(iii) Define

$$g_n\colon [-\pi,\pi] \to \mathbb{R}\colon x \mapsto \begin{cases} 0, & -\pi \le x < 1/n \\ \sqrt{n}, & 1/n \le x < 2/n \\ 0, & 2/n \le x \le \pi. \end{cases}$$

Compute

$$\lim_{n\to\infty} \|g_n\| \quad \text{and} \quad \lim_{n\to\infty} g_n(x), \quad \text{with } x \in [-\pi,\pi] \text{ fixed.}$$

(iv) Define

$$g_n\colon [-\pi,\pi] \to \mathbb{R}\colon x \mapsto \begin{cases} 0, & -\pi \le x < 1/n \\ n, & 1/n \le x < 2/n \\ 0, & 2/n \le x \le \pi. \end{cases}$$

Compute

$$\lim_{n\to\infty} \|g_n\| \quad \text{and} \quad \lim_{n\to\infty} g_n(x), \quad \text{with } x \in [-\pi,\pi] \text{ fixed.}$$

(v) For integers $n \ge 0$ let $g_n \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$ and let $g \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$. Explain how the previous parts show that knowing $\lim_{n\to\infty} g_n(x) = g(x)$ for all x implies nothing about the value, or even existence, of $\lim_{n\to\infty} \|g_n - g\|$. Explain how likewise knowing $\lim_{n\to\infty} \|g_n - g\| = 0$ implies nothing about the value, or even existence, of $\lim_{n\to\infty} g_n(x)$.

Measuring convergence of Fourier series through the norm $\|\cdot\|$ also provides more insight as to why we chose the coefficients $\mathsf{a}_k[f]$ and $\mathsf{b}_k[f]$, beyond the established facts that "they work." The reason is that they "really" work.

To understand this, it will help to have some more compact notation. Recall that for f, $g \in \mathcal{C}_{\text{pw}}([-\pi,\pi])$, we have written

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg,$$

and now, as with our prior work on heat boundary value problems, we put

$$\|f\| := \sqrt{\langle f, f \rangle} = \left(\int_{-\pi}^{\pi} [f(x)]^2 \, dx \right)^{1/2}.$$

Fourier series involve infinite sums of functions of the form $\cos(k\cdot)$ and $\sin(k\cdot)$. Each set $\{\cos(k\cdot)\}_{k=0}^{\infty}$ and $\{\sin(k\cdot)\}_{k=1}^{\infty}$ is countably infinite, and so their union is countably infinite as well. For convenience, we rescale these functions by their norms and define

$$\mathcal{U} := \left\{ \frac{\cos(k\cdot)}{\|\cos(k\cdot)\|} \right\}_{k=0}^{\infty} \cup \left\{ \frac{\sin(k\cdot)}{\|\sin(k\cdot)\|} \right\}_{k=1}^{\infty}.$$

We make things more abstract and just write

$$\mathcal{U} = \{\phi_j\}_{j=1}^{\infty}.$$

That is, for each $k \geq 1$, either there is $j \geq 0$ such that $\phi_k = \cos(j\cdot)/\left\|\cos(j\cdot)\right\|$, or there is $\ell \geq 1$ such that $\phi_k = \sin(\ell\cdot)/\left\|\sin(\ell\cdot)\right\|$; conversely, each function of the form $\cos(j\cdot)/\left\|\cos(j\cdot)\right\|$ or $\sin(\ell\cdot)/\left\|\sin(\ell\cdot)\right\|$ equals one of these ϕ_k.

44.5 Problem. **(i)** Explain why $\mathcal{U}$ is ORTHONORMAL in the sense that

$$\langle f, g\rangle = \begin{cases} 1, & f = g \\ 0, & f \neq g \end{cases}$$

for all f, $g \in \mathcal{E}$.

(ii) Let $f \in \mathcal{C}_{\mathrm{pw}}([-\pi, \pi])$. Show that

$$\{\mathsf{a}_k[f]\}_{k=0}^{\infty} \cup \{\mathsf{b}_k[f]\}_{k=0}^{\infty} = \{\langle f, \phi_j\rangle\}_{j=1}^{\infty}.$$

(iii) Explain why the limit (44.1) holds if and only if

$$\lim_{m\to\infty} \left\| f - \sum_{j=1}^{m} \langle f, \phi_j\rangle\, \phi_j \right\| = 0.$$

Now here is how the Fourier coefficients "really" work. Not only is the nth partial sum of the Fourier series for f a "good" approximation to f "on average" when n is "large," this partial sum is the *best* approximation to f out of all candidate functions in a certain class.

44.6 Theorem. *Let $f \in \mathcal{C}_{\mathrm{pw}}([-\pi, \pi])$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$. Then*

$$\left\| f - \sum_{j=1}^{m} \langle f, \phi_j\rangle\, \phi_j \right\| \leq \left\| f - \sum_{j=1}^{m} \alpha_j \phi_j \right\|.$$

Moreover, the only choice of $\alpha_1, \ldots, \alpha_m$ that achieves equality above is $\alpha_j = \langle f, \phi_j\rangle$.

Proof. We prove the equivalent statement

$$\left\| f - \sum_{j=1}^{m} \langle f, \phi_j\rangle\, \phi_j \right\|^2 \leq \left\| f - \sum_{j=1}^{m} \alpha_j \phi_j \right\|^2,$$

which removes the pesky square roots from $\|\cdot\|$. We may compute

$$\left\| f - \sum_{j=1}^{m} \alpha_j \phi_j \right\|^2 = \|f\|^2 + \sum_{j=1}^{m} \left(\alpha_j^2 - 2\alpha_j \langle f, \phi_j\rangle\right) \tag{44.2}$$

and then "complete the square" in each of the j terms:

$$\alpha_j^2 - 2\alpha_j \langle f, \phi_j\rangle = \left(\alpha_j^2 - 2\alpha_j \langle f, \phi_j\rangle + |\langle f, \phi_j\rangle|^2\right) - |\langle f, \phi_j\rangle|^2 = |\alpha_j - \langle f, \phi_j\rangle|^2 - |\langle f, \phi_j\rangle|^2.$$

Thus

$$\left\| f - \sum_{j=1}^{m} \alpha_j \phi_j \right\|^2 = \|f\|^2 + \sum_{j=1}^{m} |\alpha_j - \langle f, \phi_j \rangle|^2 - \sum_{j=1}^{m} |\langle f, \phi_j \rangle|^2.$$

The middle term is the sum of nonnegative terms $|\alpha_j - \langle f, \phi_j \rangle|^2$ and therefore nonnegative itself, which implies

$$\left\| f - \sum_{j=1}^{m} \alpha_j \phi_j \right\|^2 \geq \|f\|^2 - \sum_{j=1}^{m} |\langle f, \phi_j \rangle|^2 =: M. \tag{44.3}$$

The quantity M is a lower bound on $\left\| f - \sum_{j=1}^{m} \alpha_j \phi_j \right\|^2$. Therefore, taking $\alpha_j = \langle f, \phi_j \rangle$ gives

$$\left\| f - \sum_{j=1}^{m} \langle f, \phi_j \rangle \phi_j \right\|^2 = M. \tag{44.4}$$

That is,

$$\left\| f - \sum_{j=1}^{m} \langle f, \phi_j \rangle \phi_j \right\|^2 = M \leq \left\| f - \sum_{j=1}^{m} \alpha_j \phi_j \right\|^2$$

for any other choice of α_j. ■

44.7 Problem. Work through the details of computing (44.2). [Hint: *it will help to use different letters for the indices of summation in different slots, so write*

$$\left\| f - \sum_{j=1}^{m} \langle f, \phi_j \rangle \phi_j \right\|^2 = \left\langle f - \sum_{j=1}^{m} \langle f, \phi_j \rangle \phi_j, f - \sum_{\ell=1}^{m} \langle f, \phi_\ell \rangle \phi_\ell \right\rangle.$$

Note also that $\langle g, h \rangle = \overline{\langle h, g \rangle}$.] Highlight exactly where you use the orthonormality of $\{\phi_j\}_{j=1}^{\infty}$.

44.8 Problem. Combine (44.3) and (44.4) to conclude

$$\sum_{j=1}^{m} |\langle f, \phi_j \rangle|^2 \leq \|f\|^2$$

for all m. It is a fact that if (z_j) is a sequence in $\mathbb{C}$, and there is $C > 0$ such that $\sum_{j=1}^{m} |z_j| \leq C$ for all $m \geq 1$, then $\sum_{j=1}^{\infty} |z_j|$ converges with $\sum_{j=1}^{\infty} |z_j| \leq C$. Use this to prove that $\sum_{j=1}^{\infty} |\langle f, \phi_j \rangle|^2$ converges and that

$$\sum_{j=1}^{\infty} |\langle f, \phi_j \rangle|^2 \leq \|f\|^2. \tag{44.5}$$

Deduce from the test for convergence the **Riemann–Lebesgue lemma** for Fourier series coefficients:

$$\lim_{j\to\infty} \langle f, \phi_j \rangle = 0.$$

The inequality (44.5) is **Bessel's inequality**. This is in fact not a *strict* inequality but rather an *equality*, called **Parseval's equality**, although that is harder to prove.

44.9 Problem. Assuming Parseval's equality to be true, prove Theorem 44.3. [Hint: *again, combine* (44.3) *and* (44.4) *and use the definition of the convergence of the series* $\sum_{j=1}^{\infty} |\langle f, \phi_j \rangle|^2$ *to* $\|f\|^2$.]

Day 45: Monday, December 2.

Throughout this course, we have paid special attention to how techniques and tools from calculus (and analysis) arise in solving—and, more importantly, *understanding*—PDE. Chief among these tools was the integral as a means for representing functions and extracting data about functions. The integral still plays a critical role in our current work on Fourier series (by extracting, as it did with the Fourier transform, the Fourier modes of a function), but now we are moving from an *analytic*-focused approach to a *linear analytic* approach. Introducing more perspectives from (infinite-dimensional) linear algebra will make clear what really matters.

We have considered two boundary value problems for the finite rod heat equation. In each case, we separated variables, and the crux of that ansatz involved solving a second-order linear ODE subject to boundary conditions. Here is a summary of those results.

Dirichlet boundary conditions	***Periodic boundary conditions***
$\begin{cases} u_t = u_{xx}, \ 0 \le x \le \pi, \ t > 0 \\ u(0,t) = u(\pi,t) = 0, \ t \ge 0 \\ u(x,0) = f(x), \ 0 \le x \le \pi \end{cases}$	$\begin{cases} u_t = u_{xx}, \ -\pi \le x \le \pi, \ t > 0 \\ u(-\pi,t) = u(\pi,t), \ u_x(-\pi,t) = u_x(\pi,t), \ t > 0 \\ u(x,0) = f(x), \ -\pi \le x \le \pi \end{cases}$
Product ansatz: $u(x,t) = X(x)T(t)$	
$\begin{cases} -X'' = \lambda X \text{ on } [0,\pi] \\ X(0) = X(\pi) = 0 \end{cases}$	$\begin{cases} -X'' = \lambda X \text{ on } [-\pi,\pi] \\ X(-\pi) = X(\pi), \ X'(-\pi) = X'(\pi) \end{cases}$
Eigenvalues and eigenfunctions	
$\begin{cases} \lambda = k^2, \ k \ge 1 \\ X(x) = b_k \sin(kx) \end{cases}$	$\begin{cases} \lambda = k^2, \ k \ge 0 \\ X(x) = a_k \cos(kx) + b_k \sin(kx) \end{cases}$
Eigenfunction series expansion (ideal)	
$f(x) = \sum_{k=1}^{\infty} b_k \sin(kx)$	$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \big(a_k \cos(kx) + b_k \sin(kx)\big)$

45.1 Problem. We never discussed the validity of a "Fourier sine series" $f(x) = \sum_{k=1}^{\infty} b_k \sin(kx)$ for $f \in \mathcal{C}_{\text{pw}}([0,\pi])$. Such a series would have to be odd, so that suggests extending $f \in \mathcal{C}_{\text{pw}}([0,\pi])$ to $[-\pi,\pi]$ via

$$f_o(x) := \begin{cases} f(x), & 0 \le x \le \pi \\ -f(-x), & -\pi \le x < 0. \end{cases}$$

Define the **Fourier sine series** of $f \in \mathcal{C}_{\text{pw}}([0,\pi])$ to be $\mathsf{FSS}[f](x) := \mathsf{FS}[f_o](x)$.

(i) Determine coefficients b_k such that $\mathsf{FSS}[f](x) = \sum_{k=1}^{\infty} b_k \sin(kx)$.

(ii) Discuss the convergence of $\mathsf{FSS}[f](x)$. Specifically, develop results analogous to Theorems 42.7, 43.6, and 44.3.

45.2 Problem. Explain how looking for product solutions to the "insulated rod" problem

$$\begin{cases} u_t = u_{xx}, & 0 \le x \le \pi, \ t > 0 \\ u_x(0,t) = u_x(\pi,t) = 0, & t > 0 \\ u(x,0) = f(x), & 0 \le x \le \pi \end{cases}$$

motivates a "Fourier cosine series" $\mathsf{FCS}[f](x) = a_0/2 + \sum_{k=1}^{\infty} a_k \cos(kx)$. Specify what boundary conditions arise in the product ansatz. Give formulas for a_k and develop convergence results analogous to Theorems 42.7, 43.6, and 44.3. The rod is "insulated" because the boundary conditions prevent any flow of heat into or out of the rod at the left and right ends.

Now we isolate and abstract the key elements of these constructions. We first need an environment in which to work.

45.3 Definition. *A* **vector space over** $\mathbb{C}$ *is a nonempty set* $\mathcal{V}$ *such that for each* f, $g \in \mathcal{V}$ *and* $\alpha \in \mathbb{C}$, *there exist elements* $f + g \in \mathcal{V}$ *(the* **vector addition** *of* f *and* g*) and* $\alpha f \in \mathcal{V}$ *(the* **scalar multiplication** *of* α *and* f*) that satisfy the "usual" rules of arithmetic.*

Axioms for vector addition.

1. *Commutativity:* $v + w = w + v$ *for all* v, $w \in \mathcal{V}$.

2. *Associativity:* $v + (w + u) = (v + w) + u$ *for all* v, w, $u \in \mathcal{V}$.

3. *Identity: there exists* $0 \in \mathcal{V}$ *such that* $v + 0 = v$ *for all* $v \in \mathcal{V}$.

4. *Inverse: for each* $v \in \mathcal{V}$, *there exists* $-v \in \mathcal{V}$ *such that* $v + (-v) = 0$.

Axioms for scalar multiplication.

5. *Identity:* $1v = v$ *for all* $v \in \mathcal{V}$.

6. *Associativity:* $\alpha(\beta v) = (\alpha\beta)v$ *for all* α, $\beta \in \mathbb{C}$ *and* $v \in \mathcal{V}$.

Axioms relating vector addition and scalar multiplication.

7. *Distributivity:* $(\alpha + \beta)v = (\alpha v) + (\beta v)$ *for all* α, $\beta \in \mathbb{C}$ *and* $v \in \mathcal{V}$.

8. *Distributivity again:* $\alpha(v + w) = (\alpha v) + (\beta w)$ *for all* $\alpha \in \mathbb{C}$ *and* v, $w \in \mathcal{V}$.

45.4 Example. **(i)** Most of the function classes that we have so far considered are vector spaces, including $\mathcal{C}^r(I)$ for any interval $I \subseteq \mathbb{R}$, $\mathcal{C}_{\text{pw}}([a,b])$, $\mathcal{C}_{\text{pw}}(\mathbb{R})$, $\mathcal{C}^1_{\text{pw}}([a,b])$, $\mathcal{C}^1_{\text{pw}}(\mathbb{R})$, L^1, and L^∞. For each class, most of the vector space axioms follow at once because they are inherited from function addition and multiplication of functions by constants, and all of that arithmetic behaves "as usual" because it in turn is inherited from arithmetic in $\mathbb{C}$. Typically the hardest part is proving that, with $f + g$ defined pointwise as $(f + g)(x) = f(x) + g(x)$, we have $f + g \in \mathcal{V}$ for all f, $g \in \mathcal{V}$.

(ii) For the purposes of Fourier analysis, we have been working with spaces that encode boundary conditions, like

$$\left\{f \in \mathcal{C}^2([0,\pi]) \;\middle|\; f(0) = f(\pi) = 0\right\} \tag{45.1}$$

and

$$\left\{f \in \mathcal{C}^2([-\pi,\pi]) \;\middle|\; f(-\pi) = f(\pi),\;\; f'(-\pi) = f'(\pi)\right\}. \tag{45.2}$$

(iii) More generally, we could work with the so-called REGULAR STURM–LIOUVILLE BOUNDARY CONDITIONS and consider the space

$$\left\{f \in \mathcal{C}^2([a,b]) \;\middle|\; c_1 f(a) + c_2 f'(a) = 0,\;\; c_3 f(b) + c_4 f'(b) = 0\right\}$$

with

$$c_1^2 + c_2^2 \neq 0 \quad \text{and} \quad c_3^2 + c_4^2 \neq 0.$$

or with PERIODIC STURM–LIOUVILLE BOUNDARY CONDITIONS via

$$\left\{f \in \mathcal{C}^2([a,b]) \;\middle|\; f(a) = f(b),\;\; f'(a) = f'(b)\right\}.$$

We add little extra difficulty by allowing the functions in these spaces to be complex-valued, and so we do.

45.5 Problem. Show that the space (45.1) satisfies regular Sturm–Liouville boundary conditions and (45.2) satisfies periodic conditions.

Next, we need additional structure in our environment: a tool for extracting data about vectors.

45.6 Definition. *Let $\mathcal{V}$ be a vector space over $\mathbb{C}$. An* **INNER PRODUCT** *on $\mathcal{V}$ is a function*

$$\langle \cdot, \cdot \rangle : \{(f, g) \mid f,\ g \in \mathcal{V}\} \to \mathbb{C}$$

such that the following hold.

(i) [Distributivity] $\langle f_1 + f_2, g\rangle = \langle f_1, g\rangle + \langle f_2, g\rangle$ *for all* $f_1,\ f_2,\ g \in \mathcal{V}$.

(ii) [Homogeneity] $\langle \alpha f, g\rangle = \alpha \langle f, g\rangle$ *for all* $\alpha \in \mathbb{C}$ *and* $f,\ g \in \mathcal{V}$.

(iii) [Conjugacy] $\overline{\langle f, g\rangle} = \langle g, f\rangle$ *for all* $f,\ g \in \mathcal{V}$.

(iv) [Nonnegativity] $\langle f, g\rangle \geq 0$ *for all* $f \in \mathcal{V}$.

(v) [Definiteness] *If* $\langle f, f\rangle = 0$, *then* $v = 0$.

If an inner product is defined on a vector space $\mathcal{V}$, then we call $\mathcal{V}$ an **INNER PRODUCT SPACE** *relative to that inner product.*

45.7 Example. The primary inner product that we will employ is

$$\langle f, g\rangle := \int_a^b f(x)\overline{g(x)}\ dx,$$

defined on $\mathcal{C}([a,b])$. All of the axioms are easy to check from the algebra of integrals; perhaps the most thought is needed for definiteness, but we have done that many times. Indeed, if $\langle f, f\rangle = 0$, then $\int_a^b |f(x)|^2\ dx = 0$; the integrand is nonnegative, so we know well that $|f(x)|^2 = 0$ for all x, and thus $f = 0$.

45.8 Problem. **(i)** Why does defining

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

not give an inner product on $\mathcal{C}^1([0,1])$?

(ii) Suppose that we remove the conjugacy axiom from the definition of the inner product but keep all of the others. What contradiction results from considering $\langle if, if\rangle$?

Now we begin to extract data about vectors from inner products. The nicest vectors in inner product spaces have very simple behaviors relative to the inner product.

45.9 Definition. *Let $\mathcal{V}$ be an inner product space.*

(i) *A subset $\mathcal{U} \subseteq \mathcal{V}$ is* **ORTHOGONAL** *if* $\langle f, g\rangle = 0$ *for all* $f,\ g \in \mathcal{U}$ *with* $f \neq g$.

(ii) *A subset* $\mathcal{U} \subseteq \mathcal{V}$ *is* **ORTHONORMAL** *if*

$$\langle f, g \rangle = \begin{cases} 1, & f = g \\ 0, & f \neq g. \end{cases}$$

45.10 Example. The sets $\{\sin(k\cdot)\}_{k\in\mathbb{Z}}$, $\{\cos(k\cdot)\}_{k\in\mathbb{Z}}$, and $\{\sin(k\cdot)\}_{k\in\mathbb{Z}} \cup \{\cos(k\cdot)\}_{k\in\mathbb{Z}}$ are orthogonal but not orthonormal in $\mathcal{C}([-\pi,\pi])$ with the inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx$.

45.11 Problem. Which of the three sets of functions from the previous example are orthogonal in $\mathcal{C}([0,\pi])$ with inner product $\langle f, g \rangle = \int_0^{\pi} f(x)\overline{g(x)}\, dx$?

45.12 Remark. *We adopt the following convention: if* $\mathcal{V}$ *is an inner product space and* $\mathcal{U} \subseteq \mathcal{V}$ *is orthonormal and finite, then when we write* $\mathcal{U} = \{\phi_k\}_{k=1}^m$*, we assume that the* ϕ_k *are distinct. That is,* $\phi_j \neq \phi_k$ *for* $j \neq k$*, and so when we say that* $\{\phi_k\}_{k=1}^m$ *is orthonormal, we have*

$$\langle \phi_k, \phi_j \rangle = \begin{cases} 1, & k = j \\ 0, & k \neq j. \end{cases}$$

This avoids irritating and awkward redundancies with labeling the same element of a set twice. For example, with $\mathcal{V} = \mathcal{C}([-\pi,\pi])$ *and inner product* $\langle f, g \rangle = \int_{-\pi}^{\pi} f\overline{g}$*, the set* $\mathcal{U} := \{1/\sqrt{2\pi}, \cos(\cdot)/\pi\}$ *is orthonormal. But if we put* $\phi_1 = \phi_2 = 1/\sqrt{2\pi}$ *and* $\phi_3 = \cos(\cdot)/\pi$*, then* $\mathcal{U} = \{\phi_k\}_{k=1}^3$ *and yet* $\langle \phi_1, \phi_2 \rangle \neq 0$.

Here is the great convenience of orthonormal sets: we know easily how to represent vectors as linear combinations of elements of those sets.

45.13 Theorem. *Let* $\mathcal{V}$ *be an inner product space and* $\{\phi_k\}_{k=1}^m$ *be orthonormal. Suppose that we can write* $f \in \mathcal{V}$ *as*

$$f = \sum_{k=1}^n \alpha_k \phi_k$$

for some $\alpha_k \in \mathbb{C}$*. Then*

$$\alpha_k = \langle f, \phi_k \rangle .$$

Proof. Fix j with $1 \leq j \leq n$ and compute

$$\langle f, \phi_j \rangle = \left\langle \sum_{k=1}^n \alpha_k \phi_k, \phi_j \right\rangle = \sum_{k=1}^n \alpha_k \langle \phi_k, \phi_j \rangle = \alpha_j.$$

The third equality is orthonormality with $\langle \phi_k, \phi_j \rangle = 1$ for $k = j$ and 0 otherwise, per the convention in Remark 45.12. ∎

In addition to facilitating convenient representations of vectors, inner products give a natural way of measuring *size* of vectors.

45.14 Theorem. *Let $\mathcal{V}$ be an inner product space. The* NORM *on $\mathcal{V}$ induced by the inner product is the map*

$$\|\cdot\| : \mathcal{V} \to \mathbb{R} \colon f \mapsto \sqrt{\langle f, f \rangle},$$

and it satisfies the following properties.

(i) [Nonnegativity] $\|f\| \geq 0$ *for all* $f \in \mathcal{V}$.

(ii) [Definiteness] $\|f\| = 0$ *if and only if* $f = 0$.

(iii) [Homogeneity] $\|\alpha f\| = |\alpha| \, \|f\|$ *for all* $\alpha \in \mathbb{C}$, $f \in \mathcal{V}$.

(iv) [Triangle inequality] $\|f + g\| \leq \|f\| + \|g\|$ *for all* $f, g \in \mathcal{V}$.

From now on, when we use the symbol $\|\cdot\|$, it will always be for the norm induced by the underlying inner product.

45.15 Problem. Let $\mathcal{V}$ be an inner product space. Suppose that the CAUCHY–SCHWARZ INEQUALTY holds for all $f, g \in \mathcal{V}$:

$$|\langle f, g \rangle| \leq \|f\| \, \|g\| \, .$$

(It does.) Use this to prove the triangle inequality. [Hint: *start by squaring both sides of the desired triangle inequality.*]

We now have (almost) all of the underlying structure that we need to state and govern our Fourier series problems except for a way of actually representing the differential equations that arise from the product ansatz. For that, we need linear operators; in general, these are maps between vector spaces that "respect" the linear structure of each space. Our operators will be slightly different and more specific. First, we will consider operators that map from a "subspace" of a given space back to itself.

45.16 Definition. *Let $\mathcal{V}$ be a vector space. A subset $\mathcal{W} \subseteq \mathcal{V}$ is a* SUBSPACE *of $\mathcal{V}$ if $0 \in \mathcal{W}$ and if $f + g$, $\alpha f \in \mathcal{W}$ for all $f, g \in \mathcal{W}$ and $\alpha \in \mathbb{C}$.*

45.17 Example. Let $I \subseteq \mathbb{R}$ be an interval and $0 \leq r_1 \leq r_2$ be integers. Then $\mathcal{C}^{r_2}(I)$ is a subspace of $\mathcal{C}^{r_1}(I)$.

45.18 Problem. Explain why L^1 and L^∞ are not subspaces of each other.

Now we are ready for our flavor of linear operators.

45.19 Definition. *Let $\mathcal{V}$ be a vector space. A* **LINEAR OPERATOR IN** $\mathcal{V}$ *is a map $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$, where $\mathcal{D}(\mathcal{A}) \subseteq \mathcal{V}$ is a subspace of $\mathcal{V}$, called the* **DOMAIN** *of $\mathcal{A}$, such that the following hold.*

(i) [Additivity] $\mathcal{A}(f+g) = \mathcal{A}f + \mathcal{A}g$ *for all f, $g \in \mathcal{D}(\mathcal{A})$.*

(ii) [Homogeneity] $\mathcal{A}(\alpha f) = \alpha \mathcal{A} f$ *for all $\alpha \in \mathbb{C}$, $f \in \mathcal{D}(\mathcal{A})$.*

45.20 Example. Let $\mathcal{V} = \mathcal{C}([0,1])$, $\mathcal{D}(\mathcal{A}) = \mathcal{C}^1([0,1])$, and $\mathcal{A}f := f'$. Then $\mathcal{A}$ is a linear operator in $\mathcal{V}$ with domain $\mathcal{D}(\mathcal{A})$.

The linear operators that matter to us will interact very well with the underlying inner products.

45.21 Definition. *Let $\mathcal{V}$ be an inner product space and let $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ be an operator in $\mathcal{A}$.*

(i) *$\mathcal{A}$ is* **SELF-ADJOINT** *if*

$$\langle \mathcal{A}f, g\rangle = \langle f, \mathcal{A}g\rangle$$

for all f, $g \in \mathcal{D}(\mathcal{A})$.

(ii) *$\mathcal{A}$ is* **POSITIVE SEMIDEFINITE** *if*

$$\langle \mathcal{A}f, f\rangle \geq 0$$

for all $f \in \mathcal{D}(\mathcal{A})$.

The following example is prototypical in proving symmetry and positivity of a differential operator: integrate by parts and use the boundary conditions.

45.22 Example. Let

$$\mathcal{V} := \mathcal{C}([0,\pi]) \quad \text{and} \quad \langle f, g\rangle := \int_0^\pi f(x)\overline{g(x)}\,dx.$$

Let

$$\mathcal{D}(\mathcal{A}) := \big\{ f \in \mathcal{C}^2([0,\pi]) \;\big|\; f(0) = f(\pi) = 0 \big\} \quad \text{and} \quad \mathcal{A}f := -f''.$$

(i) We show that $\mathcal{A}$ is self-adjoint. Integrating by parts once gives, for f, $g \in \mathcal{D}(\mathcal{A})$,

$$-\langle \mathcal{A}f, g\rangle = \int_0^\pi f''(x)\overline{g(x)}\,dx = f'(x)\overline{g(x)}\Big|_{x=0}^{x=\pi} - \int_0^\pi f'(x)\overline{g'(x)}\,dx = -\int_0^\pi f'(x)\overline{g'(x)}\,dx,$$

since $g(0) = g(\pi)$ for $g \in \mathcal{D}(\mathcal{A})$. A second integration by parts gives

$$\int_0^\pi f'(x)\overline{g'(x)}\,dx = f(x)\overline{g'(x)}\Big|_{x=0}^{x=\pi} - \int_0^\pi f(x)\overline{g''(x)}\,dx = -\int_0^\pi f(x)\overline{g''(x)}\,dx$$

since $f(0) = f(\pi)$ for $f \in \mathcal{D}(\mathcal{A})$. Chasing through the negatives, we have

$$\langle \mathcal{A}f, g \rangle = \int_0^\pi f'(x)\overline{g'(x)}\, dx = -\int_0^\pi f(x)\overline{g''(x)}\, dx = \int_0^\pi f(x)[\overline{-g''(x)}]\, dx = \langle f, \mathcal{A}g \rangle.$$

(ii) We show that $\mathcal{A}$ is positive semidefinite. For $f \in \mathcal{D}(\mathcal{A})$, we compute

$$-\langle \mathcal{A}f, f \rangle = \int_0^\pi f''(x)\overline{f(x)}\, dx = f'(x)\overline{f(x)}\big|_{x=0}^{x=\pi} - \int_0^\pi f'(x)\overline{f'(x)}\, dx = -\int_0^\pi |f'(x)|^2\, dx$$

That is,

$$\langle \mathcal{A}f, f \rangle = \|f'\|^2 \geq 0.$$

45.23 Problem. In the example above, what happens if we redefine $\mathcal{A}$ as $\mathcal{A}f = f''$?

45.24 Problem. Redevelop the results of this example for the following operators.

(i) $\mathcal{A}f = -f''$, $\mathcal{V} = \mathcal{C}([-\pi,\pi])$, $\mathcal{D}(\mathcal{A}) = \{f \in \mathcal{C}^2([-\pi,\pi]) \mid f(-\pi) = f(\pi),\ f'(-\pi) = f'(\pi)\}$.

(ii) $\mathcal{A}u = -\Delta u$, $\mathcal{V} = \{u \in \mathcal{C}(\overline{\mathcal{R}}) \mid u = 0 \text{ on } \partial\mathcal{R}\}$, $\mathcal{D}(\mathcal{A}) = \mathcal{C}^2(\mathcal{R}) \cap \mathcal{V}$, $\mathcal{R} = \{(x,y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1,\ 0 \leq y \leq 1\}$. [Hint: *Lemma 39.5.*]

We can now finally state the abstract version of the first problem that arose from a product ansatz: given a vector space $\mathcal{V}$ and a linear operator $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ in $\mathcal{V}$, find $f \in \mathcal{D}(\mathcal{A})$ and $\lambda \in \mathbb{C}$ such that

$$\mathcal{A}f = \lambda f.$$

More precisely, we always wanted $f \neq 0$ to avoid trivial product solutions.

45.25 Definition. *Let $\mathcal{V}$ be a vector space and let $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ be a linear operator in $\mathcal{V}$. Suppose that $f \in \mathcal{D}(\mathcal{A}) \setminus \{0\}$ and $\lambda \in \mathbb{C}$ satisfy*

$$\mathcal{A}f = \lambda f.$$

Then λ is an **EIGENVALUE** *of $\mathcal{A}$ with corresponding* **EIGENVECTOR** *f. When $\mathcal{V}$ is a space of functions, we often say* **EIGENFUNCTION** *instead of eigenvector.*

45.26 Example. Classical results from ODE tell us that if $\mathcal{V} = \mathcal{C}(I)$ for any subinterval $I \subseteq \mathbb{R}$, and $\mathcal{A}f = f''$ with $\mathcal{D}(\mathcal{A}) = \mathcal{C}^2(I)$, then every $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathcal{A}$ with corresponding eigenfunctions $f(x) = c_1 e^{\sqrt{\lambda}x} + c_2 e^{-\sqrt{\lambda}x}$ if $\lambda > 0$, $f(x) = c_1 + c_2 x$ if $\lambda = 0$, and $f(x) = c_1 \cos(\sqrt{|\lambda|}x) + c_2 \sin(\sqrt{|\lambda|}x)$ if $\lambda < 0$. Finding eigenfunctions for $\lambda \in \mathbb{C}$ is possible, but it would require the annoying task of computing square roots of complex, nonreal numbers.

45.27 Example. Let $\mathcal{V}$ be a vector space and let $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ be a linear operator in $\mathcal{A}$. Suppose that 0 is *not* an eigenvalue of $\mathcal{A}$, and let $g \in \mathcal{V}$. Then there exists at most one $f \in \mathcal{D}(\mathcal{A})$ such that $\mathcal{A}f = g$. For if there were two distinct f_1, $f_2 \in \mathcal{D}(\mathcal{A})$ with $\mathcal{A}f_k = g$, $k = 1, 2$, then $f_1 - f_2 \neq 0$ but $\mathcal{A}(f_1 - f_2) = 0$. Then $f_1 - f_2$ would be an eigenvector for 0.

45.28 Problem. Let $\mathcal{V} = \mathcal{D}(\mathcal{A}) = \mathcal{C}([0,1])$ and $(\mathcal{A}f)(x) = xf(x)$. Prove that no $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathcal{A}$. [Hint: *if* $(x-\lambda)f(x) = 0$ *for some* $\lambda \in [0,1]$, *continuity forces* $f(\lambda) = 0$.]

45.29 Problem. Let $\mathcal{V}$ be a vector space and let $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ be a linear operator in $\mathcal{A}$. Suppose that $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathcal{A}$. Prove that the **EIGENSPACE**

$$\mathcal{E}_\lambda := \{f \in \mathcal{D}(\mathcal{A}) \mid \mathcal{A}f = \lambda f\}$$

is a subspace of $\mathcal{V}$. Is every vector in $\mathcal{E}_\lambda$ an eigenvector corresponding to λ?

There is a deep connection between how an operator interacts with an inner product and how its eigenvalues and eigenvectors behave. The next theorem and example show that seemingly special properties of solutions to $f'' + \lambda f = 0$ with various boundary conditions arise naturally from the operator-theoretic properties of $\mathcal{A}f := -f''$ on those function spaces.

45.30 Theorem. *Let $\mathcal{V}$ be an inner product space and let $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ be an operator in $\mathcal{V}$.*

(i) *If $\mathcal{A}$ is self-adjoint, then any eigenvalue of $\mathcal{A}$ is real: if $\mathcal{A}f = \lambda f$ for some $f \neq 0$, then $\lambda \in \mathbb{R}$.*

(ii) *If $\mathcal{A}$ is self-adjoint, then eigenvectors corresponding to distinct eigenvalues are orthogonal. That is, if $\mathcal{A}f_1 = \lambda_1 f_1$ and $\mathcal{A}f_2 = \lambda_2 f_2$ with $\lambda_1 \neq \lambda_2$, then $\langle f_1, f_2 \rangle = 0$.*

(iii) *If $\mathcal{A}$ is positive semidefinite, then its eigenvalues are nonnegative: if $\mathcal{A}f = \lambda f$ for some $f \neq 0$, then $\lambda \geq 0$.*

Proof. **(i)** We would like to show $\overline{\lambda} = \lambda$. We know that conjugates show up when we work in the second slot of the inner product, so we take an inner product with the only relevant vector at hand:

$$\langle f, \lambda f \rangle = \overline{\lambda} \langle f, f \rangle = \overline{\lambda} \|f\|^2.$$

But also

$$\langle f, \lambda f \rangle = \langle f, \mathcal{A}f \rangle = \langle \mathcal{A}f, f \rangle = \langle \lambda f, f \rangle = \lambda \langle f, f \rangle = \lambda \|f\|^2.$$

Thus

$$\overline{\lambda} \|f\|^2 = \langle f, \lambda f \rangle = \lambda \|f\|^2,$$

so

$$(\lambda - \overline{\lambda}) \|f\|^2 = 0.$$

Since f is an eigenvector, $f \neq 0$, and therefore $\|f\| \neq 0$. Thus $\lambda - \overline{\lambda} = 0$.

(ii) We want to control $\langle f_1, f_2 \rangle$, and that quantity almost shows up when we consider multiples like $\langle \lambda_1 f_1, f_2 \rangle$ or $\langle f_1, \lambda_2 f_2 \rangle$, which are the natural multiples to consider, given the eigenrelations assumed here. We compute

$$\langle \lambda_1 f_1, f_2 \rangle = \langle \mathcal{A} f_1, f_2 \rangle = \langle f_1, \mathcal{A} f_2 \rangle = \langle f_1, \lambda_2 f_2 \rangle = \overline{\lambda_2} \langle f_1, f_2 \rangle = \lambda_2 \langle f_1, f_2 \rangle,$$

since $\mathcal{A}$ is self-adjoint and therefore has real eigenvalues. That is,

$$\lambda_1 \langle f_1, f_2 \rangle = \langle \lambda_1 f_1, f_2 \rangle = \lambda_2 \langle f_1, f_2 \rangle,$$

thus

$$(\lambda_1 - \lambda_2) \langle f_1, f_2 \rangle = 0,$$

and since $\lambda_1 \neq \lambda_2$, we must have $\langle f_1, f_2 \rangle = 0$.

(iii) Since $\mathcal{A}$ is positive, we work with the only natural quantity relating $\mathcal{A}$, λ, and an inequality:

$$0 \le \langle \mathcal{A} f, f \rangle = \langle \lambda f, f \rangle = \lambda \langle f, f \rangle = \lambda \|f\|^2.$$

Since f is an eigenvector, $f \neq 0$, and so $\|f\| \neq 0$. Dividing, we have $0 \le \lambda$. ■

45.31 Problem. Adapt the proof above to show that if $\mathcal{A}$ is **POSITIVE DEFINITE** in the sense that $\langle \mathcal{A} f, f \rangle > 0$ for all $f \neq 0$, then all of the eigenvalues of $\mathcal{A}$ are positive.

45.32 Example. **(i)** In Example 45.22, we showed that $\mathcal{A} f = -f''$ is self-adjoint and positive as an operator in $\mathcal{V} = \mathcal{C}([0, \pi])$ with domain $\mathcal{D}(\mathcal{A}) = \{f \in \mathcal{C}^2([0,\pi]) \mid f(0) = f(\pi) = 0\}$. We also know (from separation of variables) that its eigenvalues are k^2 for $k \ge 1$, and any eigenfunction corresponding to k^2 is a scalar multiple of $\sin(k\cdot)$, and the set $\{\sin(k\cdot)\}_{k=1}^{\infty}$ is orthogonal. All of this agrees with the previous theorem on the eigenbehavior of self-adjoint positive operators. Here $\mathcal{A}$ has rather fewer eigenvalues than the second derivative operator in Example 45.26. This is because $\mathcal{A}$ here is a *different* operator: its domain is not all of $\mathcal{C}^2([0, \pi])$.

(ii) We have also shown that $\mathcal{A}$ is self-adjoint and positive as an operator in $\mathcal{V} = \mathcal{C}([-\pi, \pi])$ with domain $\mathcal{D}(\mathcal{A}) = \{f \in \mathcal{C}^2([-\pi,\pi]) \mid f(-\pi) = f(\pi),\ f'(-\pi) = f'(\pi)\}$. Here, however, the eigenvalues are k^2 for $k \ge 0$, and every eigenfunction corresponding to k^2 can be written as $f(x) = c_1 \cos(kx) + c_2 \sin(kx)$. *Domains matter.*

This example suggests that we should pay attention to the "size" of eigenspaces. (Counting elements does us no good, for any nontrivial subspace $\mathcal{W}$, i.e., $\mathcal{W} \neq \{0\}$, of a vector space $\mathcal{V}$ has infinitely many elements—why?) In each case, k^2 was an eigenvalue of the operator $\mathcal{A}$, but in the first case the eigenspace for k^2 was

$$\{c \sin(k\cdot) \mid c \in \mathbb{R}\}$$

while in the second it was

$$\{c_1 \cos(k\cdot) + c_2 \sin(k\cdot) \mid c_1,\ c_2 \in \mathbb{R}\}.$$

Both eigenspaces contain infinitely many functions, but the second should appear "larger" than the first as it admits two free constants. (Strictly speaking, in the context of Example 45.32, the first is a subspace of $\mathcal{C}([0,\pi])$, while the second is a subspace of $\mathcal{C}([-\pi,\pi])$; that is, the domains are different.)

45.33 Problem. Prove that $\{c_1 \cos(k\cdot) + c_2 \sin(k\cdot) \mid c_1,\ c_2 \in \mathbb{R}\} \neq \{c \sin(k\cdot) \mid c \in \mathbb{R}\}$.

The following theorem spells out a situation in which an eigenspace has this "finite" behavior.

45.34 Theorem. *Let $\mathcal{V}$ be an inner product space and $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ be an operator in $\mathcal{V}$ with eigenvalue $\lambda \in \mathbb{C}$. There exists at most one integer $m \geq 1$ such that the eigenspace $\mathcal{E}_\lambda$ (Problem 45.29) has the form*

$$\mathcal{E}_\lambda = \left\{ \sum_{k=1}^{m} \alpha_k \phi_k \;\middle|\; \alpha_1, \ldots, \alpha_m \in \mathbb{C} \right\}$$

for some orthonormal subset $\{\phi_k\}_{k=1}^m \subseteq \mathcal{D}(\mathcal{A})$. This integer m is the **GEOMETRIC MULTIPLICITY** *of λ as an eigenvalue of $\mathcal{A}$. (The subset $\{\phi_k\}_{k=1}^m$, unlike m, is not unique.)*

We will not prove this theorem here. Of course, it really is a statement about dimension and orthonormal basis: $\mathcal{E}_\lambda$ above is m-dimensional, and $\{\phi_k\}_{k=1}^m$ is an orthonormal basis for $\mathcal{E}_\lambda$. One would first prove the uniqueness of m, i.e., that dimension is well-defined, and then apply the Gram–Schmidt process to construct an orthonormal basis from an "ordinary" basis. In the interest of "getting to the good stuff," we will not do that here.

45.35 Problem. What is the geometric multiplicity of each eigenvalue of the operators in Example 45.32?

Here, at last, is the "good stuff"—or, at least, what we would like to be true.

45.36 Untheorem. *Let $\mathcal{V}$ be an inner product space and $\mathcal{A}\colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ be a self-adjoint operator in $\mathcal{V}$ whose eigenvalues are the set $\{\lambda_k\}_{k=1}^\infty$. Assume the following.*

(i) *$\lambda_k < \lambda_{k+1}$ for all k.*

(ii) *$\lim_{k\to\infty} \lambda_k = \infty$.*

(iii) *Each eigenvalue λ_k has finite geometric multiplicity m_k and the eigenspace corre-*

sponding to λ_k can be written as

$$\mathcal{E}_k = \left\{ \sum_{j=1}^{m_k} \alpha_i \phi_j^k \;\middle|\; \alpha_1, \ldots, \alpha_{m_k} \in \mathbb{C} \right\}$$

with $\{\phi_j^k\}_{j=1}^{m_k} \subseteq \mathcal{D}(\mathcal{A})$ orthonormal.

Define

$$\mathcal{P}_k \colon \mathcal{V} \to \mathcal{E}_k \colon f \mapsto \sum_{j=1}^{k} \left\langle f, \phi_j^k \right\rangle \phi_j^k. \tag{45.3}$$

Then

$$\lim_{n\to\infty} \left\| f - \sum_{k=1}^{n} \mathcal{P}_k f \right\| = 0 \tag{45.4}$$

for any $f \in \mathcal{V}$.

45.37 Problem. Review the two motivating boundary value problems that arose from the Dirichlet and periodic boundary conditions for the heat equation. Explain how each of these problems meets the hypotheses of Untheorem 45.36. [Hint: *for the periodic boundary conditions, it will be helpful to write the eigenvalues as $\lambda_k = (k-1)^2$ for $k \geq 1$.*] For each, write out an explicit formula for $\sum_{k=1}^{3} \mathcal{P}_k f$.

Like other optimistic statements in this course, Untheorem 45.36 is not a real theorem because it is not really true. The hypotheses here alone do not, in fact, guarantee the desired conclusion. We need more—more on the space $\mathcal{V}$ and the domain $\mathcal{D}(\mathcal{A})$ and the operator $\mathcal{A}$. It turns out that knowing a different characterization of the eigenvalues as minimizing values of a certain map is enough.

45.38 Lemma (Rayleigh quotient). *Assume the hypotheses of Untheorem 45.36 and define*

$$\mathcal{R}_{\mathcal{A}} \colon \mathcal{D}(\mathcal{A}) \setminus \{0\} \to \mathbb{R} \colon f \mapsto \frac{\langle \mathcal{A} f, f \rangle}{\|f\|^2}.$$

Suppose also that the eigenvalues satisfy

$$\lambda_1 = \min\{\mathcal{R}_{\mathcal{A}}[f] \mid \|f\| = 1\} \tag{45.5}$$

and, for $k \geq 2$,

$$\lambda_k = \min\{\mathcal{R}_{\mathcal{A}}[f] \mid \|f\| = 1 \text{ and } \mathcal{P}_j f = 0,\ 1 \leq j \leq k-1\}, \tag{45.6}$$

where $\mathcal{P}_k$ was defined in (45.3). Then the limit (45.4) holds.

Proving this lemma requires quite a bit more analysis and extra hypotheses on $\mathcal{V}$, $\mathcal{D}(\mathcal{A})$, and $\mathcal{A}$. However, if one accepts that the sets in (45.5) and (45.6) *do* have minimum values, then it is possible to prove that those minima are eigenvalues and the vector at which the minimum is attained is a corresponding eigenvector.

45.39 Problem. To motivate the set (45.5), suppose that λ is an eigenvalue of $\mathcal{A} \colon \mathcal{D}(\mathcal{A}) \to \mathcal{V}$ with eigenvector f and that $\mathcal{V}$ is an inner product space. Show that

$$\lambda = \frac{\langle \mathcal{A}f, f \rangle}{\|f\|^2}.$$