MATH 4310: PARTIAL DIFFERENTIAL EQUATIONS

Daily Log for Lectures and Readings

Timothy E. Faver

April 22, 2025

CONTENTS

How to Use This Daily Log

This document is our primary reference for the course. It contains all of the material that we discuss in class along with some supplementary remarks that may not be mentioned in a class meeting. Each individual day has references, when applicable, to relevant material from the text *Basic Partial Differential Equations* by David Bleecker and George Csordas.

This log contains three classes of problems.

(!) Problems marked (!) are meant to be attempted *immediately*. They will directly address or reinforce something that we covered (or perhaps omitted) in class. It will be to your great benefit to pause and work (!)-problems as you encounter them. You should attempt, and be able to complete, all (!)-problems whether or not they are assigned for problem sets.

(⋆) Problems marked (⋆) are intentionally more challenging and deeper than (!)-problems. The (⋆)-problems will summarize and generalize ideas that we have discussed in class and give you broader, possibly more abstract perspectives. You should attempt the (⋆)-problems on a second rereading of the lecture notes, after you have completed the (!)-problems. Completing all of the (⋆)-problems constitutes the *minimal* preparation for exams.

(+) Problems marked (+) are candidates for the portfolio project. These are meant to be more challenging than the (!)- and (⋆)-problems and will take you deeper into calculations and proofs and make connections to concepts across and beyond the course. Some (+)-problems do presume knowledge of other classes (e.g., linear algebra, ODE, real or complex analysis, topology), but the majority do not. It is not necessary to do all (+)-problems in preparation for an exam; instead, you should look out for (+)-problems that you find interesting and exciting, as that will make the portfolio project more meaningful (and palatable) for you.

Day 1: Monday, January 6.

> **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 (which we use as both a singular and a plural noun, depending on context) because many interesting quantities in life depend on more than one variable and because the study of PDE employs and motivates many interesting concepts in mathematics. In this course, we will usually study functions of two variables, typically one for space and one for time (or sometimes both for two-dimensional space), and usually the unknown function in our equations will be u; we will write $u = u(x,t)$ to emphasize that u depends on x and t, and x will denote space and t time. Major challenge of PDE involve *data* and *geometry*: when the domain of our unknown function is two-dimensional (or higher-dimensional), we must keep track of much more data from the inputs, and there are many more options available for the domain's geometry as a subset of $\mathbb{R}^2$ (or $\mathbb{R}^n$). We will avoid these challenges by taking fairly banal domains when working with one spatial variable and one temporal variable; there, each variable will belong to some subinterval of $\mathbb{R}$, possibly infinite, possibly closed and bounded.

This course will provide many opportunities to revisit topics from single and multivariable calculus; we will become stronger students of familiar calculus because of these opportunities, and we will develop new appreciation for things that we previously learned, most especially the integral. We will also have many opportunities to ask, but not fully answer, questions that connect PDE to other courses—in particular real and complex analysis, linear algebra, and topology. Questions from PDE motivate many of the rigorous results from those courses that we will not prove, fully or even partially, here. But we will prove many results here; after all, a proof is just an argument that we are correct about something.

We will devote significant attention to the following four canonical linear PDE:

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

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. In fact, once we know how to solve the transport equation as written above, a versatile "rescaling" technique will allow us to solve

$$au_t + bu_x = 0$$

for any choice of a, $b \in \mathbb{R}$. And, thanks to a clever "factoring" technique, we will be able to import many ideas from the transport equation to the wave equation, and so transport and

wave morally belong to the same "class" of PDE. But heat and Laplace are totally different, both from each other and from the transport/wave class. In particular, the difference between wave and Laplace, which is just the choice of $\pm$, is remarkable—a banal change in the algebraic structure of the PDE produces a profound change in the behavior of solutions and in the mathematical techniques and tools necessary for their analysis.

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 $c_1 u + c_2 v$ 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:

$$u_t + u u_x = 0 \qquad \text{Burgers's equation}$$
$$u_t + u_{xxx} + u u_x = 0 \quad \text{Korteweg–de Vries (KdV) equation.}$$

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

Here are some things that we will *not* do. 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.

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 will 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*.

We know this from calculus: 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. We 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 calculus professors), we will write $\int_a^b f$ most of the time, and we will agree that the dummy variable of integration does not 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.

($\int$**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

$$\int_a^b 1 \; dt = b - a.$$
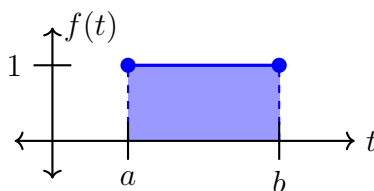
($\int$**2**) If we divide the region between the graph of f and the interval $[a, b]$ into multiple components, measure the net area of those components, and add those net areas together, we should get the total net area of the region between the graph of f and the interval

$[a, b]$. 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 $a \leq c \leq b$, then

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

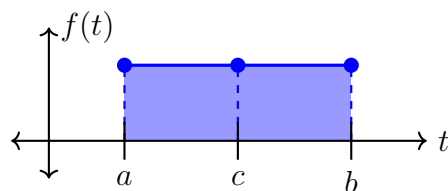($\int$**3**) If f is nonnegative, the net area of the region between the graph of f and the interval $[a, b]$ should be the genuine area of the region between the graph of f and the interval $[a, b]$, and this should be a positive quantity. So, we expect that if $0 \leq f(t)$ on $[a, b]$, then

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

($\int$**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:

$$\int_a^b f(t) \ dt + \int_a^b g(t) \ dt = \int_a^b \big[f(t) + g(t) \big] \ dt.$$
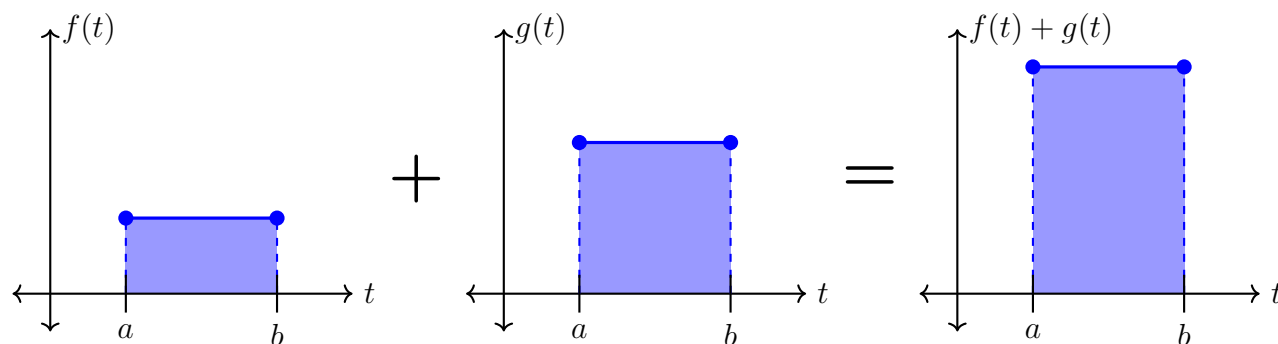
Next, multiplying a function $f \colon [a, b] \subseteq \mathbb{R} \to \mathbb{R}$ by a constant $\alpha \in \mathbb{R}$ should somehow "scale" the net area of the region between the graph of f and the interval $[a, b]$ by that factor α. For example, the area under the graph of $2f$ over $[a, b]$ should 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

$$\int_a^b \alpha f(t) \ dt = \alpha \int_a^b f(t) \ dt.$$
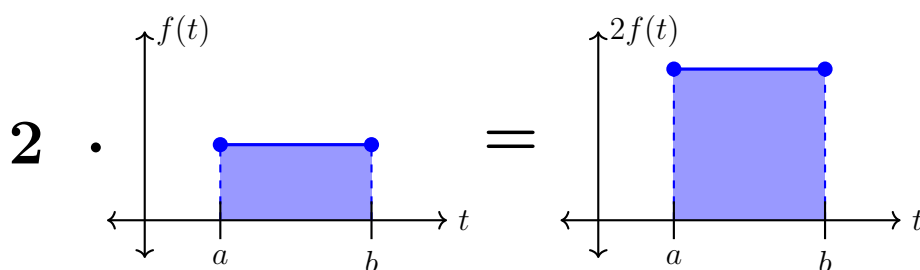
It turns out that these four properties are all that we need to prove the fundamental theorem of calculus, which we will.

Day 2: Wednesday, January 8.

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

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

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

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

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

Properties (∫4) encodes the linearity of the integral as an operator on the integrand with the limits of integration fixed, while property (∫2) is its **ADDITIVITY** over subintervals with the integrand fixed. Property (∫3) encodes the idea that a nonnegative function should have a nonnegative integral, while property (∫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 (∫1), (∫3), and (∫4) quite easily. Property (∫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 (⋆). Let $I \subseteq \mathbb{R}$ be an interval and f, $g: I \to \mathbb{R}$ 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 (∫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 (+). 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 ($\star$). 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

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

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

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

[Hint: *recall that if x, $r \in \mathbb{R}$ with $r \geq 0$, then $-|x| \leq x \leq |x|$ and $|x| \leq r$ if and only if $-r \leq x \leq r$. 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 $a \leq b$. Suppose that $f \colon I \to \mathbb{R}$ is continuous and there are m, $M \in \mathbb{R}$ such that $m \leq f(t) \leq M$ for all $t \in [a, b]$. Show that

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

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

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

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

$$F(t+h) - F(t) = \int_a^{t+h} f - \int_a^t f$$

$$= \int_a^{t+h} f + \int_t^a f$$

$$= \int_t^{t+h} f.$$

The first two terms of the numerator of the difference quotient have now collapsed into a single integral, so it would be nice if that third term, $-hf(t)$, were also an integral. First we cleverly rewrite h:

$$h = (t+h) - t = \int_t^{t+h} 1.$$

Then we use linearity of the integral to compute

$$hf(t) = f(t) \int_t^{t+h} 1 = \int_t^{t+h} f(t).$$

It may help at this point to introduce a variable of integration. Recall that t has been fixed throughout this proof, so we should not overwork it. Instead, we use τ, and so we have We then have

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

It therefore suffices to show that

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

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.*

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

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

for any $t \in I$.

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

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

We now need to show that

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

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

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

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 $0 \leq s \leq 1$, we have

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

thus (2.5) holds. ∎

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

2.9 Problem ($\star$). Prove the following "averaging" identity. Let $I \subseteq \mathbb{R}$ be an interval,

$x \in \mathcal{I}$, and $f \in \mathcal{C}(I)$. Then

$$f(x) = \lim_{r \to 0} \frac{1}{2r} \int_{x-r}^{x+r} f.$$

This says that the value of f at at point x is the limit of the average values of f on intervals centered at x as the width of those intervals shrinks to 0. [Hint: *with x fixed, put*

$$F_1(r) := \int_0^{x+r} f \quad and \quad F_2(r) := \int_0^{x-r} f$$

and show that

$$\frac{1}{2r} \int_{x-r}^{x+r} f = \frac{1}{2} \left(\frac{F_1(r) - F_1(0)}{r} - \frac{F_2(r) - F_2(0)}{r} \right).$$

Now think about difference quotients. How does this help?]

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, which we state but do not prove.

2.10 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*

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

2.11 Problem ($\star$). (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.*]

2.12 Problem ($\star$). Suppose that y solves the ODE $y' = ry$ for some $r \neq 0$ on some interval $I \subseteq \mathbb{R}$. That is, $y'(t) = ry(t)$ for all $t \in I$. Prove that $u(t) := y(t)e^{-rt}$ is constant. Explain why this justifies (what is hopefully!) our expectation that all solutions to this ODE are multiples of an exponential.

2.13 Corollary (FTC2). *Let $I \subseteq \mathbb{R}$ be an interval and let $f \colon I \to \mathbb{C}$ 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 $\varphi(b)$, and also φ should map the interval containing a and b to the domain of f, so $f \circ \varphi$ is defined and continuous. Also, the product $(f \circ \varphi)\varphi'$ should be continuous, and that requires φ' to be continuous on I.

2.14 Definition. *Let $I \subseteq \mathbb{R}$ be an interval. A function $\varphi \colon I \to \mathbb{R}$ 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 $\mathcal{C}^1(I)$.*

2.15 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. Let $F(\tau) := \int_{\varphi(a)}^\tau f$. The chain rule implies that $F \circ \varphi$ is an antiderivative of $(f \circ \varphi)\varphi'$; indeed, by FTC1,

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

Then FTC2 implies

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

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 (2.6) below.

2.16 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 $a \leq \varphi(t) \leq b$ for all t if $a \leq b$, and otherwise $b \leq \varphi(t) \leq a$ for all t if $b \leq a$. (Here is a proof of the first case, assuming $a \leq b$. Then $b - a \geq 0$ and $t \geq 0$, so $(b-a)t \geq 0$, thus $a \leq a + (b-a)t$. But also $(1-t)a \leq (1-t)b$ since $1 - t \geq 0$ and $a \leq b$, thus $(1 - t)a + tb \leq (1 - t)b + tb = b$.) 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

$$f(b) - f(a) = (b - a) \int_0^1 f'(a + (b - a)t) \, dt. \tag{2.6}$$

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 (2.6).

2.17 Problem (!). Prove the following variant of Example 2.16: if $I \subseteq \mathbb{R}$ is an interval, $f \in \mathcal{C}^1(I)$, and $t, t + h \in I$, then

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

2.18 Problem ($\star$). 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
$$F: \mathbb{R} \to \mathbb{R}: a \mapsto \int_a^{a+p} f$$
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.

2.19 Problem (!). Let $I \subseteq \mathbb{R}$ be an interval and f, $g \in \mathcal{C}^1(I)$ and a, $b \in I$. Prove the **INTEGRATION BY PARTS** identity
$$\int_a^b f g' = f(b)g(b) - f(a)g(a) - \int_a^b f'g. \tag{2.7}$$

[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.*]

2.20 Problem ($\star$). Let $f \in \mathcal{C}^2(\mathbb{R})$ with $f(0) = 0$. Prove that
$$f(x+y) - f(x) - f(y) = xy \int_0^1 \int_0^1 f''(sx + ty) \, ds \, dt$$
for all x, $y \in \mathbb{R}$. What happens in the case $f(\tau) = \tau^2$?

2.21 Problem (+). Suppose that $f \in \mathcal{C}^2(\mathbb{R})$. 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
$$|f(x) - f(y)| \leq M(|x| + |y|)|x - y|.$$
By considering the special case $f(x) = x^2$, explain why we might call this a "difference of squares" estimate. [Hint: *use Example 2.16 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 2.16 again to that integrand so that, in the end, $\mathcal{I}(x, y)$ is a double integral involving f''.*]

Day 3: Friday, January 10.

No class due to weather. You should read the material below on your own and work through it line by line.

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

$$\begin{cases} y' = f(t), \ t \in I \\ y(0) = y_0. \end{cases} \tag{3.1}$$

Here $I \subseteq \mathbb{R}$ is an interval with $0 \in I$, $f \in \mathcal{C}(I)$ is a given function, and $y_0 \in \mathbb{R}$ is also specified. 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 $t \in I$ fixed. 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

$$y(t) = y_0 + \int_0^t f(\tau)\, d\tau. \tag{3.2}$$

Thus if y solves the IVP (3.1), then y has the form above. This is a *uniqueness* result: the only possible solution is this one. But is this really a solution?

3.1 Problem (!). Use FTC1 and properties of integrals to check that defining y by (3.2) solves (3.1).

We write this up formally.

3.2 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

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

3.3 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.)

Day 4: Monday, January 13. ▬▬▬▬▬▬

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.

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 $C = y(0) = y_0$. Problem 2.12 proves this using a (nonobvious) algebraic trick, but we will also see this as a consequence of the more general result below that includes the driving term.

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. By the way, we call it "variation of parameters" because we have "varied" the parameter y_0 in the solution to the linear homogeneous IVP (i.e., the solution $y(t) = y_0 e^{at}$ when $f = 0$) via the ansatz $y(t) = u(t)e^{at}$, with u replacing y_0.

4.1 Theorem (Variation of parameters). *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*

$$\begin{cases} y' = ay + f(t) \\ y(0) = y_0 \end{cases} \tag{4.1}$$

is

$$y(t) = e^{at} \left(y_0 + \int_0^t e^{-a\tau} f(\tau) \, d\tau \right). \tag{4.2}$$

Is it?

4.2 Problem (!). (i) Check that the function y in (4.2) actually solves (4.1). (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 3.2 from Theorem 4.1 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 4.1 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.

4.3 Example. We study

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

and rather than just use the formula from (4.2), 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

$$y(t) = e^{2t}\left(\frac{3}{2} + \frac{e^{-6t}}{2}\right).$$

4.4 Problem ($\star$). We probably expect physically that two objects in motion that start "close" together should remain "close" together, at least for "some" time. We might call this "continuous dependence on initial conditions." Suppose that u and v solve

$$\begin{cases} u' = au + f(t) \\ u(0) = u_0 \end{cases} \quad \text{and} \quad \begin{cases} v' = av + f(t) \\ v(0) = v_0. \end{cases}$$

That is, u and v solve the same ODE but with possibly different initial conditions.

(i) Prove that

$$|u(t) - v(t)| \le e^{at}|u_0 - v_0|. \tag{4.3}$$

This estimate controls how close u and v are at time t in terms of how close u_0 and v_0 are.

(ii) Suppose $a < 0$. What does (4.3) say about the behavior of different solutions to $y' = ay + f(t)$ as $t \to \infty$?

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 a toy problem of how boundary behavior determines the solution.

4.5 Example. Let $f \in C(\mathbb{R})$. All solutions to $y' = f(t)$ are

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

What, if any, choices for the initial condition $y(0)$ guarantee

$$\lim_{t \to \infty} y(t) = 0?$$

We want

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

By the basic algebra of limits, this happens if and only if (1) the limit

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

exists and (2) the identity

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

holds.

We have discovered something new: the condition (1) must be met. That is to say, the improper integral $\int_0^\infty f(\tau) \, d\tau$ must converge. Nowhere in the statement of the "end behavior problem"

$$\begin{cases} y' = f(t) \\ \lim_{t \to \infty} y(t) = 0 \end{cases} \tag{4.4}$$

was it made explicit that f must be improperly integrable on $[0, \infty)$. But if this is true, we have shown that any solution to (4.4) must satisfy $y(0) = -\int_0^\infty f(\tau) \, d\tau$, and so any solution to (4.4) satisfies the IVP

$$\begin{cases} y' = f(t) \\ y(0) = -\int_0^\infty f(\tau) \, d\tau. \end{cases}$$

Is the reverse true?

4.6 Problem (!). Assume the following.

(i) $f \in C(\mathbb{R})$ is improperly integrable on $[0, \infty)$, i.e., $\lim_{b \to \infty} \int_0^b f$ exists.

(ii) Improper integrals respect the "algebraic" properties of definite integrals from Theorem

2.1, and that the derivative of $t \mapsto \int_t^\infty f$ is $-f(t)$.

Prove that the only solution to

$$\begin{cases} y' = f(t) \\ \lim_{t\to\infty} y(t) = 0 \end{cases}$$

is

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

4.7 Problem ($\star$). Here is a variation on variation of parameters that turns out to be useful. We can replace the equality in exponential growth with an inequality and still get the result that we expect. More precisely, suppose that $y \in \mathcal{C}^1(\mathbb{R})$ satisfies

$$\begin{cases} y'(t) \le ay(t), \ t \in \mathbb{R} \\ y(0) = 0 \quad 0 \le y(t), \ t \in \mathbb{R} \end{cases}$$

for some $a \in \mathbb{R}$. Then $y = 0$, which is what we would expect if we had $=$ instead of $\le$; this result with the inequality is called **GRONWALL'S INEQUALITY**. Here is how to prove this.

(i) Make the ansatz $y(t) = e^{at}u(t)$ and show that u solves

$$\begin{cases} u'(t) \le 0, \ t \in \mathbb{R} \\ u(0) = 0 \\ 0 \le u(t), \ t \in \mathbb{R}. \end{cases}$$

(ii) Deduce from this that $u(t) = 0$. [Hint: *consider rewriting $0 \le u(t) = u(0) + \int_0^t u'(\tau)\, d\tau$ and getting an upper bound on the right side.*]

Day 5: Wednesday, January 15.

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

Pages 2–3 review separation of variables for ODE.

5.1 Example. Now we consider the more general "end behavior" problem

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

where $a > 0$. Any solution to the ODE must meet

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

and so 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 $|y(t)| \leq 1$. From the formula for y, 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.

5.2 Problem ($\star$). 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 $t \geq M$, we have

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

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

(i) What if we consider $t \to -\infty$? Adapt the work in Example 5.1 to relate $\lim_{t\to-\infty} y(t)$, if this limit exists, and $y(0)$, in the case that $y' = ay + f(t)$ with $a < 0$.

(ii) 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 as $t \to \infty$ is of no help in specifying the initial condition.

5.3 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. And before doing that, we need to review a fact about continuity that will resurface many times in this course.

5.4 Lemma. *Let $I \subseteq \mathbb{R}$ be an interval and let $f \in \mathcal{C}(I)$. Suppose that $f(t_0) \ne 0$ for some $t_0 \in I$. Then there exists $\delta > 0$ such that $f(t) \ne 0$ for $t \in (t_0 - \delta, t_0 + \delta) \cap I$.*

Proof. We start with "proof by picture," which is always a good way to get an idea for the "real" proof. By continuity, the graph of f "near" t_0 should be "close" to $f(t_0)$, and so the graph should be above the t-axis.
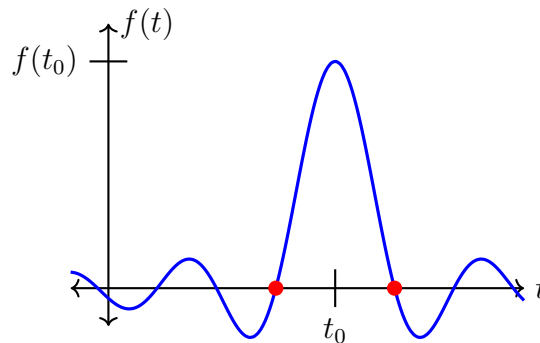
Here is the more rigorous proof. By continuity and the assumption $f(t_0) > 0$, there is $\delta > 0$ such that if $t \in (t_0 - \delta, t_0 + \delta) \cap I$, then $|f(t) - f(t_0)| < f(t_0)/2$. (In the language of classical δ-ϵ proofs, we are taking $\epsilon = f(t_0)/2$ here.) This inequality is equivalent to

$$\frac{f(t_0)}{2} < f(t) < \frac{3f(t_0)}{2},$$

and so $f(t) > f(t_0)/2 > 0$ for $t \in (t_0 - \delta, t_0 + \delta) \cap I$.

If $f(t_0) < 0$, put $g(t) := |f(t)|$ and use the previous argument to conclude that $g(t) > 0$ for $t \in t_0 - \delta, t_0 + \delta) \cap I$, thus $f(t) < 0$ for those t. ∎

5.5 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 $y_0 \ne 0$. That is, $y(0) \ne 0$. By continuity—since a solution y to this IVP is defined and differentiable at 0, thus continuous at 0—for $t \approx 0$ we have $y(t) \ne 0$. Thus we can divide to find
$$\frac{y'(t)}{(y(t))^2} = 1. \tag{5.1}$$

This is the first big idea of separation of variables: "separate the variables" so that all appearances of the unknown function y and its derivative are together on one side. The second big idea is to integrate.

Specifically, since (5.1) holds for all $t \approx 0$, we can integrate
$$\int_0^t \frac{y'(\tau)}{(y(\tau))^2}\, d\tau = \int_0^t 1\, d\tau. \tag{5.2}$$

Note our good grammar: we are integrating from 0 to t, so we have changed the independent variable from t in (5.1) to τ above. We substitute $u = y(\tau)$ on the left and use the initial condition $y(0) = y_0$ to find
$$\int_0^t \frac{y'(\tau)}{(y(\tau))^2}\, d\tau = \int_{y(0)}^{y(t)} \frac{du}{u^2} = \int_{y_0}^{y(t)} u^{-2}\, du = -u^{-1}\Big|_{u=y_0}^{u=y(t)} = \frac{1}{y_0} - \frac{1}{y(t)}.$$

Returning to (5.2), we find
$$\frac{1}{y_0} - \frac{1}{y(t)} = t,$$
and so we solve for $y(t)$ as
$$y(t) = \left(\frac{1}{y_0} - t\right)^{-1}.$$

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**.

> **5.6 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:

$$\int_0^t \frac{y'(\tau)}{g(y(\tau))}\, d\tau = \int_0^t f(\tau)\, d\tau. \tag{5.3}$$

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

$$\int_0^t \frac{y'(\tau)}{g(y(\tau))}\, d\tau = \int_{y(0)}^{y(t)} \frac{du}{g(u)} = \int_{y_0}^{y(t)} \frac{du}{g(u)}. \tag{5.4}$$

Combining (5.3) and (5.4), 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

$$H(y, t) := \int_{y_0}^{y} \frac{du}{g(u)} - \int_0^t f(\tau)\, d\tau. \tag{5.5}$$

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.

5.7 Problem (!). Consider the exponential growth problem

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

where we assume $y_0 > 0$ (but place no restrictions on a). In the context of this specific problem, what is the function H from (5.5)? Use this function H, and the assumption $y_0 > 0$, to show, as expected, that $y(t) = y_0 e^{at}$.

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

is $y(t) = y_0 e^{rt}$. (Here $r \in \mathbb{R}$ is a fixed parameter.)

It would be nice if, in general, 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 $H(y, t) = 0$?

5.8 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 $\delta, \epsilon > 0$ and a continuously differentiable function $Y : (-\delta, \delta) \to (y_0 - \epsilon, y_0 + \epsilon)$ 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 6: Friday, January 17. ━━━━━

> ### 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.

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. One can prove the following theorem by recasting the second-order linear problem as a first-order linear system and developing an analogue of variation of parameters for that system, which requires some matrix manipulations but not too much fuss otherwise. We will not purse the linear system/matrix approach here.

> **6.1 Theorem.** *Let a, b, $c \in \mathbb{R}$ with $a \neq 0$ and $f \in \mathcal{C}(I)$ with $0 \in I$. There exists a unique solution $y \in \mathcal{C}^2(I)$ to the IVP*
>
> $$\begin{cases} ay'' + by' + cy = f(t) \\ y(0) = y_0 \\ y'(0) = y_1. \end{cases}$$

This theorem does not tell us in the slightest a *formula* for y. That will come later in the case $f \neq 0$, as we simply do not need it right now. Instead, we will focus largely on the homogeneous problem with $f = 0$. The uniqueness result is a consequence of the following.

6.2 Lemma. *Let a, b, $c \in \mathbb{R}$ with $a \neq 0$. The only solution $y \in \mathcal{C}^2(\mathbb{R})$ to*

$$\begin{cases} ay'' + by' + cy = 0 \\ y(0) = 0 \\ y'(0) = 0 \end{cases} \tag{6.1}$$

is $y = 0$.

6.3 Problem (!). Use Lemma 6.2 to prove Theorem 6.1. [Hint: *suppose that the IVP in the theorem has two solutions, say, u and v. What IVP does $z := u - v$ satisfy?*]

6.4 Problem (+). This problem outlines a proof of Lemma 6.2. Suppose that y solves (??). Put

$$z(t) := \frac{(y(t))^2 + (y'(t))^2}{2}.$$

Show that $z' = (1 - c)yy' - b(y')^2$. Then use the inequality $AB \leq (A^2 + B^2)/2$, valid for all A, $B \in \mathbb{R}$, to find $a > 0$ such that $z' \leq az$. Apply Gronwall's inequality from Problem 4.7 to conclude $z = 0$.

Taking Theorem 6.1 for granted, we now focus on the homogeneous case of $f = 0$. Here 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 structure
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 $\lambda_0 = -b/2$. So, one would need to show that $y(t) = c_1 e^{-(b/2)t} + c_2 t e^{-(b/2)t}$ solves $ay'' + by' + (b^2/4a)y = 0$. This is mostly a lot of thankless algebra—so thankless that we do not even spell it out as a problem.

6.5 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 $y(t) = e^{0t}\big(c_1 \cos(t) + c_2 \sin(t)\big) = c_1 \cos(t) + c_2 \sin(t)$.

6.6 Example. Let $\lambda \in \mathbb{R}$. The IVP

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

governs the motion of an undamped, undriven simple harmonic oscillator (at least when $\lambda > 0$). We can extract two solution formulas:

$$y(t) = \begin{cases} y_0 + y_1 t, \ \lambda = 0 \\ y_0 \cos(\lambda t) + \dfrac{y_1}{\lambda} \sin(\lambda t), \\ \lambda \neq 0. \end{cases}$$

It would be nice if these were really "the same." We might wonder what happens in the limit as $\lambda \to 0$ with t fixed. Certainly $\lim_{\lambda \to 0} y_0 \cos(\lambda t) = y_0$, and L'Hospital's rule gives

$$\lim_{\lambda \to 0} \frac{\sin(\lambda t)}{\lambda} = t.$$

This should be comforting: the solution appears to be continuous in λ. Is there a more efficient way to write it? In particular, can we make a factor of t appear in the second term when $\lambda \neq 0$? Certainly:

$$\frac{\sin(\lambda t)}{\lambda} = t \left(\frac{\sin(\lambda t)}{\lambda t} \right)$$

when $t \neq 0$. Put

$$\operatorname{sinc}(x) := \begin{cases} \sin(x)/x, \ x \neq 0 \\ 1, \ x = 0. \end{cases}$$

L'Hospital's rule ensures continuity of sinc; it is, in fact, infinitely differentiable.
For $t \neq 0$, we then have

$$y(t) = y_0 \cos(\lambda t) + y_1 t \operatorname{sinc}(\lambda t).$$

This formula is also valid at $t = 0$, since it reduces to $y_0 = y(0)$ there.

6.7 Example. We can prove uniqueness of solutions to

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

directly, without relying on Lemma 6.2. What is really valuable here is not the uniqueness result but the trick that we use to get it: multiply through by a derivative. This will resurface from time to time in our study of actual PDE.

Specifically, if $y'' + \lambda^2 y = 0$, then $y''y' + \lambda^2 yy' = 0$. It may not be obvious at first glance, since doing calculus in reverse probably feels unusual, but

$$yy' = \partial_t \left[\frac{y^2}{2} \right].$$

And, similarly,

$$y''y' = \partial_t \left[\frac{(y')^2}{2} \right]$$

Thus

$$\frac{1}{2} \partial_t [y^2 + (y')^2] = 0,$$

and so $y^2 + (y')^2$ is constant. We only know the value of y and y' at one point: $t = 0$. And so

$$(y(t))^2 + (y'(t))^2 = (y(0))^2 + (y'(0))^2 = 0$$

for all t.

Now here is another trick: if $a, b \in \mathbb{R}$, and if $a^2 + b^2 = 0$, then $a = b = 0$. Otherwise if $a \neq 0$ or $b \neq 0$, then $a^2 > 0$ or $b^2 > 0$, and then we would have $0 < a^2 + b^2 = 0$, a contradiction. In particular, $y(t) = 0$ for any t.

6.8 Problem (+). Generalize the preceding work as follows. Let $\mathcal{V} \in \mathcal{C}^1(\mathbb{R})$ with $\mathcal{V}(r) > 0$ for all $r \neq 0$, $\mathcal{V}(0) = 0$, 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 $\mathcal{V}(y)$?*]

6.9 Problem (+). If we change the ODE from $y'' + \lambda^2 y$ to $y'' - \lambda^2 y$, the formula in Example 6.6 and the uniqueness proof in Example 6.7 will not work. We adapt them here.

(i) The **HYPERBOLIC SINE AND COSINE**, respectively, are

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

Show that the solution to the IVP

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

can be written in the form

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

(ii) Prove that the only solution to

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

is $y = 0$ using the following steps. Put $z(t) := y(t/\lambda)$ and show that $z'' - z = 0$ with $z(0) = z'(0) = 0$. Show next that $(z' + z)' = z' + z$, so $z'(t) + z(t) = Ce^t$ for some constant C. Take $t = 0$ to conclude $C = 0$, so $z' = -z$. How does this help?

Day 7: Wednesday, January 22. ▬▬▬

> **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.

> **7.1 Undefinition.** *A function u is a* **SOLUTION** *to a PDE if u solves that PDE at each point in its domain and if every (mixed) partial derivative of u up to the highest-order derivative in the PDE exists and is continuous.*

7.2 Example. From this convention, it is probably obvious that if u solves the transport equation

$$u_t + u_x = 0,$$

then all (both) of its first partial derivatives need to exist; we also require u_t and u_x to be continuous. It is perhaps less obvious from this convention that if u solves the heat equation

$$u_t = u_{xx},$$

then the derivatives u_{tt}, u_{tx}, and u_{xt} also need to exist and be continuous (in which case $u_{tx} = u_{xt}$), since the highest order of a derivative appearing in the heat equation is 2.

Here is some useful notation to control continuous differentiability. Below, the mention of $\mathbb{R}^2$ is not so special; rather, virtually all of our PDE will be posed in $\mathbb{R}^2$ for simplicity.

7.3 Definition. *Let $\mathcal{D} \subseteq \mathbb{R}^2$ and let $r \geq 0$ be an integer. Denote by $\mathcal{C}^r(\mathcal{D})$ the set of all r-TIMES CONTINUOUSLY DIFFERENTIABLE functions on $\mathcal{D}$ whose (mixed) partial derivatives of up to order r exist and are continuous on $\mathcal{D}$.*

So, if we have a PDE posed on $\mathcal{D} \subseteq \mathbb{R}^2$ and the highest order of the derivative in that PDE is r, we want the solution to be in $\mathcal{C}^r(\mathcal{D})$.

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.

7.4 Example. We study the PDE

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

This is a nonlinear PDE, but with a sufficiently nice domain it simplifies radically. Regardless of the domain, the remarks at the end of Example 6.7 tell us that u must meet $u_t = u_x = 0$.

(i) First suppose that we want to solve this PDE on $\mathbb{R}^2$. We expect that since u is constant in both x and t, it is simply constant. Here is why. Fix $x_0 \in \mathbb{R}$ and set $v(t) = u(x_0, t)$. By the way, in the future we might refer to this function v as $u(x_0, \cdot)$; this notation tells us to think of the x-variable as fixed at x_0, while the t-variable is the only independent variable now.

Then $v' = 0$ on $\mathbb{R}$, so v is constant. In particular, $v(t) = v(0)$ for all t, so $u(x_0, t) = u(x_0, 0)$ for all $x_0, t \in \mathbb{R}$. The same logic shows that given $t_0 \in \mathbb{R}$, we have $u(x, t_0) = u(0, t_0)$ for all $x \in \mathbb{R}$. Thus $u(x, t) = u(x, 0) = u(0, 0)$ for all $(x, t) \in \mathbb{R}^2$.

(ii) Now suppose that we want to solve this PDE on

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

This is the set drawn below.
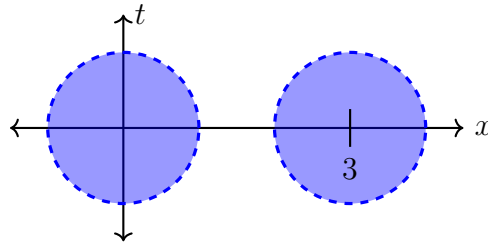
Perhaps inspired by Problem 2.11, we can put

$$u(x,t) := \begin{cases} 0, & x^2 + t^2 < 1 \\ 3, & (x-3)^2 + t^2 < 1 \end{cases}$$

to see that $u_x = u_t = 0$ but u is not constant.

(iii) What is missing in the previous situation is better geometric control. The set in the previous part is not "connected"—it is obviously in two "distinct" parts. We will get the result that we want ($u_t = u_x = 0$ implies that u is constant) if we assume more on the geometry. Say that $\mathcal{D} \subseteq \mathbb{R}^2$ is **CONNECTED** if we can find a **PATH** in $\mathcal{D}$ that connects any two points in $\mathcal{D}$. That is, given $(x_1, t_1), (x_2, t_2) \in \mathcal{D}$, there are functions $\gamma_1, \gamma_2 \in \mathcal{C}^1([0,1])$ such that

$$\begin{cases} \gamma_1(0) = x_1 \\ \gamma_2(0) = t_1, \end{cases} \qquad \begin{cases} \gamma_1(1) = x_2 \\ \gamma_2(1) = t_2, \end{cases} \qquad \text{and} \quad (\gamma_1(t), \gamma_2(t)) \in \mathcal{D} \text{ for all } t \in \mathcal{D}.$$

Say that $\mathcal{D}$ is connected and $u \in \mathcal{C}^1(\mathcal{D})$ with $u_x = u_t = 0$; we show that u is constant. Pick $(x_1, t_1), (x_2, t_2) \in \mathcal{D}$ and γ_1, γ_2 satisfying the above and set

$$v(s) := u(\gamma_1(s), \gamma_2(s))$$

for $s \in [0,1]$. The multivariable chain rule gives

$$v'(s) = u_x(\gamma_1(s), \gamma_2(s))\gamma_1'(s) + u_t(\gamma_1(s), \gamma_2(s))\gamma_2'(s) = 0$$

and so, since v is defined on an interval, v is constant. Thus

$$u(x_1, t_1) = v(0) = v(1) = u(x_2, t_2).$$

7.5 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.

7.6 Problem (!). Find all solutions to the following PDE. [Hint: *Example 6.6 and Problem 6.9.*] What regularity is necessary for f and g?

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

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

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

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

$$(7.1)$$

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

$$\int_a^b u(x,t)\ dx, \tag{7.2}$$

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

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

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

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

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

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

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

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

7.8 Problem (!). Use Problem 2.9 to explain why if the amount of substance in $[a,b]$ is given by (7.2), then the approximation (7.1) is valid.

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

$$\partial_t \left[\int_a^b u(x,t)\ dx \right] = \int_a^b u_t(x,t)\ dx. \tag{7.3}$$

This equality is called "differentiating under the integral," and it will be a hugely useful technique for us in the future. Broadly, it is a refinement of the notion that integrals are approximately sums, and derivatives commute with finite sums. We will revisit this at length in the future. For now, accept it as true and note that the variable of integration is not the same as the variable of differentiation, so we cannot invoke either of the fundamental theorems to simplify (7.3) further.

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. Previously we consider the rate of change of the amount of the substance within a spatial region; now we are considering the rate of change of the substance at a single point in space. This, too, is a little strange: 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 (7.3). That is,

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

and so

$$\int_a^b \left[u_t(x,t) + q_x(x,t) \right] \ dx = 0. \tag{7.4}$$

Now here is a marvelous fact about integrals.

7.9 Problem ($\star$). 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 $a \leq b$. 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 (7.4) 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.

Day 8: Friday, January 24.

> **Material from *Basic Partial Differential Equations* by Bleecker & Csordas**
>
> Pages 58–67 treat the somewhat broader problem $a_u x + b u_t + c u = f(x, t)$. 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.

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.

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

$$\begin{cases} u_t + u_x = 0, \ (x, t) \in \mathbb{R}^2 \\ u(x, 0) = f(x), \ x \in \mathbb{R}. \end{cases}$$

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

$$\begin{pmatrix} x \\ t \end{pmatrix} + s \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{8.1}$$

And u should be constant on any such line (as drawn in blue below).
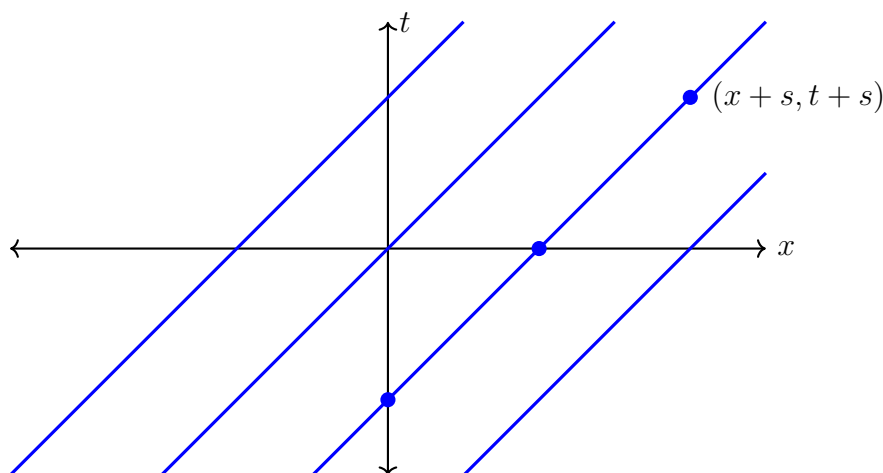
That is, for any $s, \widetilde{s} \in \mathbb{R}$, we have

$$u(x + s, t + s) = u(x + \widetilde{s}, t + \widetilde{s}). \tag{8.2}$$

Is there a point on the line through (x, t) parallel to $(1, 1)$ that is particularly "convenient"? Possibly the points with least data—with one coordinate equal to 0.

Can we pick s and/or $\widetilde{s}$ in a "convenient" way to exploit these coordinates? Why should we? We might add some data to the problem and impose the initial condition

$$u(x, 0) = f(x).$$

Then in (8.2) we could take $s \in \mathbb{R}$ arbitrary and $\widetilde{s} = -t$ to get

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

And we may as well put $s = 0$ to conclude

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

We can do this more efficiently and cut out some of the handwaving. 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.

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

$$\begin{cases} u_t + u_x = 0, \ (x, t) \in \mathbb{R}^2 \\ u(x, 0) = f(x), \ x \in \mathbb{R}. \end{cases} \tag{8.3}$$
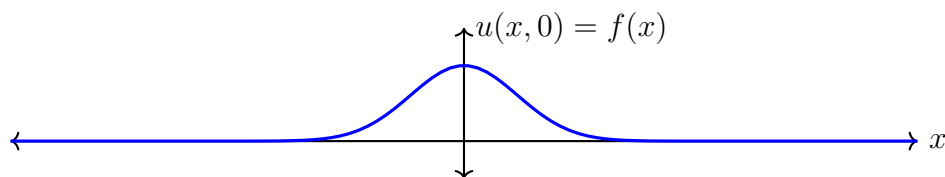
Then

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

This is a uniqueness result: the only possible solution to the IVP (8.3) is the one above. But is it really a solution?

8.2 Problem (!). Check that. Be sure to explain the importance of the regularity requirement $f \in \mathcal{C}^1(\mathbb{R})$.

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. This is best seen through some pictures.
Here is a graph for the initial data f.

Consider the solution u at time $t_1 > 0$. Then $u(x, t_1) = f(x - t_1)$, and this graph is just the graph of f "shifted" by t_1 units to the right.
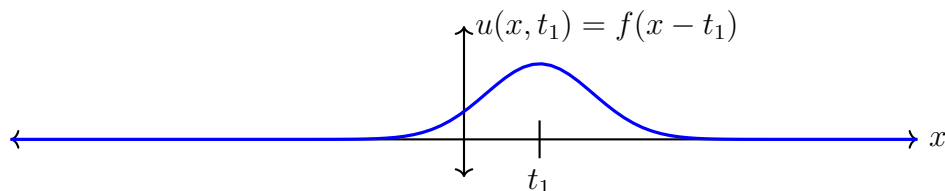
We let time evolve more and the graph gets shifted more.

$$u(x, t_2) = f(x - t_2)$$
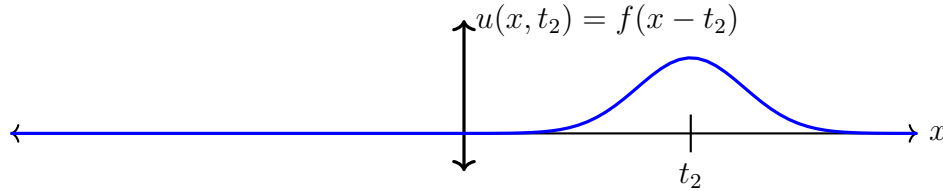
t_2

x

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.

More generally, we claim that the only solution to

$$\begin{cases} u_t + cu_x = 0, \ (x, t) \in \mathbb{R}^2 \\ u(x, 0) = f(x), \ x \in \mathbb{R} \end{cases}$$

is

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

That this u is a solution can be checked as in Problem 8.2. (Do that.) That this u is the *only* solution still needs proof, which we will provide later.

Also, we do not really need an initial condition: every solution to $u_t + cu_x = 0$ has the form

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

for some function p. Just take $p(X) = u(X, 0)$.

8.3 Example. The only solution to

$$\begin{cases} u_t + 3u_x = 0, \ (x, t) \in \mathbb{R}^2 \\ u(x, 0) = \sin(x), \ x \in \mathbb{R} \end{cases}$$

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 $\mathbb{R}^2$—that is, on a parametrized set $\{(x(s), t(s)) \mid s \in I\}$ for some interval $I \subseteq \mathbb{R}$.

8.4 Example. We consider the problem

$$\begin{cases} u_t + 3u_x = 0, \ (x, t) \in \mathbb{R}^2 \\ u(s, s) = \sin(s), \ s \in \mathbb{R}. \end{cases}$$

This prescribes the behavior of u on the line $x = t$. We know that if $u_t + 3u_x = 0$, then

$u(x, t) = p(x - 3t)$, where $p(X) = u(X, 0)$. Working backward, if we have a solution with the side condition, then

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

All we have to do is figure out p. Here is where some algebraic trickery helps: put $X = -2s$, so $s = -X/2$. Then

$$p(X) = p(-2s) = \sin(s) = \sin\left(-\frac{X}{2}\right) = -\sin\left(\frac{X}{2}\right).$$

And there is p.

So, we expect that the solution is

$$u(x, t) = p(x - 3t) = -\sin\left(\frac{x - 3t}{2}\right),$$

and we could always check that explicitly.

8.5 Problem (!). In what sense is any initial condition a side condition?

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

8.6 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 $u(x, t) = p(x - 3t)$ for some $p \in \mathcal{C}^1(\mathbb{R})$, and this p must satisfy

$$\sin(s) = u(3s, s) = p(3s - 3s) = p(0) \tag{8.4}$$

for all $s \in \mathbb{R}$. This is impossible, as $\sin(\cdot)$ is not constant. For example, (8.4) would require $0 = \sin(0) = p(0) = \sin(\pi/2) = 1$.

8.7 Problem ($\star$). 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*

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

*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*

$$h'(S) = \frac{1}{\sigma'(h(S))} \tag{8.5}$$

for all $S \in \mathbb{R}$. *(The identity (8.5) 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 show that

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

has a solution of the form

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

for some $h \in \mathcal{C}^1(\mathbb{R})$.

Day 9: Monday, January 27.

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

9.1 Example. We revisit the side conditions of Examples 8.4 and 8.6 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. Now we demand that u meet the side condition $u(s,s) = g(s)$ for some function g. 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.
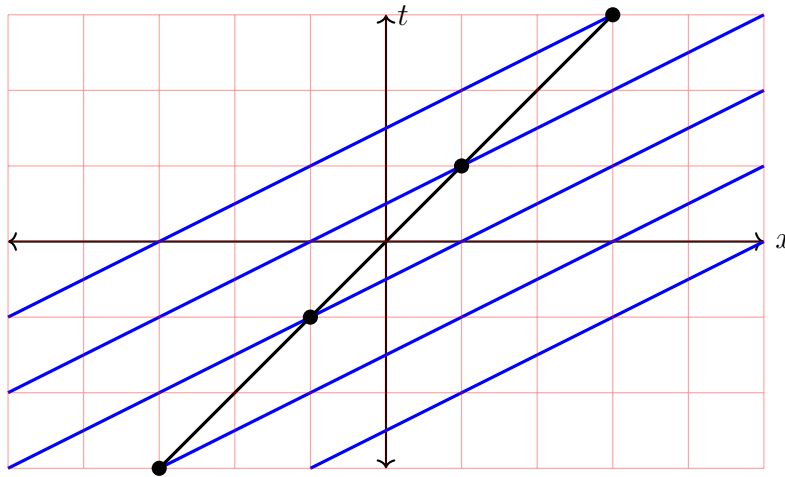
Since u is constant on each blue line, its value on any such line equals its value at the intersection of the blue line with the black line. There is only one such intersection, and so there is no ambiguity in the value of u.

(ii) Again we graph in blue lines with slope $1/3$. And, again, any solution u to $u_t + 3u_x = 0$ is constant on these lines. Now we demand that u meet the side condition $u(3s, s) = g(s)$ for some function g. We graph in black the curve parametrized by $(3s, s)$ with $s \in \mathbb{R}$, i.e., the line $t = x/3$. This black line completely overlaps with the line of slope $1/3$ that intersects the origin $(0, 0)$.
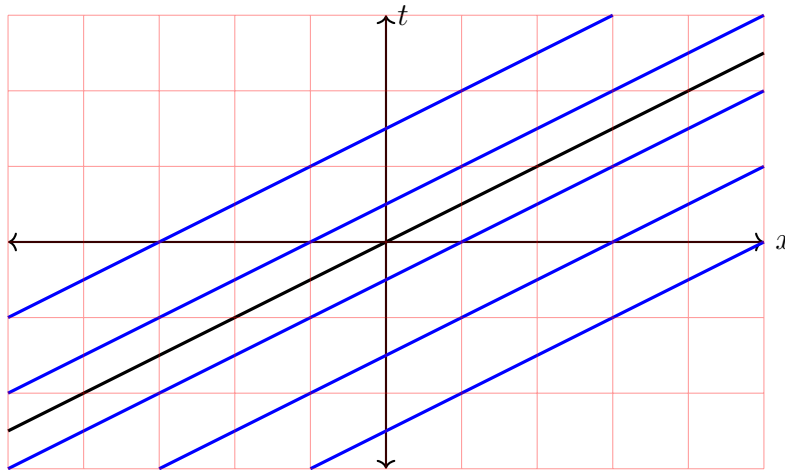
The problem is that u is supposed to be constant on all blue lines, and the black line now—but u is also supposed to agree with g on the black line. If g is not constant, a contradiction will result.

9.2 Problem ($\star$). (i) Revisit the side condition problem from Problem 8.7. Draw the side condition curve $(s, -s^3)$ and discuss how it intersects lines of slope $1/3$.

(ii) What goes wrong with the problem

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

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 8.6 and the geometry discussed in part (ii) of Example 9.1.

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 $u_t + c(x,t)u_x = 0$. Now we return to the dangling problem of solving the more general transport equation. Consider the IVP

$$\begin{cases} au_t + bu_x = 0, \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty. \end{cases} \tag{9.1}$$

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

$$\begin{cases} v_t + v_x = 0, \ -\infty < x, \ t < \infty \\ v(x,0) = g(x), \ -\infty < x < \infty. \end{cases} \iff v(x,t) = g(x - t). \tag{9.2}$$

How can we use (9.2) to solve (9.1)? (In (9.2), 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 (9.1) is equivalent to

$$\begin{cases} u_t + cu_x = 0, \ -\infty < x, \ t < \infty \\ u(x,0) = f(x), \ -\infty < x < \infty, \end{cases} \quad c = \frac{b}{a}. \tag{9.3}$$

Now we assume that u solves (9.3). 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 (9.2), 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

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

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

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

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

$$U(X,T) = u(cX, T) \quad \text{and} \quad u(x,t) = U\left(\frac{x}{c}, t\right). \tag{9.4}$$

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 (9.2) we have

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

By (9.4), 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 (9.1), 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.

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

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

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 $\alpha, \beta \in \mathbb{R}$. What values of α and β make U solve the "simpler" heat equation

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

9.4 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

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

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.

9.5 Problem (+). Let $a, b, \alpha, \beta, \gamma \in \mathbb{R}$ with both $a \neq 0$ and $b \neq 0$ and at least one of α or β nonzero. Let $f \in \mathcal{C}(\mathbb{R})$. Suppose that u solves $au_x + bu_t = 0$. What conditions on a, b, α, β, and γ ensure that $u(x,t) = f(x)$ whenever $\alpha x + \beta t = \gamma$? Interpret these conditions geometrically as well as algebraically.

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 the nonhomogeneous wave equation 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, $s \in \mathbb{R}$. (We are running out of variables, and σ looks and sounds like s.)

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

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

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.

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

$$\begin{cases} u_t + u_x = g(x,t), & -\infty < x, \ t < \infty \\ u(x,0) = f(x), & -\infty < x < \infty. \end{cases} \tag{9.5}$$

Then

$$u(x,t) = f(x - t) + \int_0^t g(x - t + \tau, \tau) \, d\tau. \tag{9.6}$$

9.7 Problem ($\star$). The integral term in (9.6) may look strange. Here is a way to see it as an analogue of something more familiar.

(i) Fix $a \in \mathbb{R}$ and $f \in \mathcal{C}(\mathbb{R})$. For t, $t_0 \in \mathbb{R}$, put $\mathcal{P}(t)y_0 := e^{at}y_0$. Show that

$$\begin{cases} y' = ay \\ y(0) = y_0 \end{cases} \iff y(t) = \mathcal{P}(t)y_0$$

and

$$\begin{cases} y' = ay + f(t) \\ y(0) = 0 \end{cases} \iff y(t) = \int_0^t \mathcal{P}(t-\tau)f(\tau)\, d\tau.$$

[Hint: *just import results from Theorem 4.1.*]

(ii) For t, $x \in \mathbb{R}$ and $f \in \mathcal{C}(\mathbb{R})$, let $\mathcal{P}(t,x)f := f(x-t)$. Show that

$$\begin{cases} u_t + u_x = 0, \ (x,t) \in \mathbb{R}^2 \\ u(x,0) = f(x), \ x \in \mathbb{R} \end{cases} \iff u(x,t) = \mathcal{P}(t,x)f$$

and

$$\begin{cases} u_t + u_x = g(x,t), \ (x,t) \in \mathbb{R}^2 \\ u(x,0) = 0, \ x \in \mathbb{R} \end{cases} \iff u(x,t) = \int_0^t \mathcal{P}(t-\tau,x)g(\cdot,\tau)\, d\tau.$$

Here $g \in \mathcal{C}^1(\mathbb{R}^2)$ and $g(\cdot,\tau)$ denotes the map $X \mapsto g(X,\tau)$. [Hint: *just import results from Theorem 9.6.*]

We can think of each $\mathcal{P}$ as a "propagator" that, in the case of the homogeneous problems, "propagates" the initial data forward in time and thereby gives a formula for homogeneous solutions. Here is how the propagator shows up in the nonhomogeneous problems. Recall that one version of the **CONVOLUTION** of the functions ϕ, $\psi \in \mathcal{C}(\mathbb{R})$ is the map $(\phi*\psi)(t) := \int_0^t \phi(t-\tau)\psi(\tau)\, d\tau$. The results above show that a particular solution to the nonhomogeneous problems (specifically, the solution with 0 initial condition) is given by convolving the propagtor (in time) with the driving term. This "convolve with propagator" approach will help us make useful (and correct) guesses about solving more complicated nonhomogeneous problems. We can also think of the propagators as linear operators: for each fixed t, $\mathcal{P}(t)$ is a linear operator on $\mathbb{R}$ (which is not very exciting: just scalar multiplication), whereas for each t and x, $\mathcal{P}(t,x)$ is a linear operator on $\mathcal{C}(\mathbb{R})$ that maps a function f to the scalar $f(x-t)$. (Thus $\mathcal{P}(t,x)$ is really the "evaluate at $x-t$" linear functional in that it maps the vector space $\mathcal{C}(\mathbb{R})$ to the underlying scalars $\mathbb{R}$.)

Day 10: Wednesday, January 29.

> **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.
>
> 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.

However, we did not show that any function u in the form (9.6) actually solves (9.5). 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 (7.3) 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

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

where I 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 $x \in I$. 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 (7.3)? 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

$$\lim_{h \to 0} \int_a^b \frac{f(x+h, s) - f(x, s)}{h}\ ds = \int_a^b \lim_{h \to 0} \frac{f(x+h, s) - f(x, s)}{h}\ ds. \tag{10.1}$$

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

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

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

is defined and differentiable on J and

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

10.2 Problem $(\star)$. 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 (10.1) to prove Leibniz's rule.

(ii) Go further and show (using, perhaps, Problem 2.17) that to prove (10.1), it suffices to establish that

$$\lim_{h \to 0} \int_a^b \int_0^1 \left[f_x(x + th, s) - f_x(x, s) \right]\ dt\ ds = 0. \tag{10.2}$$

(iii) Proving (10.2) 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 (10.2) holds. [Hint: *use Problem 2.17 again and watch out for the triple integral that shows up.*]

10.3 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 9.6. (It is also asking more of g than we did of the forcing term in the ODE from Theorem 4.1. 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

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

by FTC1 and

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

by Leibniz's rule.

We have proved the following.

10.4 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 $t \in \mathbb{R}$.

10.5 Problem ($\star$). 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

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

10.6 Problem ($\star$). Find all solutions to

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

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.

10.7 Problem (+). Let $c \in \mathbb{R}$ and $f \in \mathcal{C}^1(\mathbb{R})$. Suppose that u solves

$$\begin{cases} u_t + c u_x = 0, & x, \ t \in \mathbb{R} \\ u(x, 0) = f(x), & x \in \mathbb{R}. \end{cases}$$

Fix $x_0, x_1, t_0, t_1 \in \mathbb{R}$ and let

$$\mathcal{J}(t) := \int_{x_0 + c(t - t_0)}^{x_1 + c(t - t_1)} u(s, t) \, ds.$$

Give two proofs that $\mathcal{J}$ is constant as follows. (We might call the transport equation a "conservation law" because the quantity $\mathcal{J}$ is constant, or "conserved.")

(i) Use substitution to show

$$\mathcal{J}(t) = \int_{x_0 - c t_0}^{x - c t_1} f(\tau) \, d\tau.$$

(ii) Compute $\mathcal{J}'$ by differentiating under the integral and recalling that $u(b,t) - u(a,t) = \int_a^b u_x(s,t)\, ds$.

We now ommence 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

$$v_t - v_x = 0.$$

10.8 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 9.6 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

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

This would probably benefit from some cleaning up.

Day 11: Friday, January 31.

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

$$\int_0^t \left[g(x - t + 2\tau) + f'(x - t + 2\tau) \right] \, d\tau = \frac{1}{2} \int_{x-t}^{x+t} \left[g(s) + f'(s) \right] \, ds$$

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

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.

11.1 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*

$$\begin{cases} u_{tt} = c^2 u_{xx}, & -\infty < x, \ t < \infty \\ u(x, 0) = f(x), & -\infty < x < \infty \\ u_t(x, 0) = g(x), & -\infty < x < \infty \end{cases} \tag{11.1}$$

is the function

$$u(x, t) = \frac{f(x + ct) + f(x - ct)}{2} + \frac{1}{2c} \int_{x-ct}^{x+ct} g. \tag{11.2}$$

11.2 Problem ($\star$). Prove this.

(i) First, check that u as defined in (11.2) actually solves the wave IVP (11.1). 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 (11.1) and setting $U(X, T) = u(\alpha X, \beta T)$ for some $\alpha, \beta \in \mathbb{R}$. 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 (11.2) for u.

11.3 Example. We solve the wave IVP (11.1) for $c = 1$ and some choices of f and g and graph some results.
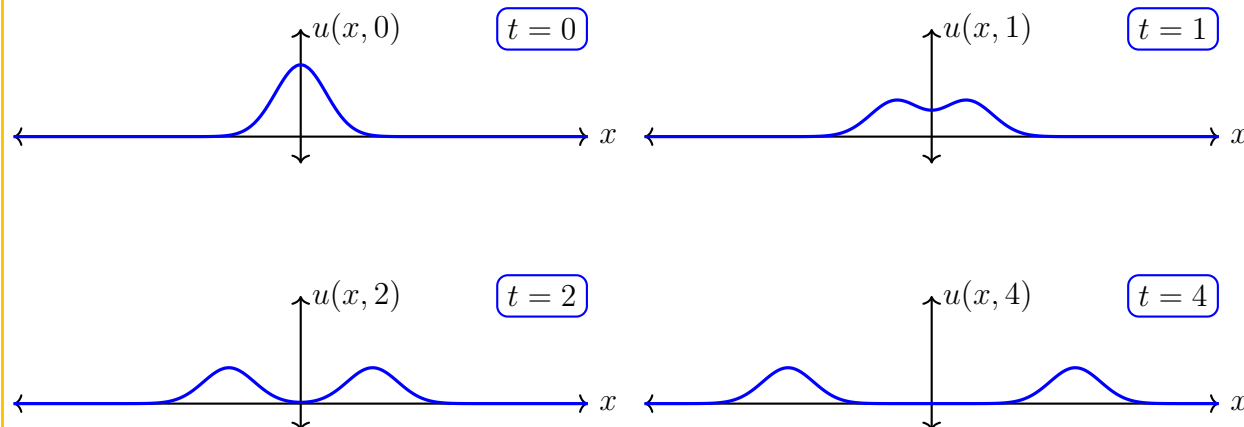
(i) Take
$$f(x) = 2e^{-x^2} \quad \text{and} \quad g(x) = 0.$$

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.

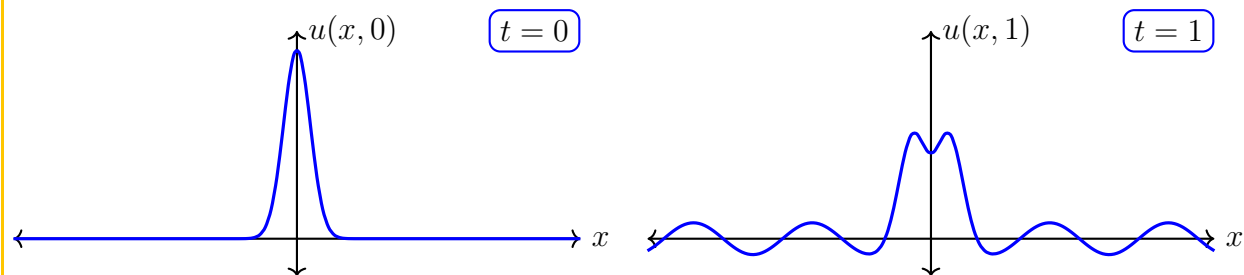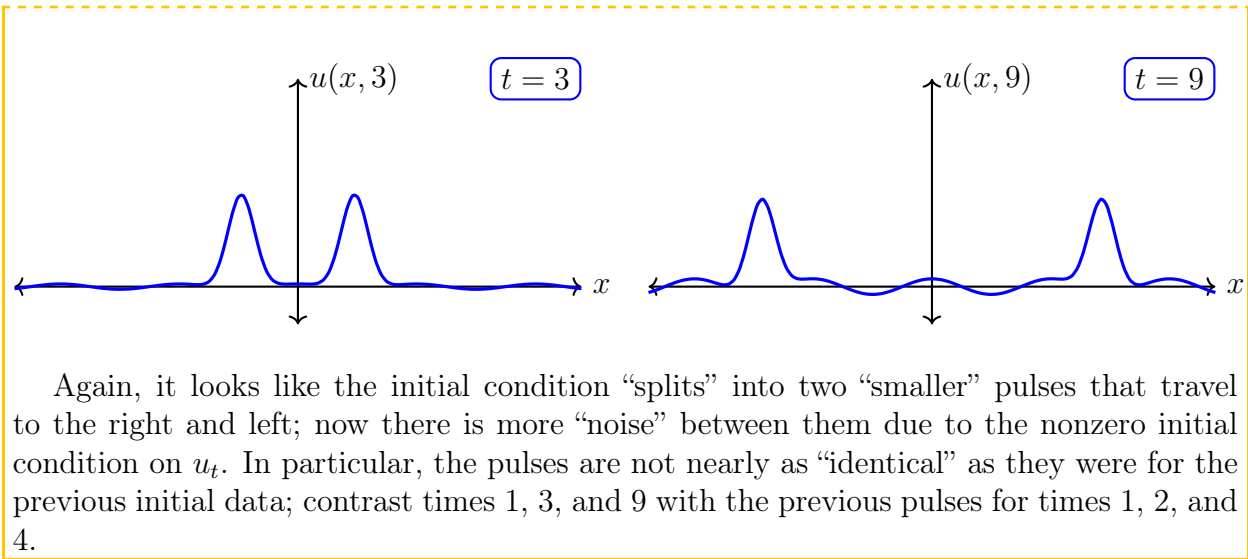**(ii)** Take
$$f(x) = 10e^{-x^2} \quad \text{and} \quad g(x) = \cos(x).$$

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

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

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

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

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

and abbreviate

$$L(X) := \frac{1}{2}\left(f(X) + \frac{1}{c}\int_0^X g\right) \quad \text{and} \quad R(X) := \frac{1}{2}\left(f(X) - \frac{1}{c}\int_0^X g\right).$$

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

$$u(x,t) = L(x+ct) + R(x-ct). \tag{11.3}$$

This is the superposition of the "profiles" L and R with L translated left with "speed" c and R translated "right." And this is why the graphs in Example 11.3 break up into two "counterpropagating" profiles.

11.4 Remark. *The profiles F and G above are definitely not the initial data f and g in general. In fact, the formula (11.3) makes sense without any initial data. Just assume that u solves $u_{tt} = c^2 u_{xx}$ 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 (11.3), and we can forget about f and g if we want.*

The structure in (11.3) is really a sum of traveling waves.

11.5 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.

11.6 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.

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.)

11.7 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

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

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 $p \neq 0$. Separating variables and integrating, we find

$$\int_0^X \frac{p'(s)}{(p(s))^2} \, ds = \int_0^X \frac{1}{c - 1} \, ds,$$

thus

$$\int_{p(0)}^{p(X)} \sigma^{-2} \, d\sigma = \frac{X}{c-1}$$

and so

$$(p(0))^{-1} - (p(X))^{-1} = \frac{X}{c-1}.$$

We conclude

$$p(X) = \left((p(0))^{-1} - \frac{X}{c-1} \right)^{-1}.$$

Replacing $p(0)$ with an arbitrary $K \in \mathbb{R}$, we conclude that traveling wave solutions are

$$u(x,t) = \left(K + \frac{x-ct}{1-c} \right)^{-1}.$$

11.8 Problem ($\star$). 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.*]

11.9 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 $c = 1$?

11.10 Problem ($\star$). 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.

Day 12: Monday, February 3.

> **Material from *Basic Partial Differential Equations* by Bleecker & Csordas**
>
> Pages 326–327 give physical motivation for the driven wave equation. Pages 327–328 give a different motivation for Duhamel's formula (i.e., the "propagator operators"). Page 329 states and proves the formula, and Example 6 on p. 330 goes through the calculations for concrete initial and driving data.

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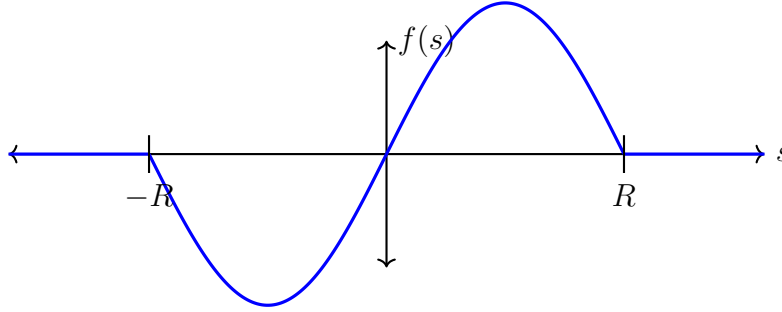
If the string is governed by $u_{tt} = c^2 u_{xx}$, then we expect that $c > 0$ is the speed of the wave(s) moving through the string. After t units of time, data or disturbances should only propagate ct units along the x-axis from where they were at time 0. This is born out by D'Alembert's formula.

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

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

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.

12.1 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 (11.1), then $u(x, t) = 0$ for $|x| > R + c|t|$.*

12.2 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.

12.3 Problem ($\star$). 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 12.1.

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

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

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:

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

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.

12.4 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

$$\begin{cases} y'' + by' + cy = f(t) \\ y(0) = y_0 \\ y'(0) = y_1 \end{cases} \tag{12.2}$$

is

$$y(t) = \mathcal{P}'(t)y_0 + \mathcal{P}(t)(y_1 + y_0 b) + \int_0^t \mathcal{P}(t - \tau)f(\tau) \ d\tau. \tag{12.3}$$

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

$$\begin{cases} z'' + bz' + cz = 0 \\ z(0) = y_0 \\ z'(0) = y_1 \end{cases} \quad and \quad \begin{cases} y_\star'' + by_\star' + cy_\star = f(t) \\ y_\star(0) = 0 \\ y_\star'(0) = 0. \end{cases} \tag{12.4}$$

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.

12.5 Problem (+). **(i)** Check that the formula (12.3) does yield a solution to (12.2). [Hint: *Lemma 10.4. For the initial condition, use* $\mathcal{P}'' + b\mathcal{P}' + c\mathcal{P} = 0$ *to compute* $\mathcal{P}''(0)$.]

(ii) Let $\lambda \in \mathbb{R} \setminus \{0\}$. What does Theorem 12.4 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 6.9?

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

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

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 (12.3) 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.

12.6 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

$$\mathcal{P}(t,x)g := \frac{1}{2} \int_{x-t}^{x+t} g(s) \ ds. \tag{12.5}$$

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 (12.3)!

Consequently, by analogy with (12.4) we are led to conjecture that

$$u(x,t) := \int_0^t \mathcal{P}(t - \tau, x)h(\cdot, \tau) \ d\tau \tag{12.6}$$

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 will check the PDE and leave the initial conditions as an exercise.

Day 13: Wednesday, February 5. ▬▬▬

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

Example 6 on pp. 307–308 and the remark on p. 308 discusses how to solve the semi-infinite string problem with the boundary condition $u_x(0,t) = 0$.

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 (12.6), we have

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

Here we use the formula (12.5) 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,

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

and

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

Thus

$$u_{xx}(x,t) = \frac{1}{2}\int_0^t \left[h_x(x+t-\tau,\tau) - h_x(x-t+\tau,\tau)\right]\,d\tau. \tag{13.1}$$

Now we work on the time derivative. We have

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

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,

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

Then

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

so

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

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

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

All together,

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

after comparison to (13.1).

13.1 Problem ($\star$). (i) Show that the function u defined in (12.6) satisfies

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

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

$$u(x, t) = \partial_t[\mathcal{P}(t, x)f] + \mathcal{P}(t, x)g + \int_0^t \mathcal{P}(t - \tau, x)h(\cdot, \tau) \, d\tau \tag{13.2}$$

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 (13.2) 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 (12.6) as a double integral.

13.2 Problem ($\star$). 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 (12.5). [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.*]

13.3 Problem (+). Rederive the solution formula

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

for the driven wave equation (with zero initial conditions) by "factoring" the wave equation as in our derivation of D'Alembert's formula and using results from the transport equation. That is, assume $u_{tt} - u_{xx} = h(x,t)$, put $v = u_t + u_x$, solve for v, and then solve for u.

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

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \leq x < \infty, \ t \in \mathbb{R} \\ u(x,0) = f(x), \ 0 \leq x < \infty \\ u_t(x,0) = g(x), \ 0 \leq x < \infty \\ u(0,t) = 0, \ t \in \mathbb{R} \end{cases} \tag{13.3}$$

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 just use D'Alembert's formula. The problem is that D'Alembert's formula does not make sense when f and g are only defined on $[0, \infty)$; for example, taking $x = 1$ in the formula, we would need to know the values of f at $1 \pm t$ for all $t \in \mathbb{R}$, thus f would have to be defined on all of $\mathbb{R}$. (Even restricting to nonnegative time does not help due to the $\pm t$ terms.)

The right idea is to work backward and *extend* spatial dependence in x from $[0, \infty)$ to all of $\mathbb{R}$. We will assume that we have a solution u to (13.3) and that we can construct functions $\widetilde{u}$, $\widetilde{f}$, and $\widetilde{g}$ such that the following hold.

1. $\widetilde{u}(x,t) = u(x,t)$ for all $x \geq 0$ and $t \in \mathbb{R}$. Likewise, $\widetilde{f}(x) = f(x)$ and $\widetilde{g}(x) = g(x)$ for all $x \geq 0$.

2. $\widetilde{f} \in \mathcal{C}^2(\mathbb{R})$ and $\widetilde{g} \in \mathcal{C}^1(\mathbb{R})$.

3. $\widetilde{u} \in \mathcal{C}^2(\mathbb{R}^2)$ solves

$$\begin{cases} \widetilde{u}_{tt} = \widetilde{u}_{xx}, \ -\infty < x, \ t < \infty \\ \widetilde{u}(x,0) = \widetilde{f}(x), \ -\infty < x < \infty \\ \widetilde{u}_t(x,0) = \widetilde{g}(x), \ -\infty < x < \infty. \end{cases}$$

Then D'Alembert's formula implies

$$\widetilde{u}(x,t) = \frac{\widetilde{f}(x+t) + \widetilde{f}(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} \widetilde{g}(s)\, ds. \tag{13.4}$$

Taking $x \geq 0$, we get a formula for u as well.

But how do we get these extensions, and why are they sufficiently differentiable for everything to work? There are many ways to extend a function from $[0, \infty)$ to all of $\mathbb{R}$. For example, putting

$$\widetilde{f}(x) = \begin{cases} f(x), & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{13.5}$$

is certainly an extension of f, and maybe one that makes sense (it keeps the initial data "turned off" for $x < 0$), but it may not be twice continuously differentiable, especially depending on the behavior of f at 0. It will pay off to learn more about what (13.3) tells us about f and g, especially in the context of the boundary condition.

To begin, if we have a solution u to (13.3), our usual conventions about the continuous differentiability of u imply $f = u(\cdot, 0) \in \mathcal{C}^2([0, \infty))$ and $g = u_t(\cdot, 0) \in \mathcal{C}^1([0, \infty))$. Recall that the convention here is that at the left endpoint 0 we only assume that limits from the right hold, e.g.,

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

Moreover, we may compute

$$f(0) = u(0,0) = 0,$$

where the first equality is the initial condition and the second is the boundary condition. Likewise,

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

Again, the first equality is the initial condition, and the second is the boundary condition. These calculations immediately tell us something new that the "ordinary" wave equation did not require: $f(0) = g(0) = 0$. Not every $f \in \mathcal{C}^2([0, \infty))$ or $g \in \mathcal{C}^1([0, \infty))$ will be compatible with the problem (13.3). What else might the structure of (13.3) tell us about f and g, in particular so that we get meaningful extensions of them to $\mathbb{R}$?

Day 14: Friday, February 7.

We return to the question of how to extend *correctly* the initial data f and g for (13.3) to functions defined on all of $\mathbb{R}$. We know that if this problem has a solution, the new feature of the boundary condition means that f and g must satisfy $f(0) = g(0) = 0$. Is there anything else that we can learn about f and g? We could try taking derivatives. We have $f'(x) = u_x(x, 0)$, and u_x does not appear in the wave equation, but

$$f''(x) = u_{xx}(x, 0) = u_{tt}(x, 0).$$

In particular, $f''(0) = u_{tt}(0, 0)$, and

$$u_{tt}(0,0) = \partial_t^2[u(0,t)]\big|_{t=0} = 0$$

by the boundary condition. Thus $f''(0) = 0$.

14.1 Problem (⋆). Can you obtain any information about g' or g'' from (13.3)? [Hint: *no. Why?*]

We will refer to these properties of f and g as **COMPATIBILITY CONDITIONS** for (13.3): if the problem has a solution, then the initial data needs to meet

$$f(0) = g(0) = f''(0) = 0. \tag{14.1}$$

It does not appear (Problem 14.1) that we can eke out anything else on f and g, and so trying to extend one of these functions naively via (13.5) will not really work.

14.2 Problem (⋆). Why not? If $f \in \mathcal{C}^2([0,\infty))$ and

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

is twice-continuously differentiable at 0, what does that imply about the value of $f'(0)$? Does that agree with what, if anything, (13.5) demands about $f'(0)$?

The right idea is something new: symmetry. Often in mathematics it is helpful to introduce and/or exploit some kind of symmetric or "reflective" structure. Even and odd functions are designed for just that.

14.3 Definition. (i) *A function $h \colon \mathbb{R} \to \mathbb{R}$ is* **EVEN** *if $h(-x) = h(x)$ for all $x \in \mathbb{R}$ and* **ODD** *if $h(-x) = -h(x)$ for all $x \in \mathbb{R}$.*

(ii) *The* **EVEN EXTENSION (REFLECTION)** *of a function $h \colon [0,\infty) \to \mathbb{R}$ is the function*
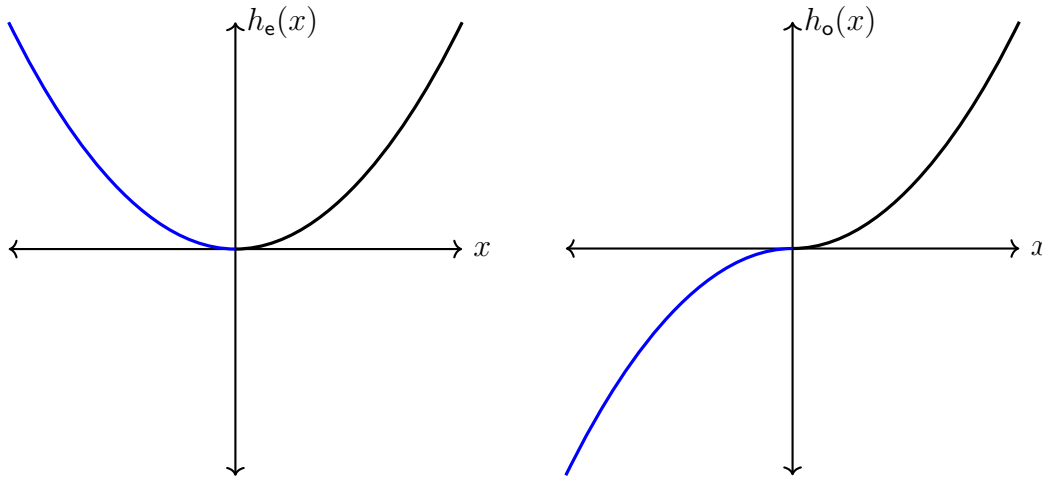
$$h_{\mathsf{e}} \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \begin{cases} h(x), & x \geq 0 \\ h(-x), & x < 0, \end{cases}$$

and the **ODD EXTENSION (REFLECTION)** *of $h \colon [0,\infty) \to \mathbb{R}$ is the function*

$$h_{\mathsf{o}} \colon \mathbb{R} \to \mathbb{R} \colon x \mapsto \begin{cases} h(x), & x \geq 0 \\ -h(-x), & x < 0. \end{cases}$$

Below the original function h is sketched in black and the even and odd extensions are

continued in blue.

14.4 Problem (!). Check that h_e is actually even and h_o is actually odd.

Which of these extensions should we use on the initial data in (13.3)? Some further properties of even and odd functions will help us make the decision.

14.5 Problem (!). (i) Suppose that $h \in \mathcal{C}^1(\mathbb{R})$ is even. Prove that h' is odd. [Hint: *differentiate both sides of $h(x) = -h(-x)$.*]

(ii) Suppose that $h \in \mathcal{C}^1(\mathbb{R})$ is odd. Prove that h' is even. [Hint: *differentiate both sides of $h(x) = h(-x)$.*]

(iii) Suppose that $h \colon \mathbb{R} \to \mathbb{R}$ is odd. Prove that $h(0) = 0$. [Hint: *if $h(x) = -h(-x)$, what happens when $x = 0$?*]

The fact that the compatibility conditions (14.1) require $f(0) = f''(0) = g(0) = 0$ suggest that we use odd extensions, as the second derivative of an odd function will be odd and thus vanish at 0. So here is our task: assume that u solves (13.3), put

$$\widetilde{u}(x,t) = \begin{cases} u(x,t), \ x \geq 0, \ t \in \mathbb{R} \\ -u(-x,t), \ x < 0, \ t \in \mathbb{R}, \end{cases} \qquad f_o(x) = \begin{cases} f(x), \ x \geq 0 \\ -f(-x), \ x < 0, \end{cases}$$

$$\text{and} \quad g_o(x) = \begin{cases} g(x), \ x \geq 0 \\ -g(-x), \ x < 0, \end{cases}$$

show that $f_o \in \mathcal{C}^2([0,\infty))$ and $g_o \in \mathcal{C}^1([0,\infty)$, show that $\widetilde{u} \in \mathcal{C}^2(\mathbb{R}^2)$, and show that

$$\begin{cases} \widetilde{u}_{tt} = \widetilde{u}_{xx}, \ -\infty < x, \ t < \infty \\ \widetilde{u}(x,0) = f_o(x), \ -\infty < x < \infty \\ \widetilde{u}_t(x,0) = g_o(x), \ -\infty < x < \infty. \end{cases}$$

Not too much to ask, right? We can then use D'Alembert's formula as in (13.4) to get a formula for $\widetilde{u}$ in terms of f_o and g_o, thus a formula for u in terms of f and g if we really need that. (By the way, we are sticking with $\widetilde{u}$, not u_o, to avoid too many subscripts.)

Basically all of the challenges here boil down to studying *what happens when $x = 0$.* This is the challenge of every piecewise function that we ever met in calculus: not so much what is happening on the individual "pieces" (usually the functions are pretty nice there) but rather what is happening where the pieces "meet" (usually this involves consideration of both left and right limits). Here is one kind of result that we need; the value of the following proof for our course is that it provides a healthy review of difference quotients and left and right limits.

14.6 Lemma. *Let $g \in \mathcal{C}^1([0, \infty))$ with $g(0) = 0$. Then $g_\text{o} \in \mathcal{C}^1(\mathbb{R})$.*

Proof. The chain rule tells us that $g_\text{o} \in \mathcal{C}^1(\mathbb{R} \setminus \{0\})$ with

$$g_\text{o}'(x) = \begin{cases} g'(x), & x > 0 \\ g'(-x), & x < 0. \end{cases}$$

We need to consider carefully what happens at $x = 0$.

Since $g \in \mathcal{C}^1([0, \infty))$ with $g(0) = 0$, we know

$$0 = g(0) = \lim_{x \to 0^+} g(x) \quad \text{and} \quad g'(0) = \lim_{h \to 0^+} \frac{g(0 + h) - g(0)}{h} = \lim_{h \to 0^+} \frac{g(h)}{h}.$$

We first want to show that

$$\lim_{h \to 0^\pm} \frac{g_\text{o}(0 + h) - g_\text{o}(0)}{h}$$

exist and are equal.

For $h > 0$, we have

$$\frac{g_\text{o}(0 + h) - g_\text{o}(0)}{h} = \frac{g_\text{o}(h)}{h} = \frac{g(h)}{h} = \frac{g(0 + h) - g(0)}{h}$$

and so

$$\lim_{h \to 0^+} \frac{g_\text{o}(0 + h) - g_\text{o}(0)}{h} = \lim_{h \to 0^+} \frac{g(0 + h) - g(0)}{h} = g'(0).$$

Next, for $h < 0$,

$$\frac{g_\text{o}(0 + h) - g_\text{o}(0)}{h} = \frac{-g(-h)}{h} = \frac{g(0 - h) - g(0)}{-h}.$$

Here is the only time that we really needed $g(0) = 0$. This permitted the second equality and the shuffling of the factor of -1 into the denominator. Thus

$$\lim_{h \to 0^-} \frac{g_\text{o}(0 + h) - g_\text{o}(0)}{h} = \lim_{h \to 0^-} \frac{g(0 - h) - g(0)}{-h} = \lim_{h \to 0^+} \frac{g(0 + h) - g(0)}{h} = g'(0).$$

Here we have used the helpful fact that

$$\lim_{h \to 0^-} \phi(-h) = \lim_{h \to 0^+} \phi(h).$$

Thus the desired limits from the left and right exist and are equal, and so g_o is differentiable on all of $\mathbb{R}$ with

$$g_o'(x) = \begin{cases} g'(x), & x \geq 0 \\ g'(-x), & x < 0. \end{cases}$$

Now we check continuity at $x = 0$:

$$\lim_{x \to 0^+} g_o'(x) = \lim_{x \to 0^+} g'(x) = g'(0),$$

while

$$\lim_{x \to 0^-} g_o'(x) = \lim_{x \to 0^-} g'(-x) = \lim_{x \to 0^+} g'(x) = g'(0).$$

This finishes the proof of continuity of g_o'. ∎

Everything else is essentially "more of the same": careful examination of left and right limits at $x = 0$.

14.7 Problem (+). Do all that.

(i) Show that if $f \in C^2([0, \infty))$ with $f(0) = f''(0) = 0$, then $f_o \in C^2(\mathbb{R})$. [Hint: *try to reduce the argument to repeated invocations of Lemma 14.6, or develop something new for even functions.*]

(ii) Show that if

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \leq x < \infty, \ t \in \mathbb{R} \\ u(x, 0) = f(x), \ 0 \leq x < \infty \\ u_t(x, 0) = g(x), \ 0 \leq x < \infty \\ u(0, t) = 0, \ t \in \mathbb{R} \end{cases} \quad \text{and} \quad \widetilde{u}(x, t) = \begin{cases} u(x, t), & x \geq 0, \ t \in \mathbb{R} \\ -u(-x, t), & x < 0, \ t \in \mathbb{R}, \end{cases}$$

then $\widetilde{u} \in C^2(\mathbb{R}^2)$ and

$$\begin{cases} \widetilde{u}_{tt} = \widetilde{u}_{xx}, \ -\infty < x, \ t < \infty \\ \widetilde{u}(x, 0) = \widetilde{f}(x), \ -\infty < x < \infty \\ \widetilde{u}_t(x, 0) = \widetilde{g}(x), \ -\infty < x < \infty. \end{cases}$$

[Hint: *again, this is mostly careful consideration of left and right limits at $x = 0$.*]

We should be careful with the flow of logic here. We have proved that if the wave equation for the semi-infinite string has a solution u, then a suitable extension of u solves the wave equation for the infinite string; in the process this gives us a formula for u. But why is there a solution for the semi-infinite problem in the first place? Is that formula from the extension process really a solution?

We are thinking that

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

with $x \geq 0$ and $t \in \mathbb{R}$. Since $f_o \in \mathcal{C}^2(\mathbb{R})$ and $g_o \in \mathcal{C}^1(\mathbb{R})$, we not need to do any new work to check that $u_{xx} = u_{tt}$, $u(x,0) = f(x)$, and $g_t(x,0) = g(x)$ for $0 \leq x < \infty$ and $t \in \mathbb{R}$. Rather, why does this formula meet the boundary condition $u(0,t) = 0$ for all t?

We compute

$$f_o(0+t) + f_o(0-t) = f_o(t) + f_o(-t) = f_o(t) - f_o(t) = 0$$

by the oddness of f_o. Next, we want $\int_{-t}^{t} g_o(s)\ ds = 0$. This is also true by oddness.

14.8 Problem ($\star$). 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).
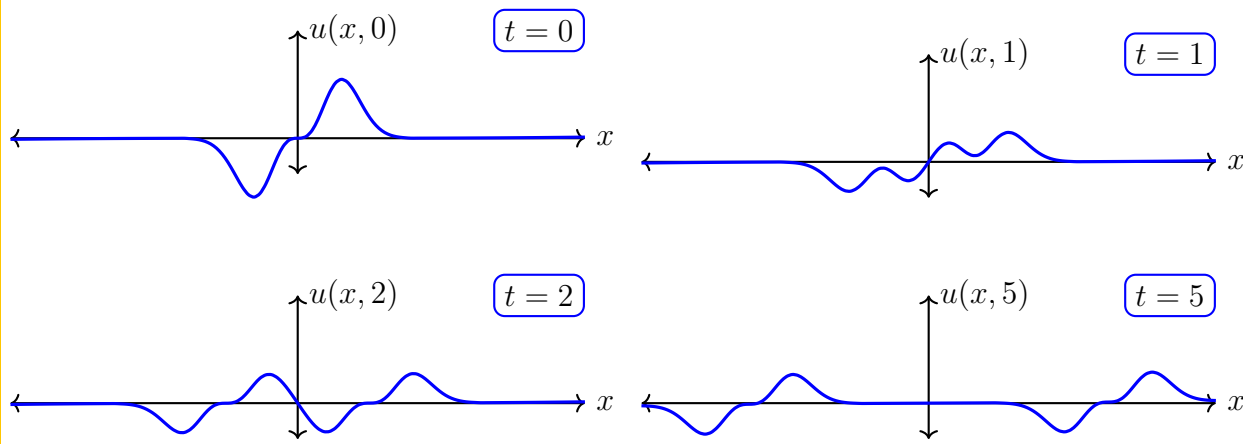
14.9 Example. We solve

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \leq x < \infty, \ t \in \mathbb{R} \\ u(x,0) = 4x^3 e^{-x^2} \\ u_t(x,0) = 0 \\ u(0,t) = 0. \end{cases}$$

Here the initial data is already odd, so we do not need to go to any great lengths to calculate odd extensions. Indeed, D'Alembert's formula for the infinite string with initial data given by the odd extensions says

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

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.

So there we are: the "odd extension method" showed that a solution to the semi-infinite wave equation can be extended to a solution to an infinite wave equation, at which point D'Alembert's formula could be invoked, and conversely good old calculus proves that that formula does indeed solve all of the semi-infinite wave equation—most notably the new feature of the boundary conditions.

14.10 Problem (+). 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?*]

Day 15: Monday, February 10.

> **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 $f_{\text{o}}^{\text{per}}$ and $g_{\text{o}}^{\text{per}}$ 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.
>
>   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.

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)

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ u(x,0) = f(x), \ 0 \le x \le L \\ u_t(x,0) = g(x), \ 0 \le x \le L \\ u(0,t) = u(L,t) = 0, \ -\infty < t < \infty. \end{cases} \tag{15.1}$$

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 $\{(x,t) \in \mathbb{R}^2 \mid 0 \leq x \leq L, \ t \in \mathbb{R}\}$. 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 $\mathbb{R}$ and use D'Alembert's formula. By "carefully," we mean that the extensions should be sufficiently differentiable, and our experience with the semi-infinite problem suggests that this regularity will rely also on knowing some exact values of f and g at 0, and maybe L.

15.1 Problem (!). Prove that if (15.1) has a solution, then the "compatibility conditions"

$$f(0) = g(0) = f(L) = g(L) = f''(0) = f''(L) = 0 \tag{15.2}$$

are all true.

That all of these values vanish at $x = 0$ suggests again odd extensions. The problem is that the solution and the initial data are only defined on $[0, L]$, so at best we could do odd extensions to $[-L, L]$.
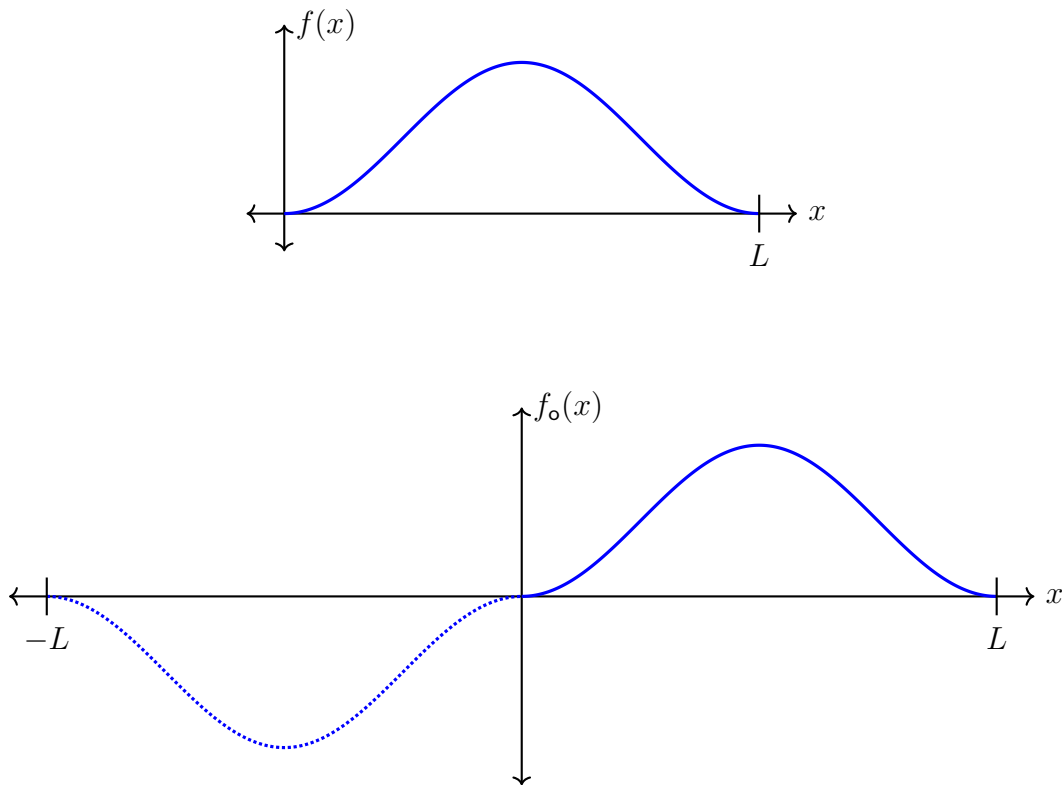
Here is the new idea. We are going to extend the odd extensions from $[-L, L]$ to $(-\infty, \infty)$ *periodically*. Informally, we "copy and paste" the graphs from $[-L, L]$ to the intervals $[(2k+1)L, (2k+3)L]$ for $k \in \mathbb{Z}$. For the initial data, call these periodic extensions $f_{\mathrm{o}}^{\mathrm{per}}$ and $g_{\mathrm{o}}^{\mathrm{per}}$; we require them to satisfy

$$f_{\mathrm{o}}^{\mathrm{per}}(x + 2L) = f_{\mathrm{o}}^{\mathrm{per}}(x) \quad \text{and} \quad g_{\mathrm{o}}^{\mathrm{per}}(x + 2L) = g_{\mathrm{o}}^{\mathrm{per}}(x)$$

for all $x \in \mathbb{R}$.
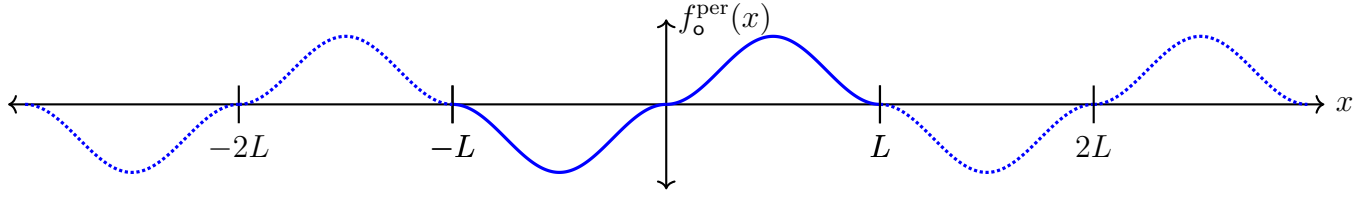
Likewise, assuming that a solution u to (15.1) exists, let $\widetilde{u}(\cdot, t)$ be its odd extension to $[L, L]$ and then let $\widetilde{u}^{\text{per}}(\cdot, t)$ be the periodic extension to $\mathbb{R}$. If $\widetilde{u}^{\text{per}} \in \mathcal{C}^2(\mathbb{R}^2)$, $f_{\text{o}}^{\text{per}} \in \mathcal{C}^2(\mathbb{R})$, and $g_{\text{o}}^{\text{per}} \in \mathcal{C}^1(\mathbb{R})$, and if

$$\begin{cases} \widetilde{u}_{tt}^{\text{per}} = \widetilde{u}_{xx}^{\text{per}}, & -\infty < x, \ t < \infty \\ \widetilde{u}^{\text{per}}(x, 0) = f_{\text{o}}^{\text{per}}(x), & -\infty < x < \infty \\ \widetilde{u}_t^{\text{per}}(x, 0) = g_{\text{o}}^{\text{per}}(x), & -\infty < x < \infty, \end{cases} \tag{15.3}$$

then we can invoke D'Alembert's formula. Conversely, we would need to check that the formula satisfies (15.1), in particular the boundary conditions.

The hard part is no longer differentiability at $x = 0$, or, indeed, at $x = 2Lk$ for $k \in \mathbb{Z}$. We handled that with the differentiability of odd extensions when we studied the semi-infinite problem. The new challenge is differentiability at $x = L$, and, more generally, $x = (2k+1)L$ for $k \in \mathbb{Z}$. However, this is not a terrible challenge. We are assuming

$$f(L) = f''(L) = g(L) = 0,$$

and the suggestive sketches above illustrate a more general truth: the odd periodic extensions are "odd at L."

More precisely, it turns out to be the case that $f_{\text{o}}^{\text{per}}(L - x) = -f_{\text{o}}^{\text{per}}(L + x)$ for $0 \le x \ leL$. (If $L = 0$, this would just be oddness.) Here is a closer sketch.
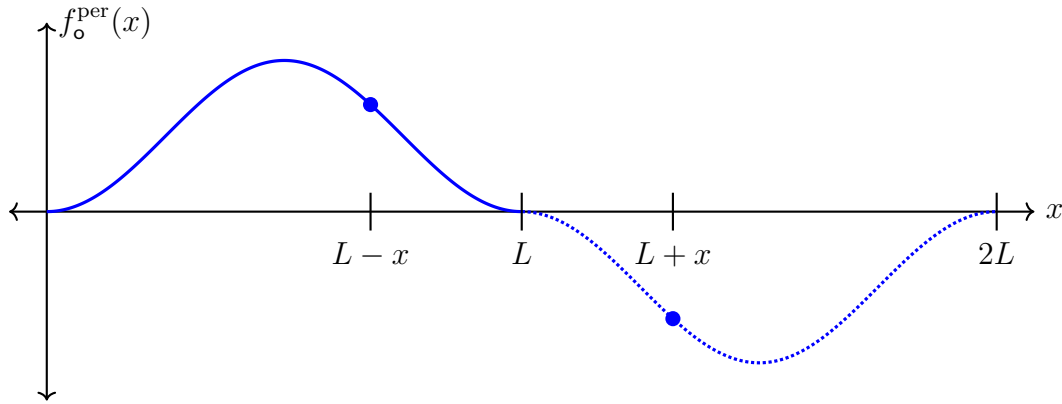
We leave checking the minutiae as a nontrivial, but manageable, problem.

15.2 Problem (+). Carry out the details of the program above, assuming $f \in \mathcal{C}^2([0, L])$ and $g \in \mathcal{C}^1([0, L])$ with (15.2) true. Remember that at the endpoints $x = 0$, L, we only know limits from the left/right for f and g and their derivatives.

(i) Show that $g_{\circ}^{\text{per}} \in \mathcal{C}^1(\mathbb{R})$. It may help to argue first that $g_{\circ}^{\text{per}}$ is "odd" about L in the sense that $g_{\circ}^{\text{per}}(L - x) = -g_{\circ}^{\text{per}}(L + x)$ for $0 \le x \le L$. Use this to show that $g_{\circ}^{\text{per}}$ is continuously differentiable at $x = L$. From this, argue by periodicity that $g_{\circ}^{\text{per}} \in \mathcal{C}^1(\mathbb{R})$.

(ii) Prove that $f_{\circ}^{\text{per}} \in \mathcal{C}^2(\mathbb{R})$. [Hint: *try to invoke the previous part.*]

(iii) Solve

$$\begin{cases} u_{tt} = u_{xx}, & -\infty < x, \ t < \infty \\ u(x,0) = f_{\circ}^{\text{per}}(x), & -\infty < x < \infty \\ u_t(x,0) = g_{\circ}^{\text{per}}(x), & -\infty < x < \infty \end{cases}$$

with D'Alembert's formula and check that the formula meets the boundary conditions $\widetilde{u}^{\text{per}}(0,t) = \widetilde{u}^{\text{per}}(L,t) = 0$. [Hint: *use the oddness of $f_{\circ}^{\text{per}}$, and also of $g_{\circ}^{\text{per}}$, at L, as discussed above.*]

The point of these extension methods is less the actual results and formulas and more the *methods* themselves—how to reduce a new problem to one already solved, and what techniques from calculus (left and right limits) appear along the way. Both extension procedures for the semi-infinite and finite wave equations are simultaneously existence and uniqueness results. There is another uniqueness, but not existence, method for the finite equation that is worth knowing. 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 $w := u - v$.

15.3 Problem (!). Check that

$$\begin{cases} w_{tt} = w_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ w(x,0) = 0, \ 0 \le x \le L \\ w_t(x,0) = 0, \ 0 \le x \le L \\ w(0,t) = w(L,t) = 0. \end{cases} \tag{15.4}$$

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.

15.4 Problem ($\star$). 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

$$w_t(x, t)^2 + w_x(x, t)^2 = 0 \tag{15.5}$$

for all $x \in [0, L]$ and $t \in \mathbb{R}$. We saw in Example 7.4 that (15.5) implies $w = 0$. (More precisely, that example presumed that (15.5) holds for all x, $t \in \mathbb{R}$; what really matters is that (15.5) holds for $x \in I$ and $t \in J$, with I, $J \subseteq \mathbb{R}$ intervals.)

Our last task is to justify the earlier claim that $E' = 0$. We have

$$E'(t) = \partial_t \left[\int_0^L \left[w_t(x, t)^2 + w_x(x, t)^2 \right] \, dx \right] = \int_0^L \partial_t \left[w_t(x, t)^2 + w_x(x, t)^2 \right] \, dx$$

$$= 2 \int_0^L \left[w_t(x, t) w_{tt}(x, t) + w_x(x, t) w_{xt}(x, t) \right] \, dx.$$

We replace w_{tt} with w_{xx} from the wave equation, so the integrand is

$$w_t(x, t) w_{tt}(x, t) + w_x(x, t) w_{xt}(x, t) = w_t(x, t) w_{xx}(x, t) + w_x(x, t) w_{xt}(x, t).$$

Here is the tricky recognition: it would be nice if this were a perfect derivative in x. We have

$$w_t(x, t) w_{xx}(x, t) + w_x(x, t) w_{xt}(x, t) = w_t(x, t) \partial_x[w_x](x, t) + w_x(x, t) \partial_x[w_t](x, t).$$

Perhaps we would see the product rule more clearly by rewriting

$$w_t(x, t) \partial_x[w_x](x, t) + w_x(x, t) \partial_x[w_t](x, t) = \partial_x[w_x](x, t) w_t(x, t) + w_x(x, t) \partial_x[w_t](x, t) = \partial_x[w_x w_t](x, t).$$

Thus

$$E'(t) = 2 \int_0^L \partial_x[w_x w_t](x, t) \, dx = 2 w_x(L, t) w_t(L, t) - 2 w_x(0, t) w_t(0, t).$$

We are given the boundary conditions $w(0, t) = w(L, t) = 0$ for all t, so differentiating we have $w_t(0, t) = w_t(L, t) = 0$ for all t. While we know nothing about w_x, this is enough to conclude

$$E'(t) = 2w_x(L, t) \cdot 0 - 2w_x(0, t) \cdot 0 = 0,$$

and that is all that we need.

Here is what we have shown.

15.5 Theorem. *The only solution to* (15.4) *is* $w = 0$.

15.6 Problem ($\star$). We did not introduce energy integrals with the transport equation because we did not really consider boundary conditions on that PDE. However, just for practice with differentiating under the integral, assume that $u_t + u_x = 0$ and let $a, b \in \mathbb{R}$. In each case, the identity $yy' = (y^2/2)'$, for $y \in \mathcal{C}^1(\mathbb{R})$, will be helpful.

(i) Let

$$E_1(t) := \int_a^b u(x, t)^2 \, dx.$$

Show that $E_1'(t) = u(a, t) - u(b, t)$.

(ii) Let

$$E_2(t) := \int_a^b \left(u_t(x, t)^2 + u_x(x, t)^2 \right) \, dx.$$

If $u \in \mathcal{C}^2(\mathbb{R}^2)$, show that $E_2'(t) = 2\left(u_t(a, t)^2 - u_t(b, t)^2 \right)$.

Day 16: Wednesday, February 12.

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 3.2, Theorem 4.1, Problem 6.8, Theorem 8.1, Theorem 9.6, Theorem 11.1, Theorem 12.4, Problem 13.1). 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."*

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

$$\mathcal{W}[f, g](x, t) := \frac{f(x + t) + f(x - t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} g(s) \, ds. \tag{16.1}$$

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

$$\begin{cases} u_{tt} = u_{xx}, \ -\infty < x, \ t < \infty \\ u(x,0) = f_1(x), \ -\infty < x < \infty \\ u_t(x,0) = g_1(x), \ -\infty < x < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, \ -\infty < x, \ t < \infty \\ v(x,0) = f_2(x), \ -\infty < x < \infty \\ v_t(x,0) = g_2(x), \ -\infty < x < \infty. \end{cases} \quad (16.2)$$

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.")
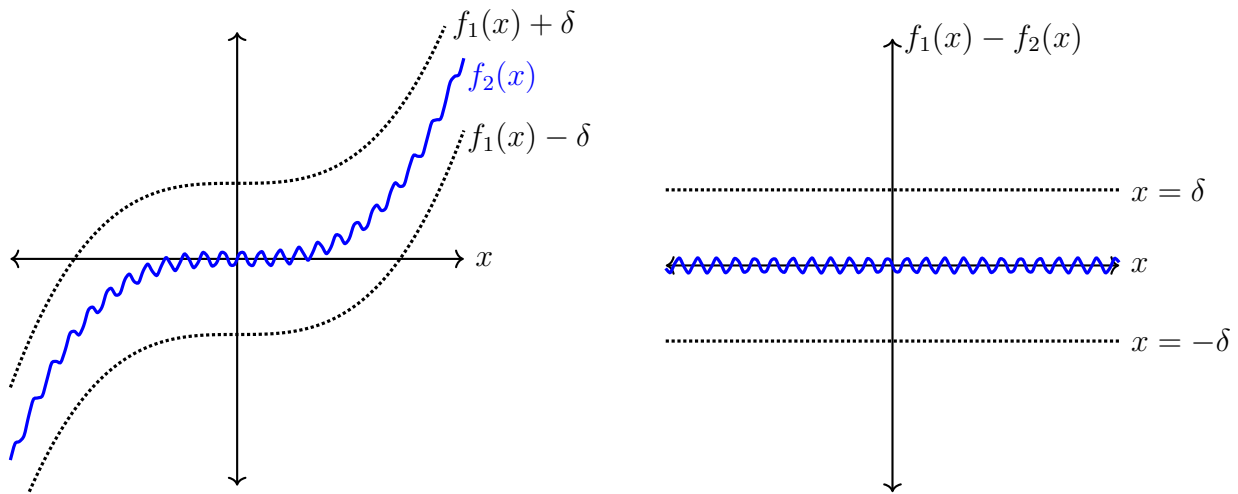
16.1 Problem (!). Suppose that u and v solve the wave IVP in (16.2). Let $w = u - v$, $f = f_1 - f_2$, and $g = g_1 - g_2$. Show that $w = \mathcal{W}[f, g]$ with $\mathcal{W}$ defined in (16.1).

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

$$|w(x,t)| \le \frac{|f(x+t) + f(x-t)|}{2} + \frac{1}{2}\left| \int_{x-t}^{x+t} g(s) \ ds \right|.$$

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 $t \ge 0$.

16.2 Problem (!). Show that if $t < 0$, then

$$\left|\int_{x-t}^{x+t} g(s)\ ds\right| < 2|t|\epsilon.$$

We conclude

$$|u(x,t) - v(x,t)| = |w(x,t)| = |\mathcal{W}[f,g](x,t)| < \delta + |t|\epsilon. \tag{16.3}$$

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.

16.3 Problem ($\star$). Here is a somewhat silly example of how this could occur. Let $\delta, \epsilon > 0$. 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 (16.2), then

$$u(x,t) = 0 \quad \text{and} \quad v(x,t) = \frac{\delta + \epsilon t}{2}.$$

Check explicitly that (16.3) still holds, but explain informally how u and v "grow apart" in time.

Day 17: Friday, February 14. ▬▬▬▬

We took Exam 1.

Day 18: Monday, February 17. ▬▬▬▬

> **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.
>
> 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.

The factor of $|t|$ in (16.3) 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 (16.3) 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

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ u(x,0) = f_1(x), \ 0 \le x \le L \\ u_t(x,0) = g_1(x), \ 0 \le x \le L \\ u(0,t) = u(L,t) = 0, \ -\infty < t < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ v(x,0) = f_2(x), \ 0 \le x \le L \\ v_t(x,0) = g_2(x), \ 0 \le x \le L \\ v(0,t) = v(L,t) = 0, \ -\infty < t < \infty. \end{cases}$$
$$(18.1)$$

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 $x \in [0, L]$.

18.1 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.

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

$$\int_c^{c+2L} h = 0 \tag{18.2}$$

for all $c \in \mathbb{R}$; in words, the integral of h over any interval of length $2L$ vanishes.

18.2 Problem ($\star$). Prove this. [Hint: *use the results of Problems 2.18 and 14.8.*]

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 $2L < b < 3L$.
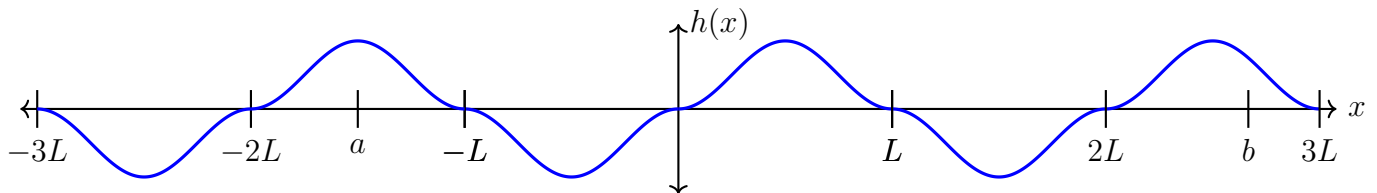
We expand

$$\int_a^b h = \int_a^{-L} h + \int_{-L}^L h + \int_L^b h. \tag{18.3}$$

By (18.2), 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

$$\left| \int_a^b h \right| < \epsilon(-L - a) + \epsilon(b - L) < L\epsilon + 2L\epsilon = 3L\epsilon. \tag{18.4}$$

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

$$(2j + 1)L \le a < (2j + 3)L \quad \text{and} \quad (2k + 1)L \le b < (2k + 3)L. \tag{18.5}$$

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 $j \le k$.

18.3 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 (18.3). Every integral in the parenthetical sum is 0 by (18.2). 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 (18.5) 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 (18.4).

> **18.4 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 (18.5)?

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 (18.1), 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.

> **18.5 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 δ, $\epsilon > 0$ with*
>
> $$|f_1(x) - f_2(x)| < \delta \quad and \quad |g_1(x) - g_2(x)| < \epsilon$$
>
> *for all $x \in \mathbb{R}$. Let u and v solve*
>
> $$\begin{cases} u_{tt} = u_{xx}, & -\infty < x,\ t < \infty \\ u(x,0) = f_1(x), & -\infty < x < \infty \\ u_t(x,0) = g_1(x), & -\infty < x < \infty \end{cases} \quad and \quad \begin{cases} v_{tt} = v_{xx}, & -\infty < x,\ t < \infty \\ v(x,0) = f_2(x), & -\infty < x < \infty \\ v_t(x,0) = g_2(x), & -\infty < x < \infty. \end{cases} \quad (18.6)$$
>
> *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 δ, $\epsilon > 0$ with*
>
> $$|f_1(x) - f_2(x)| < \delta \quad and \quad |g_1(x) - g_2(x)| < \epsilon$$

for all $x \in \mathbb{R}$. Let u and v solve

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \leq x \leq L, \ -\infty < t < \infty \\ u(x,0) = f_1(x), \ 0 \leq x \leq L \\ u_t(x,0) = g_1(x), \ 0 \leq x \leq L \\ u(0,t) = u(L,t) = 0, \ -\infty < t < \infty \end{cases} \quad and \quad \begin{cases} v_{tt} = v_{xx}, \ 0 \leq x \leq L, \ -\infty < t < \infty \\ v(x,0) = f_2(x), \ 0 \leq x \leq L \\ v_t(x,0) = g_2(x), \ 0 \leq x \leq L \\ v(0,t) = v(L,t) = 0, \ -\infty < t < \infty. \end{cases}$$

$$(18.7)$$

Then

$$|u(x,t) - v(x,t)| < \delta + 2L\epsilon$$

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

18.6 Problem ($\star$). State and prove an analogue of part (i) of Theorem 18.5 for the transport equation. Is your estimate uniform in time?

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; in particular, we will only get results for $t \geq 0$, and things will be rather complicated at $t = 0$. 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, so there should be no surprises there) and improper integrals (more good news: we just need the essentials from calculus, and there should be no surprises there, either).

Day 19: Wednesday, February 19. ▬▬▬▬

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).

Here is a terrible definition of complex numbers.

19.1 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.

19.2 Definition. *Let $z \in \mathbb{C}$ with $z = x + iy$ for some $x, \, y \in \mathbb{R}$. The **REAL PART** of z is $\mathrm{Re}(z) := x$; the **IMAGINARY PART** of z is $\mathrm{Im}(z) := y$; and the **MODULUS** of z is $|z| := \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$. 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 $\mathrm{Re}(z) = \mathrm{Re}(w)$ and $\mathrm{Im}(z) = \mathrm{Im}(w)$.

19.3 Example. With $z = 2 + i$ and $w = 1 - 3i$, we multiply as we would with real numbers and remember $i^2 = -1$:

$$zw = (2+i)(1-3i) = (2+i)1 + (2+i)(-3i) = 2 + i - 6i - 3i^2 = 2 - 5i - 3(-1) = 2 - 5i + 3$$
$$= 5 - 5i = 5(1 - i).$$

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
$$f_1(t) := \mathrm{Re}[f(t)] \quad \text{and} \quad f_2(t) := \mathrm{Im}[f(t)].$$
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.

19.4 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*

$$\int_a^b f := \int_a^b f_1 + i \int_a^b f_2.$$

We now allow $\mathcal{C}^r(I)$ *to denote the set of r-times continuously differentiable functions from* $I \subseteq \mathbb{R}$ *to* $\mathbb{C}$.

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.

19.5 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.

19.6 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 $e^{2\pi i k} = 1$ for all $k \in \mathbb{Z}$.

Now we develop further results on integrals.

19.7 Definition. *Let $f \in \mathcal{C}(\mathbb{R})$. Suppose that both of the limits*

$$\int_{-\infty}^0 f := \lim_{a \to -\infty} \int_0^a f \quad and \quad \int_0^\infty f := \lim_{b \to \infty} \int_0^b f$$

exist. Then we say that f is **INTEGRABLE**, *and we define*

$$\int_{-\infty}^\infty f := \int_{-\infty}^0 f + \int_0^\infty f.$$

19.8 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

$$\lim_{a \to -\infty} \int_a^0 f = \lim_{a \to -\infty} (1 - e^a) = 1 \quad and \quad \lim_{b \to \infty} \int_0^b f = \lim_{b \to \infty} (1 - e^{-b}) = 1,$$

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. They hinge on integrating *nonnegative* functions.

19.9 Definition. *A function $f \in \mathcal{C}(\mathbb{R})$ is* **ABSOLUTELY INTEGRABLE** *if $|f|$ is integrable.*

19.10 Theorem. *Let $f \in \mathcal{C}(\mathbb{R})$.*

(i) [Absolute integrability implies integrability] *If f is absolutely integrable, then so is f, and the* TRIANGLE INEQUALITY *holds:*

$$\left| \int_{-\infty}^{\infty} f \right| \leq \int_{-\infty}^{\infty} |f|.$$

(ii) [Comparison test] *Suppose that $g \in \mathcal{C}(\mathbb{R})$ is absolutely integrable with $|f(t)| \leq |g(t)|$ for all t. Then f is absolutely integrable with*

$$\int_{-\infty}^{\infty} |f| \leq \int_{-\infty}^{\infty} |g|.$$

19.11 Example. Let $f \in \mathcal{C}(\mathbb{R})$ be absolutely integrable. Let $k \in \mathbb{R}$ and put $h_k(t) := f(t)e^{ikt}$. Since $|e^{is}| = 1$ for all $s \in \mathbb{R}$ (check it), we have

$$|h_k(t)| = |f(t)e^{ikt}| = |f(t)||e^{ikt}| = |f(t)|,$$

and so by the comparison test (with actual equality holding), the functions h and $|h|$ are integrable.

19.12 Problem ($\star$). Let $a > 0$ and let $f(x) = e^{-ax^2}$. Show that f is integrable. [Hint: first find $C > 0$ such that $e^{ax - ax^2} \leq C$ for $0 \leq x \leq 1$. Then argue that $e^{-ax^2} \leq e^{-ax}$ for $x \geq 1$. Put these estimates together to show $e^{-ax^2} \leq (C+1)e^{-ax}$ for $x \geq 0$.]

19.13 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 $f \in \mathcal{C}(\mathbb{R})$ 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

$$\text{P.V.} \int_{-\infty}^{\infty} f := \lim_{R \to \infty} \int_{-R}^{R} f.$$

(i) Give an example of $f \in \mathcal{C}(\mathbb{R})$ such that $\lim_{R \to \infty} \int_{-R}^{R} f$ exists and yet f is not integrable.

(ii) If, however, f is integrable, then $\int_{-\infty}^{\infty} f = \text{P.V.} \int_{-\infty}^{\infty} f$. Here is why. Assume that $f \in \mathcal{C}(\mathbb{R})$ 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 ($g(x_1) \leq g(x_2)$ for $0 \leq x_1 \leq x_2$), and bounded above (there is $M > 0$ such that $0 \leq g(x) \leq M$ for all $x \geq 0$), 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 $\left| \int_{-R}^{R} f \right| \leq M$ for all $R \geq 0$, then f is integrable. [Hint: *apply the result to the functions* $R \mapsto \int_{0}^{R} f$ *and* $R \mapsto \int_{-R}^{0} f$.]

(iv) Prove that if $f \in \mathcal{C}(\mathbb{R})$ and if $\lim_{R \to \infty} \int_{-R}^{R} |f|$ exists, then f is absolutely integrable.

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).

Example 19.11 assures us that the following definition makes sense. (Does it?)

19.14 Definition. *Let* $f \colon \mathbb{R} \to \mathbb{C}$ *be continuous with* $|f|$ *integrable (Definition 19.7). 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.

19.15 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 $x \in \mathbb{R}$.

Day 20: Friday, February 21.

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.

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.

20.1 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:

$$\widehat{u_t}(k,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_t(x,t)e^{-ikx} \, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \partial_t[u(x,t)e^{-ikx}] \, dx$$

$$= \partial_t\left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x,t)e^{-ikx} \, dx\right] = \partial_t[\widehat{u}](k,t).$$

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 19.15:

$$u(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{u}(k, t)e^{ikx} \, dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)e^{-k^2 t}e^{ikx} \, dk. \tag{20.1}$$

This may well be a valid candidate for a solution formula!

20.2 Problem ($\star$). (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 $|\widehat{f}(k)| \leq M$ for all k), then g is integrable, and so the integral on the right in (20.1) 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 (20.1) 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 (20.1) satisfies $u_t = u_{xx}$.

(iii) Show that u as defined by (20.1) meets $u(x, 0) = f(x)$. [Hint: *Untheorem 19.15.*]

20.3 Problem ($\star$). Repeat the work above for the transport IVP

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

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 19.15. 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 (20.1) actually gives a formula for a solution to the heat equation. Problem 20.2 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 $t \leq 0$? 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 (20.1) 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

$$u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y)e^{-iky} e^{-k^2 t} e^{ikx} \ dy \ dk. \tag{20.2}$$

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 19.12 and the comparison test ensure that this integral converges. Then our solution candidate should be

$$u(x,t) = \int_{-\infty}^{\infty} H(x-y,t)f(y) \ dy. \tag{20.3}$$

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.

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 $X \to \pm\infty$.
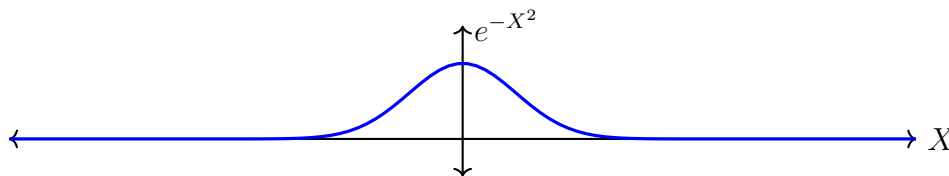
Now let $\mathcal{G}(\sqrt{t}\cdot)$ be the function

$$\mathcal{G}(\sqrt{t}\cdot) \colon \mathbb{R} \to \mathbb{R} \colon k \mapsto e^{-(k\sqrt{t})^2}.$$

Then

$$H(s,t) = \frac{1}{\sqrt{2\pi}} \widehat{\mathcal{G}(\sqrt{t}\cdot)}(-s). \tag{20.4}$$

Problem 20.2 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

$$g(\alpha\cdot) \colon \mathbb{R} \to \mathbb{C} \colon x \mapsto g(\alpha x).$$

> **20.4 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

$$\widehat{g(\alpha\cdot)}k = \frac{1}{\alpha} \widehat{g}\left(\frac{k}{\alpha}\right). \tag{20.5}$$

> **20.5 Problem ($\star$).** 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*

20.6 Problem (⋆). Let $g: \mathbb{R} \to \mathbb{C}$ be continuous with $|g|$ integrable. Let $d \in \mathbb{R}$. Prove that the "shifted map"

$$S^d g: \mathbb{R} \to \mathbb{C}: x \mapsto g(x+d)$$

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 19.7 exist and then use Problem 19.13 to express* $\int_{-\infty}^{\infty} S^d f = \lim_{R \to \infty} \int_{-R}^{R} S^d f$.]

We combine (20.4) and (20.5) to obtain

$$H(s,t) = \frac{1}{\sqrt{2\pi}} \widehat{\mathcal{G}(\sqrt{t}\cdot)}(-s) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sqrt{t}}\right) \widehat{\mathcal{G}}\left(-\frac{s}{\sqrt{t}}\right). \tag{20.6}$$

So, what is $\widehat{\mathcal{G}}$?

Day 21: Monday, February 24.

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 $\lim_{t \to 0^+} \int_{-\infty}^{\infty} H(x-y,t)f(y) \ dy = f(x)$. 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.

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

$$ik\widehat{\mathcal{G}}(k) = \frac{2}{i}\partial_k\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2}e^{-ikx} \ dx\right] = \frac{2}{i}\partial_k[\widehat{\mathcal{G}}](k). \tag{21.1}$$

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 (21.1): 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

$$\int_{-\infty}^{\infty} e^{-x^2} \ dx = \sqrt{\pi}, \tag{21.2}$$

an identity that we will not prove here. Thus

$$\widehat{\mathcal{G}}(k) = \frac{\sqrt{\pi}}{\sqrt{2\pi}}e^{-k^2/4} = \frac{e^{-k^2/4}}{\sqrt{2}}. \tag{21.3}$$

All together, we conclude from (20.6) and (21.3) 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.

21.1 Problem (!). Check that H satisfies the heat equation in the sense that $H_t = H_{ss}$.

At last we return to our candidate solution (20.3) for the heat equation:

$$u(x,t) = \int_{-\infty}^{\infty} H(x-y,t)f(y) \, dy = \int_{-\infty}^{\infty} \frac{e^{-(x-y)^2/4t}}{\sqrt{4\pi t}} f(y) \, dy, \quad x \in \mathbb{R}, \ t > 0. \qquad (21.4)$$

With this explicit formula for H, we can prove the convergence of this integral with *two* different hypotheses on f.

> **21.2 Problem ($\star$).** Prove that the (second) integral in (21.4) 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 (21.4) 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 21.1 implies

$$u_t(x,t) = \partial_t \left[\int_{-\infty}^{\infty} H(x-y,t)f(y) \, dy \right] = \int_{-\infty}^{\infty} \partial_t[H(x-y,t)f(y)] \, dy = \int_{-\infty}^{\infty} H_{ss}(x-y,t)f(y) \, dy$$

$$= \int_{-\infty}^{\infty} \partial_x^2[H(x-y,t)f(y)] \, dy = \partial_x^2 \left[\int_{-\infty}^{\infty} H(x-y,t)f(y) \, dy \right] = u_{xx}(x,t).$$

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

$$\lim_{t \to 0^+} \int_{-\infty}^{\infty} H(x-y,t)f(y) \, dy = f(x). \qquad (21.5)$$

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

$$u(x,t) = \begin{cases} \int_{-\infty}^{\infty} H(x-y,t)f(y) \, dy, & x \in \mathbb{R}, \ t > 0 \\ f(x), & x \in \mathbb{R}, t = 0, \end{cases} \qquad (21.6)$$

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.

> **21.3 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.

> **21.4 Theorem.** *Suppose that u solves the heat equation*
>
> $$\begin{cases} u_t = u_{xx}, & -\infty < x < \infty, \ t > 0 \\ u(x,0) = 0, & -\infty < x < \infty \end{cases} \tag{21.7}$$
>
> *and is continuous on $\{(x,t) \in \mathbb{R}^2 \mid x \in \mathbb{R}, \ t \geq 0\}$. Suppose also that u is **BOUNDED IN FINITE TIME** in the sense that for all $T > 0$, there is $M_T > 0$ such that $|u(x,t)| \leq M_T$ for all $x \in \mathbb{R}$ and $t \in [0,T]$. Then $u(x,t) = 0$ for all $x \in \mathbb{R}$ and $t \geq 0$.*

> **21.5 Problem ($\star$).** Give an example of a function u that is bounded in finite time, that satisfies (21.7), that is discontinuous on $\{(x,t) \in \mathbb{R}^2 \mid x \in \mathbb{R}, \ t \geq 0\}$, and that is not identically zero. Thus the continuity requirement in Theorem 21.4 really is necessary.

> **21.6 Problem ($\star$).** Which, if any, traveling wave solutions $u(x,t) = p(x - ct)$ to the heat equation $u_t = u_{xx}$ are bounded in finite time?

We will prove this theorem later after we develop some new tools. Here is how it helps.

> **21.7 Problem (!).** Suppose that u and v are continuous on $\{(x,t) \in \mathbb{R}^2 \mid x \in \mathbb{R}, \ t \geq 0\}$, 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 21.4.]

We might then ask what conditions on f guarantee that the solution to the heat IVP as given by (21.6) 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 $|f(x)| \le M$ for all $x \in \mathbb{R}$ and some $M > 0$. The previous paragraph shows that with u defined by (21.6), we have $|u(x, 0)| \le M$ for all $x \in \mathbb{R}$. Now we estimate $u(x, t)$ for $t > 0$. The triangle inequality for integrals implies

$$|u(x,t)| \le \frac{M}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4t} \, dy = \frac{M}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-[(x-y)/2\sqrt{t}]^2} \, dy.$$

21.8 Problem (!). Substitute $s = -(x-y)/2\sqrt{t}$ in the integral on the right and use (21.2) 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).

21.9 Theorem. *Let $f \in \mathcal{C}(\mathbb{R})$ be bounded in the sense that there exists $M > 0$ such that $|f(x)| \le M$ for all $x \in \mathbb{R}$. Then the only solution to the heat IVP*

$$\begin{cases} u_t = u_{xx}, & -\infty < x < \infty, \ t > 0 \\ u(x, 0) = f(x), & -\infty < x < \infty \end{cases} \tag{21.8}$$

that is continuous on $\{(x, t) \in \mathbb{R}^2 \mid x \in \mathbb{R}, \ t \ge 0\}$ is

$$u(x, t) = \begin{cases} \int_{-\infty}^{\infty} H(x - y, t) f(y) \, dy, & x \in \mathbb{R}, \ t > 0 \\ f(x), & x \in \mathbb{R}, \ t = 0, \end{cases} \qquad H(s, t) = \frac{e^{-s^2/4t}}{\sqrt{4\pi t}}.$$

Day 22: Wednesday, February 26.

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.

So what else is this solution *doing*? First, no matter what the initial temperature distribution is, eventually everything "cools all the way down."

22.1 Problem ($\star$). Let $f \in \mathcal{C}(\mathbb{R})$ be bounded and let $|f|$ be integrable. Let u solve (21.8). 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).

22.2 Problem ($\star$). Let $a, b, c, d \in \mathbb{R}$ with $a < c < d \leq b$. Suppose that $f \in \mathcal{C}([a, b])$ with $f(x) \geq 0$ for $x \in [a, b]$. Prove that

$$\int_c^d f \leq \int_a^b f. \tag{22.1}$$

If $f(x) > 0$ for $x \in [a, b]$, show that the nonstrict inequality in (22.1) becomes strict. [Hint: *start with a picture and check, using fundamental properties of integrals, 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.

22.3 Problem ($\star$). Let $f \in \mathcal{C}(\mathbb{R})$ be bounded and nonnegative and suppose that u solves (21.8). 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.*]

22.4 Problem ($\star$). Prove the following "comparison" principle for the heat equation. Suppose that $f_1, f_2 \in \mathcal{C}(\mathbb{R})$ are bounded with $f_1(x) \leq f_2(x)$ for all x and u and v solve

$$\begin{cases} u_t = u_{xx}, \ x \in \mathbb{R}, \ t > 0 \\ u(x, 0) = f_1(x), \ x \in \mathbb{R} \end{cases} \quad \text{and} \quad \begin{cases} v_t = v_{xx}, \ x \in \mathbb{R}, \ t > 0 \\ v(x, 0) = f_2(x), \ x \in \mathbb{R}. \end{cases}$$

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 $t > 0$.

22.5 Problem (+). Our initial value problems usually start at time $t = 0$ (hence the adjective "initial") but that is not strictly necessary. We might define a "backward" version of one of our favorite problems by specifying the "initial" condition at some time $t_0 > 0$ and then asking about the "genuine" initial data at $t = 0$.

(i) Suppose that u solves

$$\begin{cases} u_t + u_x = 0, \ (x,t) \in \mathbb{R}^2 \\ u(x,t_0) = g(x), \ x \in \mathbb{R} \end{cases}$$

for some $t_0 > 0$. Find a function f (in terms of g and t_0) such that $u(x,0) = f(x)$.

(ii) Prove that if u and v solve

$$\begin{cases} u_t + u_x = 0, \ (x,t) \in \mathbb{R}^2 \\ u(x,t_0) = g_1(x), \ x \in \mathbb{R} \end{cases} \quad \text{and} \quad \begin{cases} v_t + v_x = 0, \ (x,t) \in \mathbb{R}^2 \\ v(x,t_0) = g_2(x), \ x \in \mathbb{R} \end{cases}$$

for some functions g_1, g_2 with $|g_1(x) - g_2(x)| < \epsilon$ for all $x \in \mathbb{R}$, then

$$|u(x,t) - v(x,t)| < \epsilon$$

for all $(x,t) \in \mathbb{R}^2$. We might therefore say that the transport equation is "backwards stable" in the sense that if two solutions are close at some point in time (not necessarily time $t = 0$), then they are close for all time.

(iii) Let $n \geq 1$ be an integer. The function

$$u_n(x,t) := \frac{e^{n^2(1-t)} \sin(nx)}{n}$$

solves

$$\begin{cases} (u_n)_t = (u_n)_{xx}, \ (x,t) \in \mathbb{R}^2 \\ u_n(x,1) = g_n(x), \ x \in \mathbb{R}, \qquad g_n(x) := \dfrac{\sin(nx)}{n}. \end{cases}$$

Of course, the function $v(x,t) = 0$ solves

$$\begin{cases} v_t = v_{xx}, \ (x,t) \in \mathbb{R}^2 \\ v(x,0) = g_0(x), \ x \in \mathbb{R}. \end{cases} \qquad g_0(x) := 0.$$

(You do not have to check this; note that here the n on u_n is an indexing subscript, whereas the partial derivatives are $(u_n)_t$ and $(u_n)_{xx}$.) Instead, show that the heat equation is not "backwards stable" as follows. Let $\epsilon > 0$. Explain why for n large we have

$$|g_n(x) - g_0(x)| < \epsilon$$

for all $x \in \mathbb{R}$ but there will also exist $(x_n, t_n) \in \mathbb{R}^2$ with $t_n < 1$ such that

$$\frac{1}{\epsilon} < |u_n(x_n, t_n) - v(x_n, t_n)|.$$

That is, the solutions u_n and v can be arbitrarily close at time 1, but that says nothing about their closeness earlier in time. [Hint: *for n large, we will have $1/n < \epsilon$, but since* $\lim_{n \to \infty} e^{n^2}/n = \infty$, *for n large(r), we will also have $e^{n^2}/n > 1/\epsilon$.*]

We turn our attention to proving Theorem 21.4. 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.

The following result captures the more or less physically reasonable idea that the temperature in a finite rod with no internal or external heat source reaches its extreme values either at the initial time of measurement or sometime subsequently at the endpoints. Heat should flow throughout the rod and "even out," not aggregate at some internal point.

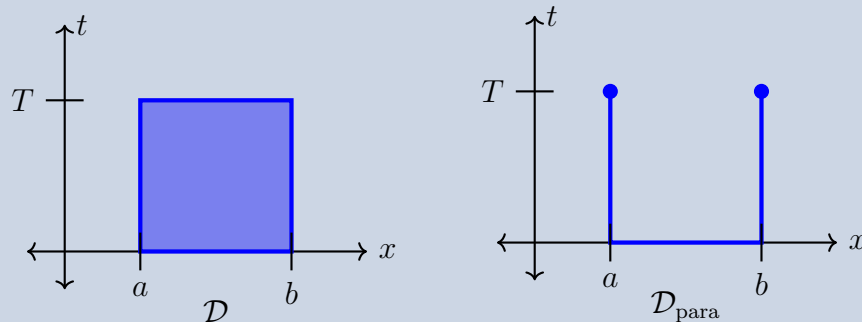**22.6 Theorem (Maximum principle).** *Suppose that u solves*

$$u_t = u_{xx}$$

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

$$\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^2 \mid a \le x \le b, \ 0 \le t \le T \right\}$$

and

$$\mathcal{D}_{\text{para}} := \left\{ (a,t) \in \mathbb{R}^2 \mid 0 \le t \le T \right\} \cup \left\{ (x,0) \in \mathbb{R}^2 \mid a \le x \le b \right\} \cup \left\{ (b,t) \in \mathbb{R}^2 \mid 0 \le t \le T \right\}.$$

Then

$$\max_{(x,t) \in \mathcal{D}} u(x,t) = \max_{(x,t) \in \mathcal{D}_{\text{para}}} u(x,t).$$

Proof. This is possibly one of the most demanding proofs in the course, so we give it in multiple separate steps.

1. 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
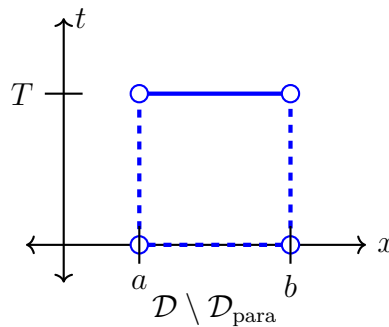
$$M := u(x_0, t_0) = \max_{(x,t) \in \mathcal{D}} u(x, t).$$

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
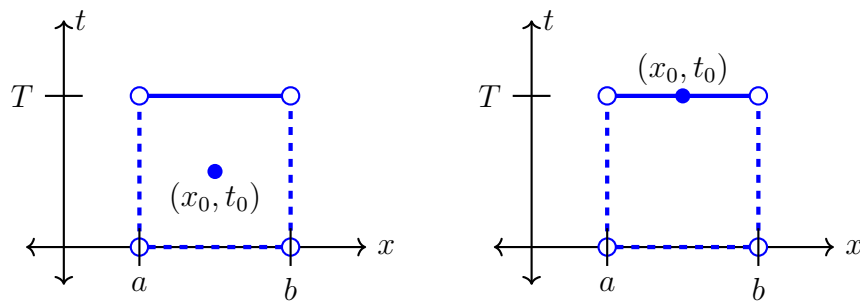
$$m := \max_{(x,t) \in \mathcal{D}_{\text{para}}} u(x, t) \le M.$$

Our goal is to show $m = M$; we proceed by contradiction and assume $m < M$. Then we must have $(x_0, t_0) \in \mathcal{D} \setminus \mathcal{D}_{\text{para}}$, as otherwise with $(x_0, t_0) \in \mathcal{D}_{\text{para}}$, we have $M \le m$ and thus $m = M$.

2. Here is a sketch of $\mathcal{D} \setminus \mathcal{D}_{\text{para}}$.

With $(x_0, t_0) \in \mathcal{D} \setminus \mathcal{D}_{\text{para}}$, we have $a < x_0 < b$ and $0 < t_0 \le T$. Here are sketches of the two possible locations for (x_0, t_0), then: either in the "interior" or on the "upper boundary."

3. Our proof largely relies on a trick and an appeal to results from single-variable calculus, which we briefly review here (and discuss further in Problem 22.8 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 $f'(x_\star) \ge 0$.

4. Here is how we use these results. 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 $u_{xx}(x_0, t_0) \leq 0$, since the function $u(\cdot, t_0)$ achieves its maximum on $[a, b]$ at $x_0 \in (a, b)$. And we also know $u_t(x_0, t_0) = u_{xx}(x_0, t_0)$. If we could get a strict inequality of the form $u_t(x_0, t_0) > 0$, or $u_{xx}(x_0, t_0) < 0$, then we would get a contradiction. (This would happen, for example, if we could rule out $t_0 = T$, then we would have $0 < t_0 < T$, and so $u_t(x_0, t_0) > 0$.) However, we do not have enough information to do any of that, and at best we know $u_t(x_0, t_0) \geq u_{xx}(x_0, t_0)$.

5. Instead, the trick of the proof is to modify u into a new function v, which has a slightly more helpful second derivative in x. Then we will apply these calculus results to v and learn something more useful. Probably the simplest modification to u is to add another function to u (since the heat equation is linear, this is a good idea), and perhaps 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.

6. Now we think about extreme values; certainly by continuity, v has a maximum on $\mathcal{D}$. We are going to show that

$$\max_{(x,t) \in \mathcal{D}} v(x, t) = v(x_1, t_1)$$

for some $(x_1, t_1) \in \mathcal{D} \setminus \mathcal{D}_{\text{para}}$. This will allow us to apply exactly the same reasoning as above to conclude

$$v_t(x_1, t_1) \geq v_{xx}(x_1, t_1). \tag{22.2}$$

But now we can compute

$$v_t(x, t) = u_t(x, t) \quad \text{and} \quad v_{xx}(x, t) = u_{xx}(x, t) + 2\epsilon.$$

Then the inequality (22.2) reads

$$u_t(x_1, t_1) \geq u_{xx}(x_1, t_1) + 2\epsilon,$$

which means

$$u_t(x_1, t_1) > u_{xx}(x_1, t_1).$$

Since u solves the heat equation and $u_t(x_1, t_1) = u_{xx}(x_1, t_1)$, that is the contradiction. Perhaps surprisingly, we will not end up showing $u_t(x_0, t_0) > u_{xx}(x_0, t_0)$, which is perhaps what we thought the original contradiction would be.

7. Now we focus on the maximum of v over $\mathcal{D}$. First, if $(x, t) \in \mathcal{D}_{\text{para}}$, then $x^2 \leq \max\{a^2, b^2\} =: C$. (This requires a bit of thought if $a < 0$ or $b < 0$.) So,

$$\max_{(x,t) \in \mathcal{D}_{\text{para}}} v(x, t) \leq m + C\epsilon. \tag{22.3}$$

Since we are assuming $m < M$, if we take $\epsilon > 0$ small enough relative to m, M, a^2, and b^2, then

$$m + C\epsilon < M. \tag{22.4}$$

And
$$v(x_0, t_0) = M + \epsilon x_0^2 \geq M.$$
(We have to keep the nonstrict inequality just in case $x_0 = 0$.) So, since $(x_0, t_0) \in \mathcal{D}$,
$$\max_{(x,t)\in\mathcal{D}} v(x,t) \geq M. \tag{22.5}$$

8. We combine (22.3), (22.4), and (22.5) 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 $(x_1, t_1) \in \mathcal{D} \setminus \mathcal{D}_{\text{para}}$. This leads to the reasoning above and the contradiction. ∎

22.7 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!)

22.8 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

$$f(a) = \max_{a \leq x \leq b} f(x).$$

(i) Use the definition of the derivative to prove that $f'(a) \leq 0$.

(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 $f'(a) \leq 0$.

22.9 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.

Day 23: Friday, February 28.

Now we start to prove Theorem 21.4. 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 $0 < t \leq T$. 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 $t_0 \in (0, T]$.

We will prove that $u(x_0, t_0) = 0$ by showing that $|u(x_0, t_0)| \leq \epsilon$ for all $\epsilon > 0$. Here is the strategy.

1. We are going to construct a "comparator" function v such that $v(x, t) \leq 0$ 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 $u(x_0, t_0) \leq \epsilon$, 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 $v(x, t) \leq 0$. By the claims above, this shows $u(x, t) \leq \epsilon$.

4. We will leave as a problem proving the inequality $-\epsilon \leq u(x, t)$ via similar means.

Here is w:
$$w(x, t) := \frac{x^2 + 2t}{x_0^2 + 2t_0}. \tag{23.1}$$

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.

> **23.1 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: $-r \leq x \leq r$ for some $r > 0$. That is, we take

$$\mathcal{D} = \left\{ (x, t) \in \mathbb{R}^2 \mid -r \leq x \leq r, \ 0 \leq t \leq T \right\}.$$
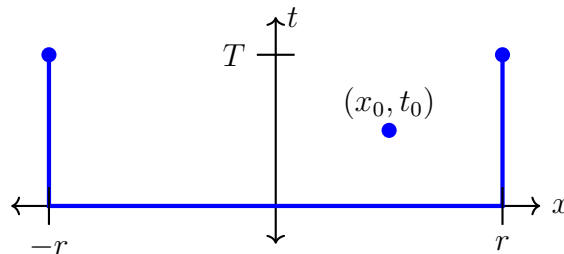
The maximum principle then guarantees that

$$\max_{(x,t) \in \mathcal{D}} v(x, t) = \max_{(x,t) \in \mathcal{D}_{\text{para}}} v(x, t),$$

where

$$\mathcal{D}_{\text{para}} = \left\{ (-r, t) \in \mathbb{R}^2 \mid 0 \leq t \leq T \right\} \cup \left\{ (x, 0) \in \mathbb{R}^2 \mid -r \leq x \leq r \right\} \cup \left\{ (r, t) \mid 0 \leq t \leq T \right\}$$

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

$$w(\pm r, t) = \frac{r^2 + 2t}{x_0^2 + 2t_0} \geq \frac{r^2}{x_0^2 + 2t_0},$$

so here, using the finite time bound on u (which says $u(x, t) \leq |u(x, t)| \leq M_T$ for $x \in \mathbb{R}$ and $t \in [0, T]$)

$$v(\pm r, t) \leq u(\pm r, t) - \epsilon \frac{r^2}{x_0^2 + 2t_0} \leq M_T - \epsilon \frac{r^2}{x_0^2 + 2t_0}.$$

If we take r sufficiently large relative to M_T, x_0, t_0, and ϵ, then

$$M_T - \frac{r^2}{x_0^2 + 2t_0} \leq 0.$$

On the horizontal side, we have $|x| \leq r$ and $t = 0$, so

$$w(x, 0) = \epsilon \frac{x^2}{x_0^2 + 2t_0} \geq 0,$$

and so here, using the initial condition $u(x, 0) = 0$,

$$v(x, 0) = u(x, 0) - w(x, 0) = -w(x, 0) \leq 0.$$

Thus $v(x, t) \leq 0$ on the parabolic boundary, as desired. The reasoning above implies $u(x_0, t_0) \leq \epsilon$, and we claim the other inequality holds.

23.2 Problem ($\star$). 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 $-\epsilon \leq u(x_0, t_0)$, which completes the argument.

Day 24: Monday, March 3.

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

Theorem 1 on p. 140 proves uniqueness for the finite rod heat equation via an energy integral argument. Theorem 1 on pp. 482–483 discusses how to use the Fourier transform to develop an explicit solution for the finite rod problem. 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.

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. One such problem is

$$\begin{cases} u_t = u_{xx}, \ a \le x \le b, \ t > 0 \\ u(x,0) = f(x), \ a \le x \le b \\ u(a,t) = \ell(t), \ u(b,t) = r(t), \ t \ge 0, \end{cases} \tag{24.1}$$

for given functions $f \colon [a,b] \to \mathbb{R}$ and $\ell, \ r \colon \mathbb{R} \to \mathbb{R}$. This models a rod for which the temperature at the left and right endpoints is known; other boundary conditions incorporate "insulation" to prevent heat flow, which introduces requirements like $u_x(a,t) = 0$ and/or $u_x(b,t) = 0$.

As with the wave equation, we can use an energy method to prove uniqueness.

24.1 Problem (!). Explain why to show uniqueness of (24.1), it suffices to show that if u solves

$$\begin{cases} u_t = u_{xx}, \ a \le x \le b, \ t \ge 0 \\ u(x,0) = 0, \ a \le x \le b \\ u(a,t) = u(b,t) = 0, \ t \ge 0, \end{cases} \tag{24.2}$$

then $u = 0$.

24.2 Theorem. *There is only one solution to (24.2).*

Proof. We assume that u solves (24.1) and study the "energy integral"

$$E(t) := \int_a^b u(x,t)^2 \ dx.$$

We have $E(t) \ge 0$ since the integrand is nonnegative and

$$E(0) = \int_a^b u(x,0)^2 \ dx = \int_a^b 0 \ dx = 0$$

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 (24.2) is really twice continuously differentiable for $0 \le x \le L$ and $t > 0$, per our PDE conventions, so Leibniz's rule applies. We have

$$E'(t) = \partial_t \left[\int_a^b u(x,t)^2 \ dx \right] = \int_a^b \partial_t[u(x,t)^2] \ dx = \int_a^b 2u(x,t)u_t(x,t) \ dx$$

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

$$E'(t) = 2 \int_a^b u(x,t) u_{xx}(x,t) \ dx.$$

If in a calculus class we encountered an antidifferentiation problem of the form $\int g g''$, we would probably integrate by parts with $u = g$ (terrible notation here, given the role of u in the heat equation) and $dv = g''$ to conclude $\int g g'' = g g' - \int g' g'$. Doing so here gives

$$E'(t) = 2u(x,t) u_x(x,t)\big|_{x=a}^{x=b} - 2 \int_a^b u_x(x,t)^2 \ dx.$$

Since $u(a,t) = u(b,t) = 0$ by the boundary conditions, the first terms are 0, and so

$$E'(t) = -2 \int_a^b u_x(x,t)^2 \ dx \le 0,$$

as desired. ∎

24.3 Problem (!). Describe in words (using as few symbols as possible) how this energy argument for the heat equation both resembles and differs from the energy argument for the wave equation that proved Theorem 15.5.

The proof of Theorem 24.2 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

$$E[w] := \int_a^b w(x,t)^2 \ dx$$

is decreasing, and so $E[w](t) \le E[w](0)$ for all $t \ge 0$.

24.4 Problem (!). Reread the proof of Theorem 24.2 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:

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ u(x,0) = a(t), \ t \ge 0 \\ u(L,t) = b(t), \ t \ge 0 \\ u(x,0) = f_1(x), \ 0 \le x \le L \end{cases} \quad \text{and} \quad \begin{cases} v_t = v_{xx}, \ 0 \le x \le L, \ t \ge 0 \\ v(x,0) = a(t), \ t \ge 0 \\ v(L,t) = b(t), \ t \ge 0 \\ v(x,0) = f_2(x), \ 0 \le x \le L. \end{cases}$$

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

$$E[w](t) = \int_a^b w(x,t)^2 \ dx = \int_a^b \left[u(x,t) - v(x,t) \right]^2 \ dx$$

and

$$E[w](0) = \int_a^b w(x,0)^2 \ dx = \int_a^b \left[f_1(x) - f_2(x) \right]^2 \ dx.$$

That is,

$$\int_a^b \left[u(x,t) - v(x,t) \right]^2 \ dx \le \int_a^b \left[f_1(x) - f_2(x) \right]^2 \ dx \tag{24.3}$$

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 19.15 (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.
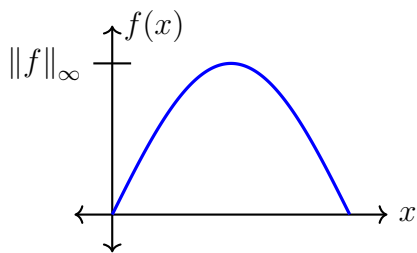
We might call this number $\|f\|_\infty$ the "maximum norm" of f. This value $\|f\|_\infty$ measures how large in a "pointwise" sense f can be.

We might also think about average value. The number

$$\frac{1}{L} \int_a^b |f(x)| \; dx$$

is the average value of $|f|$ on $[0, L]$. Put

$$\|f\|_{L^1} := \int_a^b |f(x)| \; dx.$$

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":

$$\|f\|_{L^2} := \left(\int_a^b |f(x)|^2 \; dx \right)^{1/2}.$$

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 (24.3) 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.
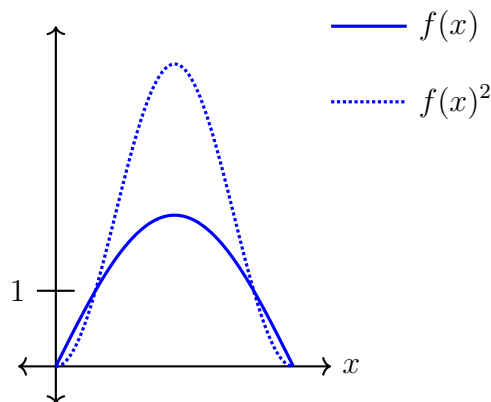
Our gut instinct is probably to want estimates in $\|\cdot\|_\infty$, as this is what we know best from calculus so far. Unfortunately, an estimate in $\|\cdot\|_{L^1}$ or $\|\cdot\|_{L^2}$ need not imply an estimate in $\|\cdot\|_\infty$. Consider the following picture, where ϵ, L, $M > 0$ are fixed.
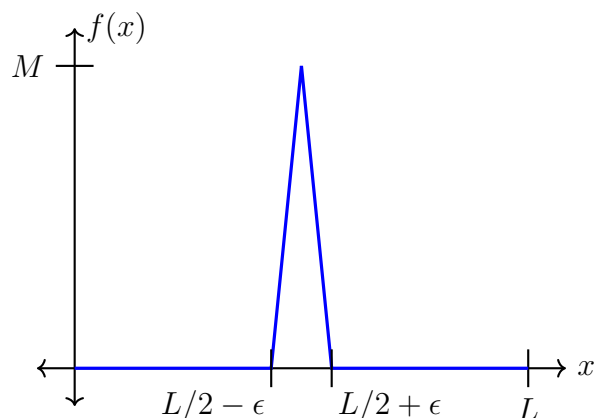
It should be the case that $\|f\|_\infty = M$ but $\|f\|_{L^1}$ and $\|f\|_{L^2}$ are "small" since the area under the graph of $f = |f|$ is quite small.

24.5 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* $M = 1/\sqrt{\epsilon}$.]

Day 25: Wednesday, March 5.

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

Pages 222–223 discuss piecewise continuous functions.

Nonetheless, we can still obtain $\|\cdot\|_\infty$ estimates on differences of solutions to the heat boundary value problems. The maximum principle and Problem 22.9 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.

$$w(x,0) = f_1(x) - f_2(x), \ 0 \le x \le L$$
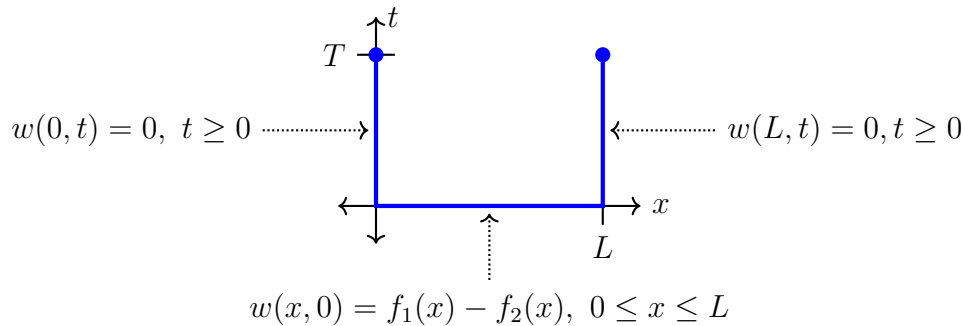
On the vertical sides of this boundary ($x = 0$ and $x = L$), $w = 0$, so those sides contribute nothing really interesting to the extreme values of w. Suppose $-\epsilon \le f_1(x) \le f_2(x) \le \epsilon$ for all $x \in [0, L]$. Then the minimum of w on the parabolic boundary is at least $-\epsilon$ 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.

25.1 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 $t \ge 0$.*

25.2 Problem (!). Prove the L^2-estimate in part (ii) of Theorem 25.1.

25.3 Problem ($\star$). Let $f \in \mathcal{C}(\mathbb{R})$ be bounded with $|f(y)| < \epsilon$ for all $y \in \mathbb{R}$. Suppose that u is continuous on $\{(x, t) \in \mathbb{R}^2 \mid x \in \mathbb{R}, \ t \ge 0\}$ and solves

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

Develop an estimate for $|u(x, t)|$ that is, ideally, independent of x and t but involves ϵ. [Hint: *Problem 21.8, Theorem 21.9.*]

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.

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.
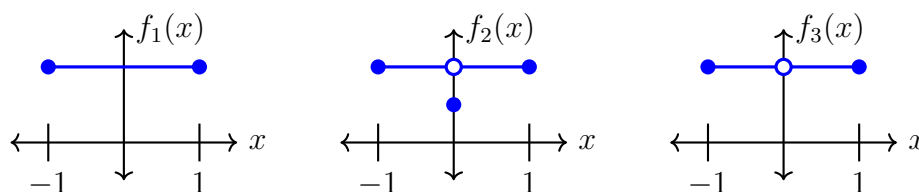
Any "good" notion of area should take into account that "the area under a point is 0." That f_2 disagrees with f_1 just at $x = 0$, and that f_3 is not even defined at $x = 0$ should not matter.

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.
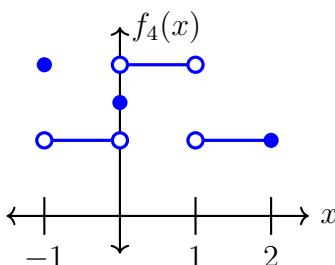
The only difference here is that sometimes $\lim_{s \to x^+} f_4(s) \neq \lim_{s \to x^-} f_4(s)$ for some $x \in (-1, 2)$.

All of these functions are what we want to call "piecewise continuous."

25.4 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:*

$$\lim_{s \to a^+} f(s), \qquad \lim_{s \to x^\pm} f(s), \ a < x < b, \qquad and \qquad \lim_{s \to b^-} f(s).$$

Then we say that f is **PIECEWISE CONTINUOUS** *on $[a, b]$. The set of all piecewise continuous 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.

25.5 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

$$\int_c^d f = \int_a^b f \chi_{[c,d]}, \qquad \chi_{[c,d]}(x) := \begin{cases} 1, & c \le x \le d \\ 0, & x < c \text{ or } c > d. \end{cases}$$

25.6 Theorem. *Every part of Theorem 2.1 remains true for $I = [a, b]$ when $\mathcal{C}([a, b])$ is replaced by $\mathcal{C}_{\mathrm{pw}}([a, b])$.*

25.7 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 $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ if $f \in \mathcal{C}_{\mathrm{pw}}([a, b])$ 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).

25.8 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 $[a, b] \subseteq \mathbb{R}$. (In particular, this holds if 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 $\mathcal{C}_{\mathrm{pw}}(\mathbb{R})$.*

25.9 Problem (+). Prove that if there is $d > 0$ such that the set $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ satisfies $|x_k - x_j| \ge d$ for all $j \ne k$, then $\{x_k\}_{k=1}^{\infty} \cap [a, b]$ is finite for any $[a, b] \subseteq \mathbb{R}$. (However, the converse is not true—consider a set whose elements are $2 - 1/2, 2 + 1/2, 3 - 1/3, 3 + 1/3, 4 - 1/4, 4 + 1/4, \ldots$. There is no "minimum distance" between points in this set, but since any $[a, b] \subseteq \mathbb{R}$ contains only finitely many integers, it will also contain only finitely many points of this set.)

As before, a piecewise continuous function on $\mathbb{R}$ does not have to be defined "on" $\mathbb{R}$, though certainly we have $\mathcal{C}(\mathbb{R}) \subseteq \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$. 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 $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ 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.

Day 26: Friday, March 7.

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

We now define improper integrals of piecewise continuous functions on $\mathbb{R}$ using the integral for piecewise continuous functions on subintervals of $\mathbb{R}$.

> **26.1 Definition. (i)** *Let $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ and suppose that the limits*
>
> $$\int_{-\infty}^0 f := \lim_{a \to -\infty} \int_a^0 f \quad \text{and} \quad \int_0^\infty f := \lim_{b \to \infty} \int_0^b f$$
>
> *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 $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ 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.

> **26.2 Theorem (Comparison test).** *Let $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ 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*
>
> $$\left| \int_{-\infty}^\infty f \right| \le \int_{-\infty}^\infty |f| \le \int_{-\infty}^\infty |g|.$$

> **26.3 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 $\mathbb{R} \setminus \{x_k, y_k\}_{k=1}^\infty$.) [Hint: *use results from Problem 19.13.*]

26.4 Problem (+). While L^1 is a vector space, it is not an "algebra" (which is a vector space in which one can multiply the vectors nicely): it is not true that if f, $g \in L^1$, then $fg \in L^1$. Define $f \colon \mathbb{R} \to \mathbb{C}$ by $f(x) = 0$ for $x < 1$, and, for $x \geq 1$, if $n \geq 1$ is an integer such that $n \leq x < n + 1$, then $f(x) = n$ for $n \leq x < n + 1/n^3$, and $f(x) = 0$ for $n + 1/n^3 \leq x < n + 1$. First graph f over $[-3, 3]$ to get a sense of what is going on. Then show that $f \in L^1$ but $f^2 \notin L^1$. [Hint: *for what values of $p \in \mathbb{R}$ does the series $\sum_{n=1}^\infty n^p$ converge?*]

26.5 Problem (+). Prove that $\|\cdot\|_{L^1}$ satisfies three of the four properties that one usually demands of a "norm" on a vector space:

$$\|f\|_{L^1} \geq 0, \qquad \|\alpha f\|_{L^1} = |\alpha| \, \|f\|_{L^1}, \qquad \text{and} \qquad \|f + g\|_{L^1} \leq \|f\|_{L^1} + \|g\|_{L^1}$$

for all f, $g \in L^1$ and $\alpha \in \mathbb{R}$. However, find an example of $f \in L^1$ such that $f \neq 0$ 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.

26.6 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!).

26.7 Example. Fix $r > 0$ and let

$$f(x) := \begin{cases} 1, & |x| \leq r \\ 0, & |x| > r. \end{cases} \tag{26.1}$$

It is a worthwhile exercise in the definition of the integral to show $f \in L^1$; we leave this as a problem. (The function f is really the **INDICATOR FUNCTION** of the interval $[-r, r]$,

and later we will write $\chi(x,r) := f(x)$.) It then follows from Problem 19.13 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 $k \neq 0$ we have

$$\widehat{f}(k) = \frac{1}{\sqrt{2\pi}} \frac{1}{-ik} \int_{-r}^{r} -ike^{-ikx} \, dx = -\frac{1}{ik\sqrt{2\pi}} \int_{-r}^{r} \partial_x[e^{-ikx}] \, dx = \frac{e^{ikr} - e^{-ikr}}{ik\sqrt{2\pi}}$$

$$= \frac{2r}{kr\sqrt{2\pi}} \frac{e^{ikr} - e^{-ikr}}{2i} = \frac{2r}{\sqrt{2\pi}} \frac{\sin(kr)}{kr}.$$

Now put

$$\mathrm{sinc}(X) := \begin{cases} \sin(X)/X, & X \neq 0 \\ 1, & X = 0. \end{cases}$$

We have shown

$$\widehat{f}(k) = \frac{2r}{\sqrt{2\pi}} \mathrm{sinc}(kr),$$

and this is a rather nicer function than f, since we recall from calculus that sinc is infinitely differentiable.

26.8 Problem (!). Do that worthwhile exercise and show that $f \in L^1$, where f is defined in (26.1).

26.9 Problem ($\star$). 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 if f is real-valued, then $\overline{\widehat{f}(k)} = \widehat{f}(-k)$. [Hint: *use the fact that* $\overline{\int_a^b f} = \int_a^b \overline{f}$. *This is true for improper integrals, too.*]

(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 26.7?

The nice properties of $\widehat{f}$ as computed in Example 26.7 carry over well to Fourier transforms in general.

26.10 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 $\widehat{f}$! ∎

26.11 Problem (+). **(i)** A natural way to try to prove continuity of $\widehat{f}$ is to rewrite

$$|\widehat{f}(k_2) - \widehat{f}(k_1)| = \frac{|k_2 - k_1|}{\sqrt{2\pi}} \left| \int_{-\infty}^\infty \left(\int_0^1 e^{-ix[(1-t)k_1 + tk_2]} \, dt \right) x f(x) \, dx \right|. \tag{26.2}$$

First show that (26.2) is true. The utility of (26.2) 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 (26.2) 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 $\pm\infty$.*

We formalize the most useful property of the Fourier transform.

26.12 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 19.13 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

$$\int_{-R}^{R} f'(x)e^{-ikx} \, dx = f(x)e^{-ikx}\big|_{x=-R}^{x=R} - \int_{-R}^{R} f(x)(-ik)e^{-ikx} \, dx$$

$$= f(R)e^{-ikR} - f(-R)e^{ikR} + ik\int_{-R}^{R} f(x)e^{-ikx} \, dx.$$

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. ∎

26.13 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.*]

26.14 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:*

$$\widehat{ODE} = algebraic, \qquad \widehat{PDE} = ODE.$$

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 26.12 presume decay of the solution at (spatial) infinity.

As we previously observed in our study of the heat equation, we can iterate Theorem 26.12 and take the transform of any derivative.

26.15 Corollary. *Let $f \in C^r(\mathbb{R})$ for some $r \geq 1$ and suppose $f^{(j)} \in L^1$ for $j = 0, \dots, r$. 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

$$\widehat{f}(k) = \frac{\widehat{f^{(r)}}(k}{(ik)^r} \quad \text{and so} \quad |\widehat{f}(k)| \leq \frac{\|f^{(r)}\|_{L^1}}{\sqrt{2\pi}} \left(\frac{1}{|k|^r} \right). \tag{26.3}$$

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!

Day 27: Monday, March 17.

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

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 $y' = ay + f(x)$.

Next we consider inverting the Fourier transform. Untheorem 19.15 gives us the (unfortunately false) hope that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k) e^{ikx} \, dk \tag{27.1}$$

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 26.7, 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.

27.1 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.

27.2 Problem (!). Explain why by considering

$$f(x) := \begin{cases} 1, & x = 0 \\ 0, & x \neq 0. \end{cases}$$

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

$$\frac{f(x^+) + f(x^-)}{2} = \frac{1}{\sqrt{2\pi}} \lim_{R \to \infty} \int_{-R}^{R} \widehat{f}(k)e^{ikx} \, dk. \tag{27.2}$$

27.3 Problem (!). Suppose that $f \in L^1 \cap \mathcal{C}(\mathbb{R})$ with $\widehat{f} \in L^1$ and (27.2) is true for some $x \in \mathbb{R}$. Why does that imply about (27.1)?

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.
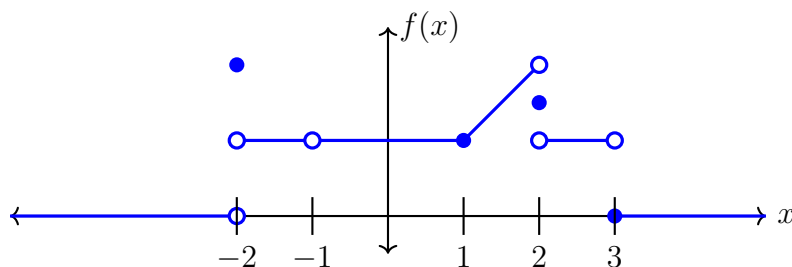
The function above is 0 on $(-\infty, -2)$ and $[3, \infty)$, and slightly more complicated on $[2, 3]$. It has some jump discontinuities, and it is not defined at -1, and it is defined and continuous but not differentiable at 1, but it is definitely the case that

$$f(x^{\pm}) := \lim_{s \to x^{\pm}} f(s)$$

exist for all $x \in \mathbb{R}$ and are equal except for $x = \pm 2, 3$. Likewise, the limits of the **LEFT AND RIGHT DIFFERENCE QUOTIENTS**

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x^+)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(x+h) - f(x^-)}{h} \tag{27.3}$$

exist at all $x \in \mathbb{R}$ and are equal except for $x = \pm 2, 1, 3$. (In fact, these limits are equal at $x = -1$ even though f is not defined there!)

Suppose that we take the unusual step of defining $f'(x)$ to be either of the difference quotients in (27.3) 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 $\mathbb{R}$ in the sense of Definition 25.8!

27.4 Definition. *Let $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ and suppose that*

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x^+)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(x+h) - f(x^-)}{h}$$

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 $\mathcal{C}_{\mathrm{pw}}(\mathbb{R})$, 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 $f' \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$. Then we say that f is **PIECEWISE CONTINUOUSLY DIFFERENTIABLE** on $\mathbb{R}$, and we denote the set of such functions by $\mathcal{C}_{\mathrm{pw}}^1(\mathbb{R})$.*

As with $\mathcal{C}_{\mathrm{pw}}(\mathbb{R})$, a function in $\mathcal{C}_{\mathrm{pw}}^1(\mathbb{R})$ need not be defined on all of $\mathbb{R}$. Additionally, and unfortunately, a function in $\mathcal{C}_{\mathrm{pw}}^1(\mathbb{R})$ 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.

27.5 Theorem (Fourier inversion). *Let $f \in L^1 \cap \mathcal{C}^1_{\mathrm{pw}}(\mathbb{R})$. 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 26.10, 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 (27.1) applies, then we have hope of checking it rigorously. (Anyway, even Theorem 27.5 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 (27.1) of Fourier inversion.

27.6 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" $\widecheck{g}(x) = \widehat{g}(-x)$.

27.7 Problem (!). Prove that if $f \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ with $\widehat{f} \in L^1$, then

$$f(x) = \widehat{\widehat{f}}(-x).$$

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 $\mathrm{Re}(ik - a) = -a \neq 0$, so $ik - a \neq 0$ for all $k \in \mathbb{C}$. Then we can solve for $\widehat{y}(k)$:

$$\widehat{y}(k) = \frac{\widehat{f}(k)}{ik - a}.$$

27.8 Problem (!). If $a = 0$, then the division strategy above fails. Show that if $y \in L^1 \cap C^1(\mathbb{R})$ with $y' \in L^1$ 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

$$\int_{-\infty}^{\infty} \widehat{f}(k)\widehat{g}(k)e^{ikx} \, dk = \sqrt{2\pi}\,\overset{\smile}{\widehat{f}\widehat{g}}(x).$$

Maybe there is something more profound that we can say about $\overset{\smile}{\widehat{f}\widehat{g}}(x)$. How does the Fourier transform interact with products?

27.9 Problem (!). Explain why you should expect $\widehat{fg}(k) \neq \widehat{f}(k)\widehat{g}(k)$ in general. [Hint: *integration is "additive" in the sense that* $\int_a^b (f + g) = \int_a^b f + \int_a^b g$, *but do we know anything about how* $\int_a^b fg$ *and* $\left(\int_a^b f\right)\left(\int_a^b g\right)$ *are related?*]

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
$$\widehat{g' - ag}(k) = 1. \tag{27.4}$$

Now it looks like we are assuming that $g \in \mathcal{C}^1_{\mathrm{pw}}(\mathbb{R})$ with $g' \in L^1$, but, remember, we are only fooling around. The identity (27.4) 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
$$\int_{-\infty}^{\infty} x(g'(x) - ag(x))e^{-ikx}\ dx = 0. \tag{27.5}$$

Put $q(x) := x(g'(x) - ag(x))$. Then (27.5) 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 $\mathsf{E}_a^+ \in \mathcal{C}^1_{\mathrm{pw}}(\mathbb{R})$ 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.

27.10 Problem ($\star$). Using only the definitions of L^1 and the Fourier transform, show that $\mathsf{E}_a^+ \in L^1$ and

$$\widehat{\mathsf{E}_a^+}(k) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{a - ik} \right).$$

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?

Day 28: Wednesday, March 19. ▬▬▬▬▬▬▬▬▬▬▬

> **Material from *Basic Partial Differential Equations* by Bleecker & Csordas**
>
> Pages 681–682 review Fubini's theorem. Pages 435–436 discuss convolution.

We pause from considering $y' = ay + f(x)$ and let f, $g \in L^1$. We want to understand the product $\widehat{f}(k)\widehat{g}(k)$. When all else fails, go back to the definition:

$$2\pi \widehat{f}(k)\widehat{g}(k) = \left(\int_{-\infty}^{\infty} f(x)e^{-ikx}\ dx \right)\left(\int_{-\infty}^{\infty} g(y)e^{-iky}\ dy \right).$$

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:

$$2\pi \widehat{f}(k)\widehat{g}(k) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x)e^{-ikx}\ dx \right) g(y)e^{-iky}\ dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)g(y)e^{-ik(x+y)}\ dx\ dy.$$

Can we 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:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s - y)g(y)e^{-iks}\ ds\ dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s - y)g(y)e^{-iks}\ dy\ ds.$$

28.1 Theorem (Fubini, highly specialized). *This works. More precisely let f, $g \in L^1$ and let $h \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ with h bounded (there is $M > 0$ such that $|h(s)| \leq M$ 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 $-\infty \leq a \leq b \leq \infty$ and $-\infty \leq c \leq d \leq \infty$.

We take the version of Fubini above for granted and will not prove it. Thus

$$2\pi \widehat{f}(k)\widehat{g}(k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s-y)g(y)e^{-iks} \ dy \ ds = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(s-y)g(y) \ dy \right) e^{-iks} \ ds.$$
(28.1)

The inner integral on the right has a special name.

28.2 Theorem. *Let $f \in L^1$ and $g \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ 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 $x \in \mathbb{R}$.

28.3 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. Now we can interpret (28.1) via convolution. Changing variables from s back to x, we have

$$2\pi \widehat{f}(k)\widehat{g}(k) = \int_{-\infty}^{\infty} (f * g)(x)e^{-ikx} \ dx,$$

but that seems to be presuming $f * g \in L^1$, no?

28.4 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*

$$\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}.$$

28.5 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:

$$\int_{-R}^{R} |(f * g)(x)| \ dx \leq \int_{-\infty}^{\infty} \int_{-R}^{R} |f(x-y)||g(y)| \ dx \ dy \leq \|f\|_{L^1} \|g\|_{L^1}.$$

Then apply part (iii) of Problem 19.13.

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.

28.6 Theorem. *Let f, $g \in L^1$ with at least one of these functions bounded. Then*

$$\widehat{f}(k)\widehat{g}(k) = \frac{\widehat{f * g}(k)}{\sqrt{2\pi}}.$$

We digressed from the story of solving $y' = ay + f(x)$ via the Fourier transform to introduce the important transform of E_a^+ and the even more important operation of convolution. Now we redo this ODE from a Fourier point of view more quickly, so we can see how everything works all together. Once more, assume $y \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ and $f' \in L^1$. Then our prior work gives

$$\widehat{y}(k) = -\sqrt{2\pi}\widehat{f}(k)\widehat{\mathsf{E}_a^+}(k),$$

and so convolution gives

$$\widehat{y}(k) = -\sqrt{2\pi} \left(\frac{1}{\sqrt{2\pi}} \right) \widehat{f * \mathsf{E}_a^+}(k) = -\widehat{f * \mathsf{E}_a^+}(k).$$

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

$$y(x) = -e^{ax} \int_{x}^{\infty} e^{-as} f(s) \, ds. \tag{28.2}$$

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 5.1 told us that the only solution to $y' = ay + f(x)$ with $\lim_{x \to \infty} y(x) = 0$ has exactly this form (28.2). *The Fourier transform will not necessarily yield all solutions to a differential equation because legitimate application of the transform presumes decay at spatial infinity.*

28.7 Problem ($\star$). 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* $\widehat{\mathsf{E}_{-a}^+}(k)$?]

28.8 Problem ($\star$). Let $\mathsf{E}(x) = e^{-|x|}$. Show that

$$\widehat{\mathsf{E}}(k) = \frac{1}{\sqrt{2\pi}} \frac{2}{k^2 + 1}. \tag{28.3}$$

Try to do this *not* by using the definition of the Fourier transform but by recognizing E as the sum of $\mathsf{E}_a^{\pm}$ for some choice of a.

28.9 Problem ($+$). Suppose throughout that $y'' - y = f(x)$.

(i) Use (28.3) to show that the candidate solution for this ODE that one obtains from the Fourier transform is

$$y(x) = -\frac{1}{2} \int_{-\infty}^{\infty} e^{-|x-s|} f(s) \, ds.$$

(ii) Show that the solution above has the form predicted by variation of parameters (Theorem 12.4). In particular, what are the values of $y(0) = y_0$ and $y'(0) = y_1$ that emerge from this analysis? [Hint: *first work out what (12.3) looks like in this case, assuming arbitrary values for* $y(0) = y_0$ *and* $y'(0) = y_1$; *at some point* $\sinh(x)$ *and* $\cosh(x)$ *should show up—recall Problem 6.9. Then rewrite the integral above as a sum of integrals of the form* $\int_{-\infty}^{x}$ *and* $\int_{x}^{\infty}$, *where the exponentials in the integrand have no absolute values. Rewrite those integrals to expose factors of* $e^{\pm x}$. *Do algebra until* $\sinh(x)$ *and* $\cosh(x)$ *appear.*]

(iii) Using the Fourier transform to find a candidate solution here presumes that y and y' vanish at $\pm\infty$ (recall Problem 26.13). Suppose that y solves $y'' - y = f(x)$ with $\lim_{x \to \pm\infty} y(x) = \lim_{x \to \pm\infty} y'(x) = 0$ and with $f \in L^1$. Argue that

$$y(0) + y'(0) = -\int_0^{\infty} e^{-s} f(s) \, ds \quad \text{and} \quad y(0) - y'(0) = -\int_{-\infty}^{0} e^s f(s) \, ds.$$

How does this compare to the values for $y(0)$ and $y'(0)$ that you found in the previous part? Note that the integrals exist by the comparison test, since $f \in L^1$ and the exponentials are decaying on the given intervals. [Hint: *from variation of parameters, express* $y(x) = A(x)e^x + B(x)e^{-x}$ *for some functions A and B. Argue that* $A_\infty := \lim_{x \to \infty} A(x)$ *and* $B_\infty := \lim_{x \to -\infty} B(x)$ *both exist as finite real numbers, and so if the limits above on y at* $\pm\infty$ *are 0, then* $A_\infty = B_\infty = 0$.]

Day 29: Friday, March 21. ▬▬▬▬

Please continue reading and working through the material from Wednesday.

Day 30: Monday, March 24. ▬▬▬▬

We are going to study an aspect of end behavior for the heat equation mostly as a way of getting further practice with properties of the Fourier transform and convolution. We need one more function space.

30.1 Definition. *A function* $f \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ *is* **BOUNDED** *if there exists* $C > 0$ *such that* $|f(x)| \leq C$ *for all x in the domain of f. Denote the set of all bounded piecewise continuous functions on* $\mathbb{R}$ *by* L^∞.

30.2 Problem (!). Prove that if $f \in L^1$ and $g \in L^\infty$, then $fg \in L^1$. This is possibly one of the most common uses of the comparison test for improper integrals in analysis. (Use the idea in Problem 26.3 to make sense of the domain of fg.)

30.3 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 $|f(x)| \leq C$ for all $x \in \mathbb{R}$. [Hint: *Fourier inversion.*]

30.4 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 $\alpha \in \mathbb{C}$. (Refer to Problem 26.3 for how to make sense of the domain of $f + g$.)

30.5 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 might 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.

Here is what we are going to do with the heat equation. Problem 22.1 gave conditions for a solution u to the infinite rod heat equation to satisfy

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

for all x (actually, uniformly in x). Now we establish the limit

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

for all t (though not uniformly in t). Here is why we should expect this: if we can interchange the integral and the limit (BIG IF), then

$$\lim_{x \to \infty} u(x,t) = \lim_{x \to \infty} \int_{-\infty}^{\infty} H(x-y,t) f(y) \ dy = \lim_{x \to \infty} \int_{-\infty}^{\infty} \frac{e^{-(x-y)^2/4t}}{\sqrt{4\pi t}} f(y) \ dy$$

$$= \int_{x \to \infty} \lim_{x \to \infty} \frac{e^{-(x-y)^2/4t}}{\sqrt{4\pi t}} f(y) \ dy = 0,$$

since

$$\lim_{x \to \infty} \frac{e^{-(x-y)^2/4t}}{\sqrt{4\pi t}} f(y) = 0$$

for all fixed $y \in \mathbb{R}$ and $t > 0$.

The first step is to view the solution to the heat equation as convolution. Specifically, let $f \in \mathcal{C}(\mathbb{R}) \cap L^\infty$. Then

$$u(x,t) := (H(\cdot,t) * f)(x)$$

solves the heat equation for $t > 0$. Here, of course, H is the heat kernel. This is all from Theorem 21.9. The notation above means that we are taking the convolution in the first variable of H and thinking of time t in that convolution just as a parameter.

30.6 Problem (!). Let $t > 0$. Check that $H(\cdot,t)$ satisfies the following:

$$H(\cdot,t) \in L^1 \cap L^\infty \quad \text{and} \quad \widehat{H}(\cdot,t) \in L^1.$$

Here $\widehat{H}(\cdot,t)$ is the Fourier transform in the first variable of H. [Hint: *use* (21.3) *to relate* $\widehat{H}(k,t)$ *and* $H(k,t)$. *Cite results from Problem 19.12 for integrability.*]

Going forward, we will ignore t and instead think about a function m satisfying the following. Really, $m = H(\cdot,t)$ for t fixed; this just cuts down on the writing.

$(\widehat{\textbf{M1}})$ $m \in L^1$ (Problem 30.6 for $m = H(\cdot,t)$).

$(\widehat{\textbf{M2}})$ $\widehat{m} \in L^1$ (Problem 30.6 for $m = H(\cdot,t)$).

And we will assume the following on f, which is not quite the same as in Theorem 21.9.

(F̂1) $f \in L^\infty$.

(F̂2) $m * f \in \mathcal{C}^1(\mathbb{R})$ (Theorem 21.9 for $m * f = H(\cdot, t) * f$).

By (M̂1), (F̂1), and Theorem 28.4, we have $g := m * f \in L^1$, and then by Theorem 28.6, we have $\widehat{m * f} = \widehat{m}\widehat{f}/\sqrt{2\pi}$. Since $\widehat{m} \in L^1$ by (M̂2) and $\widehat{f} \in L^\infty$ by (F̂1), we have $\widehat{m}\widehat{f} \in L^1$ by Problem 30.2. That is, $\widehat{g} \in L^1$. Since $g \in \mathcal{C}^1(\mathbb{R})$ by (F̂2), we may use the ideal version of Fourier inversion from Problem 27.7 to get

$$g(x) = \widehat{\widehat{g}}(-x).$$

And since $\widehat{g} \in L^1$, we have

$$\lim_{k \to \pm\infty} \widehat{\widehat{g}}(k) = 0$$

by the Riemann–Lebesgue lemma (Theorem 26.10). This gives

$$\lim_{x \to \pm\infty} g(x) = \lim_{x \to \pm\infty} \widehat{\widehat{g}}(-x) = 0.$$

Abstractly, painfully we have proved the following.

30.7 Theorem. *Let $f \in \mathcal{C}(\mathbb{R}) \cap L^\infty$ and let*

$$u(x, t) = \int_{-\infty}^{\infty} H(x - y, t)f(y) \, dy, \quad x \in \mathbb{R}, \ t > 0,$$

where H is the heat kernel. Then

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

It happens that we have omitted an argument in our study of convolutions: to have $f * g \in L^1$, we need $f * g \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$. 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.

30.8 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

$$(f * g)(x + h) - (f * g)(x) = \int_{-\infty}^{\infty} f(x + h - y)g(y) \, dy - \int_{-\infty}^{\infty} f(x - y)g(y) \, dy$$

$$= \int_{-\infty}^{\infty} \big[f(x + h - y) - f(x - y)\big]g(y) \, dy.$$

Then we estimate

$$|(f*g)(x+h)-(f*g)(x)| \leq \int_{-\infty}^{\infty} \left|[f(x+h-y)-f(x-y)]g(y)\right| dy \leq C \int_{-\infty}^{\infty} \left|f(x-y+h)-f(x-y)\right| dy,$$
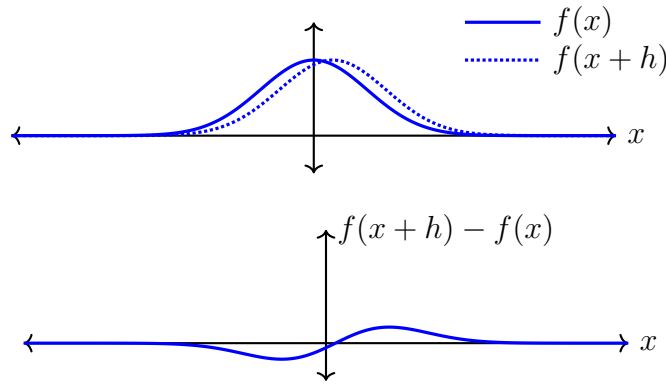
where $C > 0$ is such that $|g(y)| \leq C$ for all $y \in \mathbb{R}$. Substitute $s = x - y$ above to find

$$|(f*g)(x+h)-(f*g)(x)| \leq C \int_{-\infty}^{\infty} |f(s+h)-f(s)| \, dy = C \left\|S^h f - f\right\|_{L^1}, \qquad (S^h f)(s) := f(s+h).$$

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. ■

30.9 Problem (+). Here is how to make the last paragraph of the preceding proof precise. (This requires awareness of uniform continuity.)

    Assume $f \in L^1$ and, for simplicity, assume $f \in C(\mathbb{R})$, not just $C_{\mathrm{pw}}(\mathbb{R})$. Let $\epsilon > 0$. We want to find $\delta > 0$ such that if $0 < |h| < \delta$, then

$$\left\|S^h f - f\right\|_{L^1} < \epsilon.$$

For $M > 0$, rewrite

$$\left\|S^h f - f\right\|_{L^1} = \mathcal{I}_-(M, h) + \mathcal{I}(M, h) + \mathcal{I}_+(M, h),$$

where

$$\mathcal{I}_-(M, h) = \int_{-\infty}^{-M} |S^h f - f|, \quad \mathcal{I}(M, h) = \int_{-M}^{M} |S^h f - f|, \quad \text{and} \quad \mathcal{I}_+(M, h) = \int_{M}^{\infty} |S^h f - f|.$$

(i) Use the triangle inequality to estimate

$$\mathcal{I}_-(M,h) \le \int_{-\infty}^{h-M} |f| + \int_{-\infty}^{-M} |f|.$$

Since $f \in L^1$, choose M so large that $\int_{-\infty}^{1-M} |f|$ is small (say, a fraction of ϵ). Then with $|h| < 1$, argue that $\int_{-\infty}^{h-M} |f|$ is also small.

(ii) Do a similar argument to show that $\mathcal{I}_+(M,h)$ is small.

(iii) Use uniform continuity to show that for some $0 < \delta < 1$, if $0 < |h| < \delta$, then $\mathcal{I}(M,h)$ is also small.

Day 31: Wednesday, March 26.

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

$$\begin{cases} u_t = u_{xx}, & (x,t) \in \mathcal{D} \\ u(x,0) = f(x), \end{cases} \qquad \mathcal{D} := \left\{ (x,t) \in \mathbb{R}^2 \mid x \in \mathbb{R}, \ t > 0 \right\}.$$

Since $H(\cdot,t) \in L^1$ (by Problem 19.12), we can assume either $f \in L^1$ or $f \in L^\infty$ 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.

31.1 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 $(f * g)' = f' * g$.

If it is true that $f * g \in \mathcal{C}^1(\mathbb{R})$ for $f \in L^1 \cap \mathcal{C}^1(\mathbb{R})$ and $g \in L^\infty$, then

31.2 Problem ($\star$). For $r > 0$, let

$$\chi(x,r) := \begin{cases} 1, & |x| \leq r \\ 0, & |x| > r. \end{cases} \tag{31.1}$$

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

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

whose solution, by Example 6.6, is

$$y(t) = y_0 \cos(kt) + y_1 t \operatorname{sinc}(kt), \qquad \operatorname{sinc}(\tau) := \begin{cases} \sin(\tau)/\tau, & \tau \neq 0 \\ 1, & \tau = 0. \end{cases}$$

With that ODE result, we have

$$\widehat{u}(k,t) = \widehat{f}(k) \cos(kt) + \widehat{g}(k) t \operatorname{sinc}(kt),$$

and so Fourier inversion suggests the candidate solution

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{u}(k,t) e^{ikx} \, dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(k) \cos(kt) e^{ikx} \, dk + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k) t \operatorname{sinc}(kt) e^{ikx} \, dk.$$

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.

For the second term, put

$$\chi(x,t) = \begin{cases} |x| \leq t \\ 0, & |x| > t \end{cases}$$

as in (31.1). We calculated in Example 26.7 that its Fourier transform is

$$\widehat{\chi}(k,t) = \frac{2t}{\sqrt{2\pi}} \operatorname{sinc}(kt).$$

Then we rewrite the second term as

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k) t \operatorname{sinc}(kt) e^{ikx} \, dk = \frac{1}{2} \left[\int_{-\infty}^{\infty} \widehat{g}(k) \left(\frac{2t \operatorname{sinc}(kt)}{\sqrt{2\pi}} \right) e^{ikx} \, dk \right] = \frac{1}{2} \int_{-\infty}^{\infty} \widehat{g}(k) \widehat{\chi}(k,t) e^{ikx} \, dk$$

$$= \frac{\sqrt{2\pi}}{2} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k) \widehat{\chi}(k,t) e^{ikx} \, dk \right) = \frac{\sqrt{2\pi}}{2} \widetilde{g\widehat{\chi}(\cdot,t)}(x).$$

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

$$\phi * \psi = \overbrace{\widehat{\phi} * \widehat{\psi}} = \sqrt{2\pi}\widetilde{\widehat{\phi}\widehat{\psi}}.$$

Thus the second term in the candidate solution should be

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k)t \operatorname{sinc}(kt)e^{ikx} \, dk = \frac{(g * \chi(\cdot, t))(x)}{2}.$$

31.3 Problem (!). Use the result of Problem 31.2 to show that this second candidate term is what we expect the second term to be from D'Alembert's formula.

31.4 Problem (!). Use the Fourier transform to find a formal solution to

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

Your solution should largely resemble your solution to Problem 20.3. Simplify your answer until you arrive at the solution form from Theorem 9.6. [Hint: *the new feature here, compared to Problem 20.3, is the presence of g. It may be easier to start with $\int_0^t g(x - t + \tau, \tau) \, d\tau$, rewrite*

$$g(x - t + \tau, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g}(k, \tau)e^{ik(x-t+\tau)} \, dk,$$

and then interchange the order of integration.]

31.5 Problem (⋆). Use the Fourier transform to find a formal solution to

$$\begin{cases} u_{tt} = u_{xx} + h(x, t), & x, \ t \in \mathbb{R} \\ u(x, 0) = f(x), & x \in \mathbb{R} \\ u_t(x, 0) = g(x), & x \in \mathbb{R}. \end{cases}$$

Your work should largely resemble the steps above for the wave equation (so you should use this as an opportunity to check your understanding of those steps). Simplify your answer until you arrive at the solution form (12.6).

31.6 Problem (+). **(i)** Use the Fourier transform to find a formal solution to

$$\begin{cases} u_t = u_{xx} + g(x, t), & -\infty < x < \infty, \ t > 0 \\ u(x, 0) = f(x), & -\infty < x < \infty. \end{cases} \tag{31.2}$$

Get down to a formula that resembles (20.3) but also takes into account g; use this as an opportunity to check your understanding of the development of (20.3) back when $g = 0$.

(ii) In the spirit of the "propagator" operators for the transport and wave equations (recall the discussion following (12.1)), develop an operator $\mathcal{P}$ such that a candidate solution for

(31.2) is

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

Explain clearly how you constructed $\mathcal{P}$. You do not have to prove that your candidate solution actually solves the driven heat equation; indeed, this will be delicate due to a certain "singularity" in $\mathcal{P}(t-\tau, x)g(\cdot, \tau)$ that would make differentiation under the integral unpleasant.

Day 32: Friday, March 28.

You took Exam 2.

Day 33: Monday, March 31.

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.

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?

The goal is to prove that

$$\partial_s \left[\int_{-\infty}^{\infty} f(s,\tau)\, d\tau \right] = \int_{-\infty}^{\infty} f_s(s,\tau)\, d\tau,$$

assuming enough reasonable hypotheses on f. (This probably will not be true for all f, but we do not want the hypotheses to be so restrictive that they would not match, say, the convolution structure for the heat equation.) In particular, we need to assume enough so that all integrals involved converge.

The limit definition of the derivative means that we really want to show

$$\lim_{h \to 0} \int_{-\infty}^{\infty} \frac{f(s+h,\tau) - f(s,\tau)}{h}\, d\tau = \int_{-\infty}^{\infty} \lim_{h \to 0} \frac{f(s+h,\tau) - f(s,\tau)}{h}\, d\tau.$$

This motivates the following set-up and notation.

Let $I \subseteq \mathbb{R}$, and we put

$$\mathcal{D} := \big\{ (h,\tau) \in \mathbb{R}^2 \mid h \in I,\ \tau \in \mathbb{R} \big\}.$$

Let $q\colon \mathcal{D} \to \mathbb{C}$ and suppose that either $h_0 \in I$ or $h_0 \in \mathbb{R}$ is an endpoint of I. We want to show

$$\lim_{h \to h_0} \int_{-\infty}^{\infty} q(h, \tau) \, d\tau = \int_{-\infty}^{\infty} \lim_{h \to h_0} q(h, \tau) \, d\tau. \tag{33.1}$$

If h_0 is an endpoint, then we allow left or right limits; we also allow $h_0 = \pm\infty$. 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 (33.1) is true?

33.1 Problem ($\star$). No. Let

$$q(h, \tau) := \begin{cases} 1, & \dfrac{1}{|h|} \le \tau < \dfrac{2}{|h|}, \ h \ne 0, \ \tau \in \mathbb{R} \\[2mm] 0, & \text{otherwise.} \end{cases}$$

(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 $h \to 0$?*]

33.2 Problem ($\star$). And no. Let

$$q(h, \tau) := \begin{cases} 1, & \dfrac{1}{|h|} \le \tau < \dfrac{1}{|h|} + 1, \ h \ne 0, \ \tau \in \mathbb{R} \\[2mm] 0, & \text{otherwise.} \end{cases}$$

(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 $|q(h, \tau)| \leq |M(\tau)|$ for all $(h, \tau) \in \mathbb{R}^2$. [Hint: *again, given $\tau > 0$, compute $q(1/\tau, \tau)$ and conclude $|M(\tau)| \geq 1$ 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 $|q(h, \tau)| \leq |M(\tau)|$ 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.

33.3 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 $h_0 \in I$ or let $h_0 \in \mathbb{R}$ be an endpoint of I, and let $q \colon \mathcal{D} \to \mathbb{C}$ be a function with the following properties.*

(i) *$q(h, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ 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) *$Q \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$.*

(iv) *There is $M \in L^1$ such that $|q(h, \tau)| \leq |M(\tau)|$ for all $(h, \tau) \in \mathcal{D}$.*

Then
$$\lim_{h \to h_0} \int_{-\infty}^{\infty} q(h, \tau) \, d\tau = \int_{-\infty}^{\infty} \lim_{h \to h_0} q(h, \tau) \, d\tau = \int_{-\infty}^{\infty} Q(\tau) \, d\tau.$$

If h_0 is an endpoint of I, then the result above is still true with $\lim_{h \to h_0}$ replaced by $\lim_{h \to h_0^{\pm}}$; the result is also still true for $h_0 = \pm\infty$.

33.4 Problem (!). Assume that the hypotheses of the dominated convergence theorem are met with $h_0 \in I$ and that q is continuous on $\{(h, \tau) \in \mathbb{R}^2 \mid h \in I, \tau \in \mathbb{R}\}$. Put $\mathcal{I}(h) := \int_{-\infty}^{\infty} q(h, \tau) \, d\tau$ for $h \in I$ and prove that $\mathcal{I}$ is continuous at h_0.

We now apply this to give precise hypotheses on when differentiating under an *improper* integral is valid. Let $J \subseteq \mathbb{R}$ be an interval and let $\mathcal{D}_f := \{(s, \tau) \in \mathbb{R}^2 \mid s \in J, \tau \in \mathbb{R}\}$. Let $f \colon \mathcal{D}_f \to \mathbb{C}$. To ensure that $\int_{-\infty}^{\infty} f(s, \tau) \, d\tau$, the integral that we are differentiating, converges, assume $f(s, \cdot) \in L^1$ for all s. To ensure that the appropriate derivative of this integral exists, assume that f_s exists on $\mathcal{D}_f$. From now on, fix $s \in J$.

Let
$$I_s = \{h \in \mathbb{R} \mid s + h \in J\} \quad \text{and} \quad \mathcal{D}_s = \{(h, \tau) \in \mathbb{R}^2 \mid h \in I_s, \tau \in \mathbb{R}\}.$$

Put
$$q \colon \mathcal{D}_s \to \mathbb{C} \colon (h, \tau) \mapsto \frac{f(s + h, \tau) - f(s, \tau)}{h}.$$

In the notation of the dominated convergence theorem, we are thinking that $Q(\tau) = f_s(s, \tau)$. We are not keeping track of s anymore in the notation of q or Q.

We need $q(h, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$, and that will be true if $f(s, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ for all $s \in J$. We also need $Q \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$, and that will be true if $f_s(s, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ for all $s \in J$. All that is missing is a dominating function.

Here it would be highly annoying if we settled for finding $M \in L^1$ such that $|q(h, \tau)| \leq |M(\tau)|$ for all $(h, \tau) \in \mathcal{D}_s$. The reason is that we typically do not want to differentiate under the integral at just one s; if we had to change M for each s, our result would not be very powerful. Instead, it turns out that we just need to dominate f_s.

That is, suppose there exists $M \in L^1$ such that $|f_s(s, \tau)| \leq |M(\tau)|$ for all $(s, \tau) \in \mathcal{D}$. Then

$$q(h, \tau) = \frac{f(s + h, \tau) - f(s, \tau)}{h} = \frac{1}{h} \int_s^{s+h} f_s(\sigma, \tau) \, d\sigma,$$

and so we estimate (if $h > 0$; if $h < 0$ we just flip the limits of integration)

$$|q(h, \tau)| \leq \frac{1}{h} \int_s^{s+h} |f_s(\sigma, \tau)| \, d\sigma \leq \frac{1}{h} \int_s^{s+h} |M(\tau)| \, d\sigma = \frac{1}{h}(s + h - s)|M(\tau)| = |M(\tau)|.$$

We have therefore checked all of the hypotheses of dominated convergence. (Note that the estimate $|f_s(s, \tau)| \leq |M(\tau)|$ is not enough to guarantee $f(s, \cdot) \in L^1$.)

Here is our streamlined result.

> **33.5 Theorem (Differentiating under the integral for improper integrals).** *Let $J \subseteq \mathbb{R}$ be an interval, let $\mathcal{D}_f := \{(s, \tau) \in \mathbb{R}^2 \mid s \in J, \ \tau \in \mathbb{R}\}$, and let $f \colon \mathcal{D} \to \mathbb{C}$ be a function with the following properties.*
>
> **(i)** $f(s, \cdot) \in L^1$ *for all* $s \in J$.
>
> **(ii)** f_s *exists on* $\mathcal{D}_f$.
>
> **(iii)** $f_s(s, \cdot) \in \mathcal{C}_{\mathrm{pw}}(\mathbb{R})$ *for all* $s \in J$.
>
> **(iv)** *There exists* $M \in L^1$ *such that* $|f_s(s, \tau)| \leq |M(\tau)|$ *for all* $(s, \tau) \in \mathcal{D}_f$.
>
> *Then*
> $$\partial_s \left[\int_{-\infty}^{\infty} f(s, \tau) \, d\tau \right] = \int_{-\infty}^{\infty} f_s(s, \tau) \, d\tau.$$

Day 34: Wednesday, April 2.

> **Material from *Basic Partial Differential Equations* by Bleecker & Csordas**
>
> Pages 339–340 provide a concise historical overview of Laplace's equation. Pages 342–343 discuss steady-state solutions to heat and wave.

We apply this to check over differentiating a convolution. The goal is the identity

$$\partial_x \left[\int_{-\infty}^{\infty} f(x - y)g(y) \, dy \right] = \int_{-\infty}^{\infty} f'(x - y)g(y) \, dy.$$

Of course, we could also put all the burden of differentiability on g.

Put $F(x, y) = f(x - y)g(y)$. To have $F(x, \cdot) \in L^1$, we know that we need either

$$f \in L^1 \text{ and } g \in L^\infty \quad \text{or} \quad f \in L^\infty \text{ and } g \in L^1.$$

In the notation of Theorem 33.5, we assume that $\mathcal{D}_F = \mathbb{R}^2$ and that both f and g are defined on all of $\mathbb{R}$. (This is not always true for functions in L^p, $p = 1, \infty$.) For F_x to exist, we assume $f \in \mathcal{C}^1(\mathbb{R})$; it then follows that $F_x(x, \cdot) \in \mathcal{C}_{\text{pw}}(\mathbb{R})$ since $f' \in \mathcal{C}(\mathbb{R})$ and $g \in \mathcal{C}_{\text{pw}}(\mathbb{R})$.

Last, we want a dominating function for F_x. Because we want $(f * g)' = f' * g$, we are expecting that we need either

$$f' \in L^1 \text{ and } g \in L^\infty \quad \text{or} \quad f' \in L^\infty \text{ and } g \in L^1.$$

In the first case, we would estimate

$$|F_x(x, y)| \le C|f'(x - y)|,$$

where $|g(y)| \le C$ for all y. But then we would need to estimate $|f'(x - y)|$ just in terms of y. That is, we would want $M \in L^1$ such that

$$|f'(x - y)| \le |M(y)|$$

for all x, $y \in \mathbb{R}$. This is asking too much.

> **34.1 Problem (!).** 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.

So, the workaround is to assume $f' \in L^\infty$, which means that we should just take $g \in L^1$ all along. Here is the technical result on differentiating convolution.

> **34.2 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 absolutely integrable—and yet, if f is differentiable, then so is $f * g$.

> **34.3 Problem (!).** The discussion above proves almost all of Theorem 34.2. Under the hypotheses of that theorem, explain why $f' * g \in \mathcal{C}(\mathbb{R})$. [Hint: *Theorem 30.8.*]

We now have the tools to check that the heat kernel does indeed give a solution to the heat equation via convolution in the spatial variable. As usual,

$$H(s, t) = \frac{e^{-s^2/4t}}{\sqrt{4\pi t}}, \quad s \in \mathbb{R}, \ t > 0,$$

and we have $H_t = H_{ss}$. We know that if $f \in L^1 \cup L^\infty$, then the convolution $u(x,t) := (H(\cdot,t) * f)(x)$ is defined, since $H(\cdot,t) \in L^1$ for all $t > 0$. Also, we can check (by induction) that for any $r \geq 0$,

$$\partial_s^r[H](s,t) = p_r\left(\frac{s}{t}\right) H(s,t) \tag{34.1}$$

for a polynomial p_r of degree r (where the coefficients of p_r are independent of t). For example,

$$H_s(s,t) = \frac{e^{-s^2/4t}}{\sqrt{4\pi t}}\left(-\frac{2s}{4t}\right) = p_1\left(\frac{s}{t}\right) H(s,t), \qquad p_1(z) = -\frac{z}{2}$$

and

$$H_{ss}(s,t) = p_1'\left(\frac{s}{t}\right) H(s,t) + p_1\left(\frac{s}{t}\right) H_s(s,t) = p_2\left(\frac{s}{t}\right) H(s,t), \qquad p_2(z) = p_1'(z) + [p_1(z)]^2.$$

The result is that for each $t > 0$ and $r \geq 0$, we have $\partial_s^r[H](\cdot,t) \in L^\infty$ since the exponential in H vanishes faster as $s \to \pm\infty$ than any polynomial.

To differentiate u, we want to use our convolution theory to control the x-derivatives and plain old differentiation under the integral for the t-derivative (since the behavior in t does *not* come from a convolution). Suppose $f \in L^1$. Then we can *repeatedly* apply Theorem 34.2 (with, annoyingly, f in that theorem as $\partial_s^r[H](\cdot,t)$ here, and g in that theorem as f here) to get

$$\partial_x^r[u](x,t) = (\partial_s^r[H](\cdot,t) * f)(x).$$

For $r = 2$, this gives

$$u_{xx}(x,t) = (H_{ss}(\cdot,t) * f)(x) = (H_t(\cdot,t) * f)(x).$$

All that remains is to check the t-derivative (with $x \in \mathbb{R}$ fixed). This could be challenging if we try to do too much all at once. Rewrite

$$u(x,t) = (H(\cdot,t) * f)(x) = \int_{-\infty}^{\infty} H(x-y,t)f(y) \, dy,$$

so the t-derivative of the integrand is

$$H_t(x-y,t)f(y) = H_{ss}(x-y,t)f(y)$$

If we are careless, we would want to bound this as

$$|H_t(x-y,t)f(y)| \leq |M(x,y)| \tag{34.2}$$

for all x, $y \in \mathbb{R}$ and $t > 0$, where $M(x,\cdot) \in L^1$. Remember that we are considering $x \in \mathbb{R}$ fixed, so the dominating function can depend on x. We know that $H_{ss}(\cdot,t) \in L^\infty$ for each t, but what happens as we vary t?

Nothing good: it turns out that

$$\lim_{t \to 0^+} \max_{x \in \mathbb{R}} |\partial_s^r[H](x,t)| = \infty$$

for any $r \geq 0$. The point is that we will not be able to achieve that uniform estimate (34.2). However, there is no reason to try to control all t simultaneously. Instead, let $n \geq 2$ be an integer, and consider $t \in [1/n, n]$. Given any $t > 0$ of interest, eventually t will belong to some such interval for sufficiently large n.

Now consider

$$H_t(x - y, t)f(y) = H_{ss}(x - y, t)f(y) = p_2\left(\frac{x - y}{t}\right)H(x - y, t).$$

Let $|p_2|$ be the polynomial whose coefficients are the absolute values of those of p_2. Then

$$|H_t(x - y, t)f(y)| \leq |p_2|\left(\frac{|x - y|}{|t|}\right)\frac{e^{-(x-y)^2/4t}}{\sqrt{4\pi t}} \leq n|p_2|\left(\frac{|x - y|}{n}\right)\frac{e^{-(x-y)^2/4n}}{\sqrt{4\pi}}.$$

34.4 Problem (!). Check these estimates carefully, recalling $1/n \leq t \leq n$.

The upshot is that because the exponential vanishes faster than any polynomial, there is $C_n > 0$ such that

$$0 \leq n|p_2|\left(\frac{|x - y|}{n}\right)\frac{e^{-(x-y)^2/4n}}{\sqrt{4\pi}} \leq C_n$$

for all x, $y \in \mathbb{R}$. (We would have been content to let C_n depend on x, but that turns out not to be necessary.) And so

$$|H_t(x - y, t)f(y)| \leq C_n|f(y)|$$

for all x, $y \in \mathbb{R}$ and $t \in [1/n, n]$. This is the dominating condition that we wanted.

Let us celebrate with a theorem.

34.5 Theorem. *Let $f \in L^1$. The function*

$$u(x, t) := (H(\cdot, t) * f)(x), \qquad H(s, t) = \frac{e^{-s^2/4t}}{\sqrt{4\pi t}}, \quad s \in \mathbb{R}, \; t > 0,$$

solves the heat equation $u_t = u_{xx}$ for $x \in \mathbb{R}$ and $t > 0$.

34.6 Problem (+). The work above relied on having $f \in L^1$. What if we only want $f \in L^\infty$? Problem 30.5 tells us that $L^1 \not\subseteq L^\infty$, while Theorem 21.9 claimed that we could get a solution to the heat equation with $f \in L^\infty$. (Certainly bounded initial data was necessary for uniqueness, given our reliance on the boundedness in finite time result.)

So, how do we show that $u(x, t) = (H(\cdot, t) * f)(x)$ is sufficiently differentiable if only $f \in L^\infty$? The challenge is that H now needs to bear a double burden: it must carry the differentiability and it must ensure the integrability. The trick is to work with both x and t bounded. Let $n \geq 1$ be an integer and, from now on, suppose $1/n \leq t \leq n$, $-n \leq x \leq n$, and $f \in L^\infty$. We will prove the estimate

$$|\partial_x^r[H](x - y, t)f(y)| \leq C(f, n, r)e^{-\rho_n y^2},$$

where $C(f, n, r) > 0$ depends on f, n, and r, but not x, y, or t, and where $\rho_n > 0$ may depend on n.

(i) First explain why this estimate allows us to invoke Theorem 33.5 to prove $u_t = u_{xx}$.

(ii) Now we prove the estimate. To start, use (34.1) to bound

$$|\partial_x^r[H](x - y, t)f(y)| \leq C(f, n)|p_r\left(\frac{|x - y|}{n}\right)e^{-(x-y)^2/4n}$$

for some $C(f, n) > 0$ that depends on f and n but not x, y, or t. (Remember that $x \in [-n, n]$ and $t \in [1/n, n]$.)

(iii) Use the fact that the exponential vanishes faster than any polynomial at $\pm\infty$ to find a bound of the form

$$|p_r\left(\frac{|x - y|}{n}\right)e^{-(x-y)^2/8n} \leq C(n, r).$$

(Remember that $x \in [-n, n]$ and $t \in [1/n, n]$.)

(iv) Find C_n, $\rho_n > 0$ such that

$$e^{-(x-y)^2/8n} \leq C_n e^{-\rho_n y^2}$$

for all $x \in [-n, n]$ and $y \in \mathbb{R}$. [Hint: *show that for ρ_n appropriately chosen and $|y|$ sufficiently large (relative to n), we have*

$$\left(\rho_n - \frac{1}{8n}\right)y^2 - \left(\frac{x^2 - 2xy}{8n}\right) < 0.$$

When $|y|$ is sufficiently small (relative to n), argue that because $|x| \leq n$, the extreme value theorem in $\mathbb{R}^2$ guarantees an n-dependent bound on this expression over those x and y.]

This effectively exhausts our catalogue of things to study about the heat equation and related techniques to develop (not to mention exhausting us). 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

$$u_{tt} = u_{xx} + u_{yy} \quad \text{and} \quad u_t = u_{xx} + u_{yy},$$

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 $\mathbb{R}^2$.

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

$$u_t + u_x = 0, \qquad u_{tt} - u_{xx} = 0, \quad \text{and} \quad u_t - u_{xx} = 0.$$

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

$$u_{tt} = u_{xx} + u_{yy} \quad \text{and} \quad u_t = u_{xx} + u_{yy},$$

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, April 4.

> **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** $\big\{(x,y) \in \mathbb{R}^2 \mid y > 0\big\}$ 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 $|\widehat{u}(k, y)| \leq M(y)$. 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 $c_2(k)$.

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

$$z(y) = \begin{cases} c_2(k)e^{-|k|y}, & k > 0 \\ c_1(k)e^{-|k|y}, & k < 0 \end{cases} = c(k)e^{-|k|y}, \ k \neq 0, \qquad c(k) := \begin{cases} c_2(k), & k > 0 \\ c_1(k), & k < 0. \end{cases}$$

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 (21.5) 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 34.2 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 33.5 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 $f \in C^2((0, \infty))$. 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 11.7).

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 $f \in C^2((0, \infty))$ 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 ($\star$). 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

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}.$$

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, April 7.

> **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.

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*
>
> $$\mathcal{B}((x_0, y_0); r) := \{(x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 < r^2\}.$$
>
>
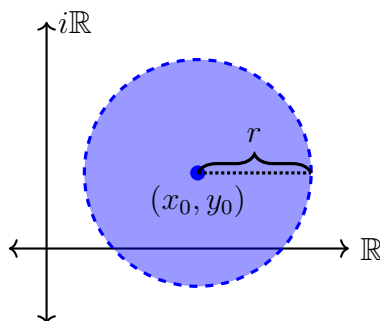
Open balls are immensely useful for controlling the geometry of certain subsets of $\mathbb{R}^2$. In single-variable calculus, most of the interesting results occurred for functions defined on open or closed intervals. Here is the two-dimensional analogue of open intervals.

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* $\mathcal{B}((x,y);r) \subseteq \mathcal{D}$.
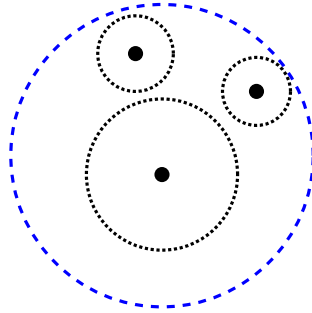
In single-variable calculus we also need to keep track of the endpoints of open intervals: the points a and b function as the "boundary" of (a,b) in the sense that perturbing even slightly from a or b lands one either in (a,b) or in $\mathbb{R} \setminus (a,b)$.

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*

$$\mathcal{B}((x,y);r) \cap \mathcal{D} \neq \varnothing \quad \text{and} \quad \mathcal{B}((x,y);r) \cap (\mathbb{R}^2 \setminus \mathcal{D}) \neq \varnothing.$$
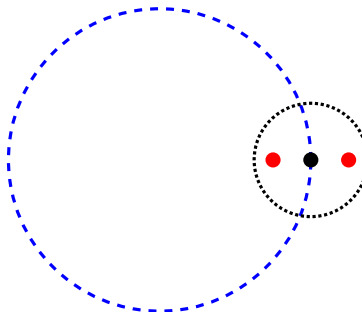
We denote the boundary of $\mathcal{D}$ by $\partial \mathcal{D}$.

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 $\mathcal{D} \subseteq$*

$\mathcal{B}((0,0); R)$.
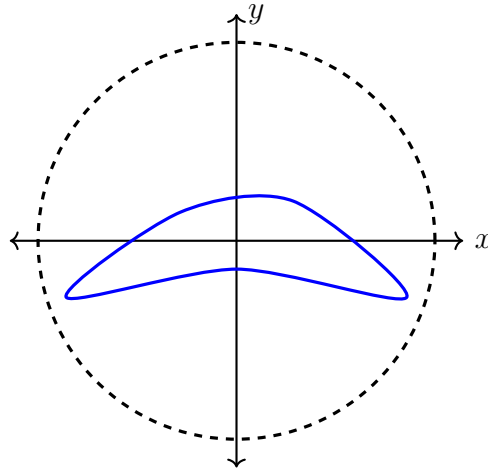
Here are some comforting and useful results.

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.

(iii) Prove that $\overline{\partial\mathcal{D}} = \partial\mathcal{D}$.

(iv) Prove that if $\mathcal{D}$ is bounded, then $\partial\mathcal{D}$ is bounded, too.

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*
>
> $$u(x_{\min}, y_{\min}) = \min_{(x,y) \in \overline{\mathcal{D}}} u(x,y) \quad and \quad u(x_{\max}, y_{\max}) = \max_{(x,y) \in \overline{\mathcal{D}}} u(x,y). \tag{36.1}$$

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/minimum 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 $v_\epsilon(x, y) := u(x, y) + \epsilon w(x, y)$ for some $\epsilon > 0$, which we will select later. We compute $\Delta v_\epsilon(x, y) = \epsilon \Delta w(x, y) > 0$ on $\mathcal{D}$. The extreme value theorem gives $(x_1^\epsilon, y_1^\epsilon) \in \overline{\mathcal{D}}$ such that v attains its maximum at $(x_1^\epsilon, y_1^\epsilon)$. If $(x_1^\epsilon, y_1^\epsilon) \in \mathcal{D}$, the second derivative test gives $\Delta v(x_1^\epsilon, y_1^\epsilon) \le 0$, a contradiction. Thus $(x_1^\epsilon, y_1^\epsilon) \in \partial \mathcal{D}$.

Now we use the nonnegativity of w to compare

$$u(x, y) \le u(x, y) + \epsilon w(x, y) = v_\epsilon(x, y) \le v_\epsilon(x_1^\epsilon, y_1^\epsilon) = u(x_1^\epsilon, y_1^\epsilon) + \epsilon w(x_1^\epsilon, y_1^\epsilon). \tag{36.2}$$

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

$$v_\epsilon(x_1^\epsilon, y_1^\epsilon) \le u(x_1^\epsilon, y_1^\epsilon) + \epsilon M. \tag{36.3}$$

We combine (36.2) and (36.3) to get

$$u(x, y) \le u(x_1^\epsilon, y_1^\epsilon) + \epsilon w(x_1^\epsilon, y_1^\epsilon) \le u(x_1^\epsilon, y_1^\epsilon) + \epsilon M \tag{36.4}$$

for all $(x, y) \in \overline{\mathcal{D}}$ and all $\epsilon > 0$.

As $\epsilon \to 0^+$, that ϵM term will disappear, and we might think that the inequality

$$u(x, y) \le \lim_{\epsilon \to 0^+} u(x_1^\epsilon, y_1^\epsilon)$$

will result. If only $x_1^\epsilon \to x_\star$ and $y_1^\epsilon \to y_\star$ with $(x_\star, y_\star) \in \partial\mathcal{D}$, then we would have

$$u(x, y) \leq u(x_\star, y_\star)$$

for all $(x, y) \in \overline{\mathcal{D}}$. That would give $u(x_\star, y_\star) = \max_{(x,y)\in\overline{\mathcal{D}}} u(x, y)$. But we do not have enough data to say anything about the behavior of x_1^ϵ and x_2^ϵ as $\epsilon \to 0$.

However, we do know $(x_1^\epsilon, y_1^\epsilon) \in \partial\mathcal{D}$ for all ϵ, and we know $\partial\mathcal{D} = \overline{\partial\mathcal{D}}$ by Problem 36.6. We can apply the extreme value theorem again to u on just $\partial\mathcal{D}$ to find $(x_2, y_2) \in \partial\mathcal{D}$ such that $u(x, y) \leq u(x_2, y_2)$ for all $(x, y) \in \partial\mathcal{D}$. Thus $u(x_1^\epsilon, y_1^\epsilon) \leq u(x_2, y_2)$ for all ϵ. We use this to rewrite (36.4) as

$$u(x, y) \leq u(x_2, y_2) + \epsilon M$$

for all $(x, y) \in \overline{\mathcal{D}}$. Sending $\epsilon \to 0^+$, we get $u(x, y) \leq u(x_2, y_2)$ for all $(x, y) \in \overline{\mathcal{D}}$. Thus

$$u(x_2, y_2) = \max_{(x,y)\in\overline{\mathcal{D}}} u(x, y) \quad \text{and} \quad (x_2, y_2) \in \partial\mathcal{D},$$

as desired. ∎

36.10 Problem (!). Reread the proof of the maximum principle for the heat equation (Theorem 22.6) 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: *how are* $\min_{(x,y)\in\overline{\mathcal{D}}} u(x, y)$ *and* $\max_{(x,y)\in\overline{\mathcal{D}}} -u(x, y)$ *related?*]


Day 37: Wednesday, April 9.

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

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.

37.1 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 $\gamma(1) = (x_1, y_1)$.*
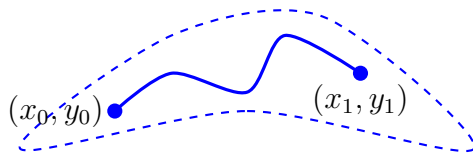
37.2 Theorem (Strong maximum/minimum principle). *Let $\mathcal{D} \subseteq \mathbb{R}^2$ be open and connected. If u is harmonic on $\mathcal{D}$, and if there is $(x_0, y_0) \in \mathcal{D}$ such that $u(x_0, y_0) = \max_{(x,y) \in \mathcal{D}} u(x, y)$ or $u(x_0, y_0) = \min_{(x,y) \in \mathcal{D}} u(x, y)$, 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...

Here is an illustration in the case that we need only three balls. Say that $z = (x_0, y_0)$ and $w = (x_1, y_1)$. We draw two intermediate points z_1 and z_2 between them so that $z_1 \in \mathcal{B}(z; r)$, $z_2 \in \mathcal{B}(z_1; r)$, and $w \in \mathcal{B}(z_2; r)$. Here is the covering of the path between z and w by balls
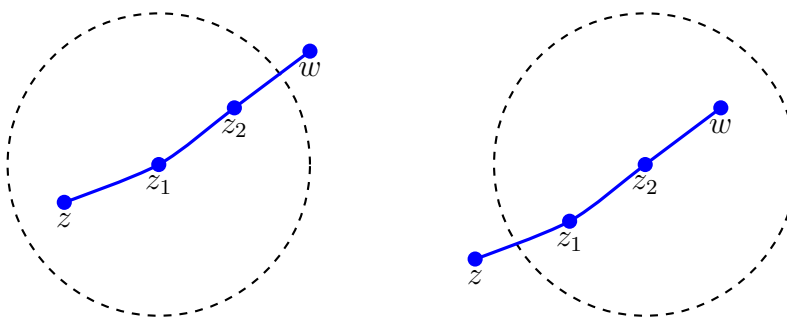
centered at z, z_1, and z_2.
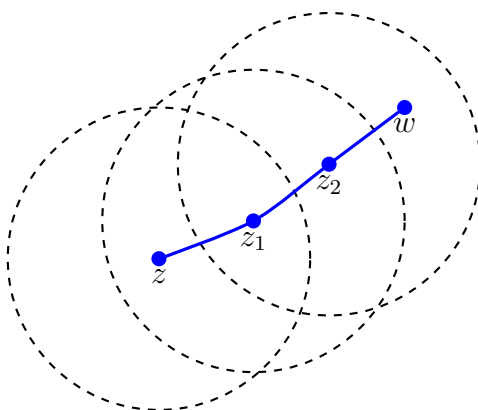
Here we draw just $\mathcal{B}(z; r)$ so we can see $z_1 \in \mathcal{B}(z; r)$.
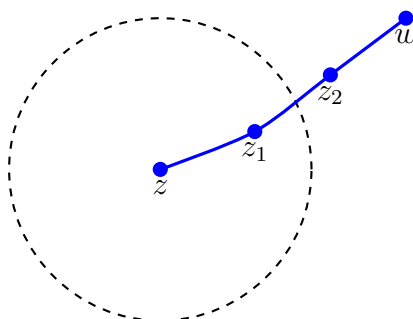
Now we draw $\mathcal{B}(z_1; r)$ with $z_2 \in \mathcal{B}(z_1; r)$ and $\mathcal{B}(z_2; r)$ with $w \in \mathcal{B}(z_2; r)$.

37.3 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 $\mathcal{D}$.*]

37.4 Problem ($\star$). Use the strong maximum principle to prove the weak maximum principle in the case that $\mathcal{D} \subseteq \mathbb{R}^2$ is open, bounded, and connected.

37.5 Problem ($\star$). Prove the "strong minimum" principle part of Theorem 37.2. [Hint: *as usual, think about $-u$.*]

37.6 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 22.6 (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 37.2. (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.)*

A nice immediate consequence of the (weak) maximum principle for Laplace's equation is a uniqueness result.

37.7 Theorem. *Let $\mathcal{D} \subseteq \mathbb{R}^2$ be open and bounded. 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 weak maximum principle 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 $0 \leq u(x,y) \leq 0$ 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 $(-\pi, \pi]$.
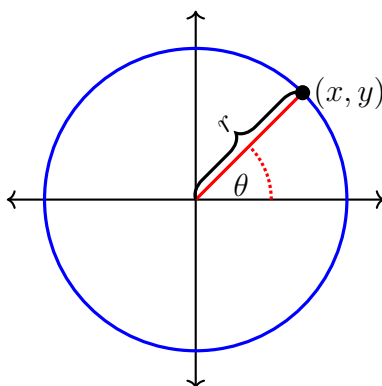
We are interested in the following situation. Suppose that $u \colon \mathcal{B}((0,0); R) \to \mathbb{R}$ is harmonic for some $R > 0$. We can "rectangularize" the domain of u by putting

$$\mathcal{R} := \big\{ (r, \theta) \in \mathbb{R}^2 \mid 0 \le r \le R, \ -\pi \le \theta \le \pi \big\}$$

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?

Day 38: Friday, April 11.

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

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. 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. Page 343 proves the rotational invariance of the Laplacian in the sense of averages of second directional derivatives.

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

$$U_{rr} + \frac{1}{r^2} U_{\theta\theta} + \frac{1}{r} U_r = 0, \ 0 < r \le R, \ -\pi \le \theta \le \pi. \tag{38.1}$$

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:

$$U_r(r, \theta) = u_x \cos(\theta) + u_y \sin(\theta), \tag{38.2}$$

$$U_{rr}(r,\theta) = \big[u_{xx}x\cos(\theta) + u_{xy}\sin(\theta)\big]\cos(\theta) + \big[u_{yx}\cos(\theta) + u_{yy}\sin(\theta)\big]\sin(\theta)$$
$$= u_{xx}\cos^2(\theta) + 2u_{xy}\sin(\theta)\cos(\theta) + u_{yy}\sin^2(\theta),$$

$$U_\theta(r,\theta) = -u_x r\sin(\theta) + u_y r\cos(\theta),$$

and

$$U_{\theta\theta}(r,\theta) = -\big[-u_{xx}r\sin(\theta) + u_{xy}r\cos(\theta)\big]r\sin(\theta)$$
$$- u_x r\cos(\theta)$$
$$+ \big[-u_{yx}r\sin(\theta) + u_{yy}r\cos(\theta)\big]r\cos(\theta)$$
$$- u_y r\sin(\theta)$$

$$= u_{xx}r^2\sin^2(\theta) - 2u_{xy}r^2\sin(\theta)\cos(\theta) + u_{yy}r^2\cos^2(\theta)$$
$$- r\big[u_x r\cos(\theta) + ru_y\sin(\theta)\big]$$

$$= u_{xx}r^2\sin^2(\theta) - 2u_{xy}r^2\sin(\theta)\cos(\theta) + u_{yy}r^2\cos^2(\theta) - rU_r(r,\theta).$$

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

$$U_{rr}(r,\theta) + \frac{1}{r^2}U_{\theta\theta}(r,\theta) + \frac{1}{r}U_r(r,\theta) = \Delta u(r\cos(\theta), r\sin(\theta)), \tag{38.3}$$

so when u is harmonic we end up with (38.1).

> **38.1 Remark.** *Our strategy with Laplace's equation in polar coordinates will be to assume that $\Delta u = 0$ and then obtain that (38.1) holds. However, one could work backwards: solve (38.1) 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.*

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.

38.2 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

$$\phi'(r) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_r \big[u(r\cos(\theta), r\sin(\theta))\big] \ d\theta$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \big[u_x(r\cos(\theta), r\sin(\theta))\cos(\theta) + u_y(r\cos(\theta), r\sin(\theta))\sin(\theta)\big] \ d\theta.$$

Now here is the advantage of polar coordinates: we recall from (38.2) 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 $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le r^2\}$. Since $\Delta u = 0$, this gives the desired identity $\phi'(r) = 0$. ■

Now here is that last calculation.

38.3 Lemma. *Let $u \in C^2(\mathcal{B}((0,0); R))$. As usual, for $0 \leq r \leq R$ and $\theta \in \mathbb{R}$, let $U(r, \theta) := u(r \cos(\theta), r \sin(\theta))$. Then*

$$\iint_{x^2+y^2 \leq \rho^2} \Delta u(x, y) \ dx \ dy = \rho \int_{-\pi}^{\pi} U_r(\rho, \theta)$$

for any $\rho \in (0, R)$.

Proof. We switch to polar coordinates. Recall that for $f \in C(\mathcal{B}(0,0); R)$ we have

$$\iint_{x^2+y^2 \leq \rho^2} f(x, y) \ dx \ dy = \int_0^{2\pi} \int_0^\rho f(r \cos(\theta), r \sin(\theta)) r \ dr \ d\theta. \tag{38.4}$$

Thus

$$\mathcal{I}(\rho) := \iint_{x^2+y^2 \leq \rho^2} \Delta u(x, y) \ dx \ dy = \int_0^{2\pi} \int_0^\rho \Delta u(r \cos(\theta), r \sin(\theta)) r \ dr \ d\theta$$

$$= \int_0^{2\pi} \int_0^\rho \left(U_{rr}(r, \theta) + \frac{1}{r^2} U_{\theta\theta}(r, \theta) + \frac{1}{r} U_r(r, \theta) \right) r \ dr \ d\theta. \tag{38.5}$$

In (38.3) we calculated

$$\Delta u(r \cos(\theta), r \sin(\theta)) = U_{rr}(r, \theta) + \frac{1}{r^2} U_{\theta\theta}(r, \theta) + \frac{1}{r} U_r(r, \theta) \tag{38.6}$$

for $r \neq 0$. Since $\Delta u \in C(\mathcal{B}((0,0); R))$, the limit as $r \to 0$ of the right side is defined, and so the integral over $[0, \rho]$ with respect to r in (38.5) 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 $\text{sinc}(x)$ as $\sin(x)/x$ even at $x = 0$.)

Then

$$\mathcal{I}(\rho) = \underbrace{\int_0^{2\pi} \int_0^\rho \left[U_r(r, \theta) + r U_{rr}(r, \theta) \right] \ dr \ d\theta}_{\mathcal{I}_1(\rho)} + \underbrace{\int_0^{2\pi} \int_0^\rho \frac{U_{\theta\theta}(r, \theta)}{r} \ dr \ d\theta}_{\mathcal{I}_2(\rho)}.$$

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 (38.3) we have

$$\frac{U_{\theta\theta}(r, \theta)}{r} = r \Delta u(r \cos(\theta), r \sin(\theta)) - r U_{rr}(r, \theta) - U_r(r, \theta),$$

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:

$$U_r(r, \theta) + r U_{rr}(r, \theta) = (1 \cdot U_r(r, \theta)) + r U_{rr}(r, \theta) = \partial_r[r U_r(r, \theta)],$$

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 28.1, 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. ∎

38.4 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* $(x,y) \in \mathcal{B}((0,0); R)$.]

38.5 Problem ($\star$). 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 $r \in (0, R)$.

Day 39: Monday, April 14. ▬▬▬

It turns out that the mean value principle also characterizes harmonic functions in the following sense.

39.1 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

$$\phi'(r) = \frac{1}{2\pi r} \iint_{(x-x_0)^2 + (y-y_0)^2 \leq r^2} \Delta u(x, y) \, dx \, dy > 0,$$

which is impossible if ϕ is constant. ∎

39.2 Problem ($\star$). Let u be harmonic on $\mathcal{B}((0,0); R)$. Prove that

$$u(0, 0) = \frac{1}{\pi r^2} \iint_{x^2 + y^2 \leq r^2} u(x, y) \, dx \, dy$$

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 the polar integral identity (38.4).]

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.

39.3 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 by Problem 15.4.

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).

39.4 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 (38.2) 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. ∎

Day 40: Wednesday, April 16.

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

$$\mathcal{D} := \left\{ (x, y) \in \mathbb{R}^2 \mid 0 < x < 1,\ 0 < y < 1 \right\}.$$

We will study solutions to the Dirichlet problem

$$\begin{cases} \Delta u = 0 \text{ in } \mathcal{D} \\ u = f \text{ on } \partial\mathcal{D}. \end{cases} \tag{40.1}$$

40.1 Problem ($\star$). Why are solutions unique?

We first prove a highly useful "integration by parts" identity.

40.2 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

$$\int_a^b f''g = -\int_a^b f'g'.$$

40.3 Lemma. *Let $u \in C^2(\mathcal{D}) \cap C(\overline{\mathcal{D}})$ and $v \in C^1(\mathcal{D}) \cap C(\overline{\mathcal{D}})$ with $v = 0$ on $\partial\mathcal{D}$. Then*

$$\iint_{\overline{\mathcal{D}}} (\Delta u)v = -\iint_{\overline{\mathcal{D}}} (u_x v_x + u_y v_y).$$

Proof. We have

$$\iint_{\overline{\mathcal{D}}} (\Delta u)v = \int_0^1 \int_0^1 (u_{xx}(x,y) + u_{yy}(x,y))v(x,y) \ dx \ dy.$$

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 $\partial\mathcal{D}$, 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

$$\iint_{\overline{\mathcal{D}}} u_{xx}(x,y)v(x,y) \ dx \ dy = -\iint_{\overline{\mathcal{D}}} u_x(x,y)v_x(x,y) \ dx \ dy.$$

A similar calculation, along with an interchange in the order of integration, shows

$$\iint_{\overline{\mathcal{D}}} u_{yy}(x,y)v(x,y) \ dx \ dy = -\iint_{\overline{\mathcal{D}}} u_y(x,y)v_y(x,y) \ dx \ dy, \tag{40.2}$$

which leads to the desired identity. ∎

40.4 Problem (!). Do that similar calculation with the interchange of integrals to obtain (40.2).

40.5 Problem (⋆). Suppose that we have this integration by parts identity for a more general $\mathcal{D} \subseteq \mathbb{R}$:

$$\iint_{\overline{\mathcal{D}}} (\Delta u)v = -\iint_{\overline{\mathcal{D}}} (u_x v_x + u_y v_y) \tag{40.3}$$

for $u \in C^2(\mathcal{D}) \cap C(\overline{\mathcal{D}})$ and $v \in C^1(\mathcal{D}) \cap C(\overline{\mathcal{D}})$ 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 (40.3) use $u = v$.*]

Day 41: Friday, April 18.

We apply this integration by parts result in the following quite possibly non-obvious way. Let
$$\mathcal{V} := \left\{ w \in \mathcal{C}^2(\mathcal{D}) \cap \mathcal{C}(\overline{\mathcal{D}}) \mid w = f \text{ on } \partial\mathcal{D} \right\}.$$

Define
$$\mathcal{E} \colon \mathcal{V} \to \mathbb{R} \colon w \mapsto \iint_{\overline{\mathcal{D}}} (w_x^2 + w_y^2).$$

We will show that any solution u to the Dirichlet problem (40.1) 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.

Proving that a solution to the Dirichlet problem minimizes $\mathcal{E}$ is not all that difficult, with one classical trick: adding 0. Let u solve (40.1) 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 that classical trick of adding 0:
$$\mathcal{E}[w] = \mathcal{E}[u + (w - u)].$$

Now we work on $\mathcal{E}[u + v]$ with $v := w - u$ and find
$$\mathcal{E}[u + v] = \mathcal{E}[u] + \mathcal{E}[v] + 2 \iint_{\overline{\mathcal{D}}} (u_x v_x + u_y v_y).$$

41.1 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 $\partial\mathcal{D}$. So, we *can* integrate by parts to find
$$\iint_{\overline{\mathcal{D}}} (u_x v_x + u_y v_y) = \iint_{\overline{\mathcal{D}}} (\Delta u) v = 0,$$

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 $\partial\mathcal{D}$, so we just need to show $\Delta u = 0$; that Δu is defined follows from the differentiability of u by its

membership in $\mathcal{V}$. This is the harder part part, and the first trick is to introduce a new function space.

For $r \geq 0$, let

$$\mathcal{C}_0^r(\mathcal{D}) := \left\{ v \in \mathcal{C}^r(\mathcal{D}) \cap \mathcal{C}(\overline{\mathcal{D}}) \mid v = 0 \text{ on } \partial\mathcal{D} \right\}.$$

41.2 Problem (!). How did $\mathcal{C}_0^r(\mathcal{D})$ show up in Lemma 40.3?

Let $v \in \mathcal{C}_0^2(\mathcal{D})$ and $s \in \mathbb{R}$. If we know $u + sv \in \mathcal{V}$, then since u minimizes $\mathcal{E}$ we must have $\mathcal{E}[u] \leq \mathcal{E}[u + sv]$.

41.3 Problem (!). Explain why $u + sv \in \mathcal{V}$.

We expand, similar to some work above,

$$\mathcal{E}[u + sv] = \mathcal{E}[u] + s^2 \mathcal{E}[v] + 2s \iint_{\overline{\mathcal{D}}} (u_x v_x + u_y v_y).$$

41.4 Problem (!). Check that.

Integrating by parts, we find

$$\mathcal{E}[u] \leq \mathcal{E}[u] + s^2 \mathcal{E}[v] + 2s \iint_{\overline{\mathcal{D}}} (u_x v_x + u_y v_y) = \mathcal{E}[u] + s^2 \mathcal{E}[v] - 2s \iint_{\overline{\mathcal{D}}} (\Delta u)v,$$

and so

$$s \iint_{\overline{\mathcal{D}}} (\Delta u)v \leq \frac{s^2}{2} \mathcal{E}[v].$$

This is true for all $s \in \mathbb{R}$. When $s > 0$, we just have

$$\iint_{\overline{\mathcal{D}}} (\Delta u)v \leq \frac{s}{2} \mathcal{E}[v]$$

and so

$$\iint_{\overline{\mathcal{D}}} (\Delta u)v \leq \lim_{s \to 0^+} \frac{s}{2} \mathcal{E}[v] = 0.$$

41.5 Problem ($\star$). By considering $s < 0$, show as well that

$$0 \leq \iint_{\overline{\mathcal{D}}} (\Delta u)v.$$

Thus

$$\iint_{\overline{\mathcal{D}}} (\Delta u)v = 0.$$

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 $\overline{\mathcal{D}}$. The tool returns the data 0 all the time. It turns out that this is enough to conclude that $\Delta u = 0$ on $\mathcal{D}$, but it is easier to see why in one dimension.

41.6 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 $r \geq 0$. Then $f = 0$.*

Proof. Suppose not. Then there is $x_0 \in [a, b]$ such that $f(x_0) \neq 0$. 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.
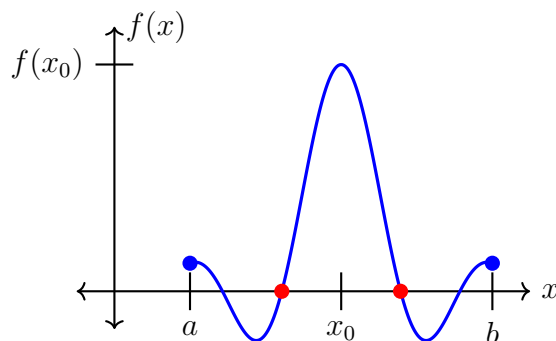
We claim there is $g_\delta \in \mathcal{C}^\infty([a, b])$ such that $g_\delta(x) = 0$ for $a \leq x \leq x_0 - \delta$ and $x_0 + \delta \leq x \leq b$ but $g_\delta(x) > 0$ for $x_0 - \delta \leq x \leq x_0 + \delta$. Here is another picture.
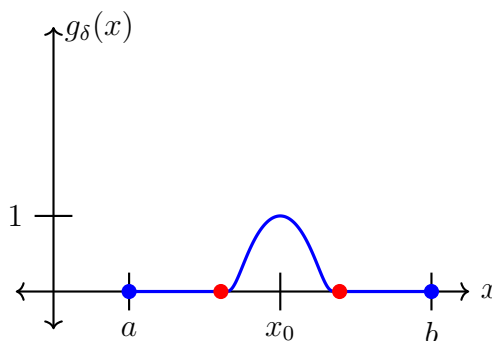
And here is a picture of how the product fg_δ behaves.
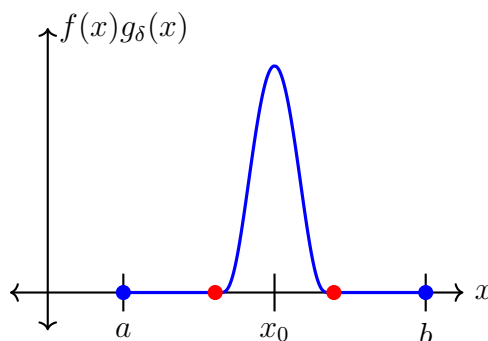
Assuming the existence of this g_δ, we have

$$0 < \int_{x_0-\delta}^{x_0+\delta} f(x)g_\delta(x) \ dx = \int_a^b f(x)g_\delta(x) \ dx = 0,$$

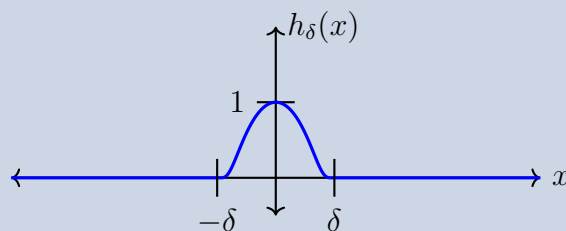where the last equality follows from the hypothesis. This is a contradiction. ∎

41.7 Problem ($\star$). Prove the preceding theorem in the case $x_0 = a$. (The case $x_0 = b$ is similar, and you do not need to treat that.) [Hint: *show that if $f(x_0) \neq 0$, then there is $x_1 \in (a,b)$ such that $f(x_1) \neq 0$. Invoke the old proof.*]

Now we show the existence of that g_δ.

41.8 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 $h_\delta \in C^\infty(\mathbb{R})$.

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 $x \to \infty$. ∎

41.9 Problem (+). Perform that technical induction argument.

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

$$(x, y) \mapsto \begin{cases} e^{1/\delta^2} \exp\left(-\dfrac{1}{\delta^2 - \left(x^2 + y^2\right)}\right), & x^2 + y^2 < \delta^2 \\ 0, & x^2 + y^2 \geq \delta^2. \end{cases}$$

Day 42: Monday, April 21.

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

$$u(x, t) = X(x)T(t) \tag{42.1}$$

with $X \in \mathcal{C}^2(\mathbb{R})$ and $T \in \mathcal{C}^1(\mathbb{R})$, then

$$u_t(x, t) = X(x)T'(t) \quad \text{and} \quad u_{xx}(x, t) = X''(x)T(t).$$

Thus $u(x, t) = X(x)T(t)$ solves the heat equation $u_t = u_{xx}$ if and only if

$$X(x)T'(t) = X''(x)T(t). \tag{42.2}$$

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 (42.2) by $X(x)T(t)$ to find

$$\frac{X''(x)}{X(x)} = \frac{T'(t)}{T(t)}. \tag{42.3}$$

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 (42.3), we find $\lambda \in \mathbb{R}$ such that

$$\frac{X''(x)}{X(x)} = \lambda \quad \text{and} \quad \frac{T'(t)}{T(t)} = \lambda \tag{42.4}$$

for all x, $t \in \mathbb{R}$ (or, at least those at which X and T are nonzero). We can view (42.1), (42.3), and (42.4) 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 (42.4) are really

$$T'(t) = \lambda T(t) \quad \text{and} \quad X''(x) - \lambda X(x) = 0, \tag{42.5}$$

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?

42.1 Problem (!). No. Convince yourself that if X and T solve (42.5), 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

$$u(x,t) = \left(c_1 e^{\lambda x} + c_2 e^{-\lambda x}\right) T(0) e^{\lambda^2 t} = \left(T(0)c_1 e^{\lambda x} + T(0)c_2 e^{-\lambda x}\right) e^{\lambda^2 t}. \tag{42.6}$$

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

$$u(x,t) = \left(c_1 e^{\lambda x} + c_2 e^{-\lambda x}\right) e^{\lambda^2 t}. \tag{42.7}$$

More precisely, every solution of the form (42.6) is of the form (42.7). 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 (42.7) in the original form (42.6).

Similarly, for the other two cases we have product solutions of the form

$$u(x,t) = c_1 x + c_2 \tag{42.8}$$

and

$$u(x,t) = \left(c_1 \cos(\lambda x) + c_2 \sin(\lambda x)\right) e^{-\lambda^2 t}, \tag{42.9}$$

where here we may assume $\lambda > 0$. (Although $\lambda < 0$ still works—why?)

42.2 Problem ($\star$). 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.

> **42.3 Problem ($\star$).** 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

$$\begin{cases} u_t = u_{xx}, \ 0 \le x \le \pi, \ t \in \mathbb{R} \\ u(0, t) = u(\pi, t) = 0, \ t \in \mathbb{R}. \end{cases} \tag{42.10}$$

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 (42.7) 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.

> **42.4 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 (42.8) to meet the boundary conditions.

> **42.5 Problem (!).** What exactly happens?

And so we are down to solutions of the form (42.9). 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

$$u(x,t) = c_2 \sin(kx)e^{-k^2 t} \tag{42.11}$$

solves (42.10). 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 (42.11) solves (42.10) with a slightly more complicated initial condition. Specifically,

$$u(x,t) := \sum_{k=1}^{n} b_k \sin(kx)e^{-k^2 t}$$

solves

$$\begin{cases} u_t = u_{xx}, \ 0 \leq x \leq \pi, \ t \in \mathbb{R} \\ u(0,t) = u(\pi,t) = 0, \ t \in \mathbb{R} \\ u(x,0) = f(x), \ 0 \leq x \leq \pi \end{cases}$$

with

$$f(x) = \sum_{k=1}^{n} b_k \sin(kx).$$

42.6 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 $f \in C^2([0,\pi])$ 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.

42.7 Problem ($\star$). What happens if we try to use product solutions of the forms (42.7) or (42.8) to solve this problem?

We get to the point and try solutions in the form (42.9). We first differentiate:

$$u_x(x,t) = \left(-c_1 \lambda \sin(\lambda x) + c_2 \lambda \cos(\lambda x)\right) e^{-\lambda^2 t}.$$

Then we want

$$\left(c_1 \cos(-\lambda\pi) + c_2 \sin(-\lambda\pi)\right) e^{-\lambda^2 t} = \left(c_1 \cos(\lambda\pi) + c_2 \sin(\lambda\pi)\right) e^{-\lambda^2 t}$$

and

$$\left(-c_1 \lambda \sin(-\lambda\pi) + c_2 \lambda \cos(-\lambda\pi)\right) e^{-\lambda^2 t} = \left(-c_1 \lambda \sin(\lambda\pi) + c_2 \lambda \cos(\lambda\pi)\right) e^{-\lambda^2 t}.$$

Simplifying, this leads to the system

$$2c_2 \sin(\lambda\pi) = 0 \quad \text{and} \quad 2c_1 \lambda \sin(\lambda\pi) = 0.$$

42.8 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

$$u(x,t) = \left(c_1 \cos(kx) + c_2 \sin(kx)\right) e^{-k^2 t}.$$

By linearity,

$$u(x,t) = \sum_{k=0}^{n} \big(a_k \cos(kx) + b_k \sin(kx)\big)e^{-k^2 t}$$

solves

$$\begin{cases} u_t = u_{xx}, \ -\pi \leq x \leq \pi, \ t \in \mathbb{R} \\ u(-\pi, t) = u(\pi, t), \ u_x(-\pi, t) = u_x(\pi, t), \ t \in \mathbb{R} \\ u(x, 0) = f(x), \ -\pi \leq x \leq \pi \end{cases}$$

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.

42.9 Problem ($\star$). 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 (23.1) seem less mysterious?

Day 43: Wednesday, April 23.

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.

43.1 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 $\left(\sum_{k=1}^{n} z_k\right)$. 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.

43.2 Example. Let $r \in \mathbb{C}$ with $|r| < 1$. Then for any integer $n \geq 1$ we have

$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r}.$$

This can be proved by multiplying both sides of the desired equality by $1 - r$ and then using the formula for a finite telescoping sum (which is quite worth knowing!). With this finite sum established, we take the limit as $n \to \infty$ and use $|r| < 1$ to obtain

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}.$$

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.

43.3 Theorem. *Let (z_k) and (w_k) be sequences with $|z_k| \leq |w_k|$. 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

$$f(x) = \sum_{k=0}^{\infty} \big(a_k \cos(kx) + b_k \sin(kx) \big)$$

and then

$$u(x, t) = \sum_{k=0}^{\infty} e^{-k^2 t} \big(a_k \cos(kx) + b_k \sin(kx) \big).$$

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 25.4, 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

$$\langle \cos(k\cdot), \sin(j\cdot) \rangle = 0 \text{ for all } k, \ j.$$

Perhaps the simplest initial temperature distribution f for the problem at hand is

$$f(x) = \sum_{k=0}^{n} \big(a_k \cos(kx) + b_k \sin(kx) \big), \tag{43.1}$$

which we call a **TRIGONOMETRIC POLYNOMIAL**. Linearity of the integral gives

$$\langle f, \cos(j\cdot) \rangle = \sum_{k=0}^{n} \big(a_k \langle \cos(k\cdot), \cos(j\cdot) \rangle + b_k \langle \sin(k\cdot), \cos(j\cdot) \rangle \big) = \begin{cases} 2\pi a_0, \ k = j = 0 \\ \pi a_j, \ k = j \ge 1 \\ 0, \ k \ne j. \end{cases}$$

Similarly, we have

$$\langle f, \sin(j\cdot)\rangle = \begin{cases} \pi b_j, & k = j \geq 1 \\ 0, & k \neq j. \end{cases}$$

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 (43.1) 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 $f \in \mathcal{C}_{\mathrm{pw}}([-\pi, \pi])$.

43.4 Definition. *Let $f \in \mathcal{C}_{\mathrm{pw}}([-\pi, \pi])$. 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.

43.5 Example. It is probably worth doing one explicit Fourier series calculation. Let

$$f\colon [-\pi, \pi] \to \mathbb{R} \colon x \mapsto \begin{cases} 1, & |x| < \pi/2 \\ 0, & |x| > \pi/2. \end{cases}$$

Here is its graph.
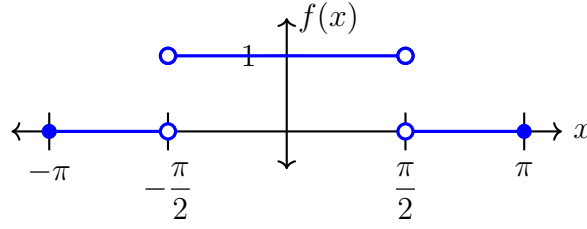
We have

$$\mathsf{a}_k[f] = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) \ dx = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \cos(kx) \ dx.$$

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,

$$\mathsf{FS}[f](x) = \frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{\infty} \left(\mathsf{a}_k[f] \cos(kx) + \mathsf{b}_k[f] \sin(kx) \right) = \frac{1}{2} + \sum_{j=0}^{\infty} \frac{2(-1)^j}{(2j + 1)\pi} \cos((2j + 1)x)$$

This formula by itself is probably not very enlightening, and while it may resemble an alternating series, proving its convergence using only classical calculus methods could be, at best, annoying.

For $f \in \mathcal{C}_{\mathrm{pw}}([-\pi, \pi])$ and $x \in [-\pi, \pi]$, we can always interpret $\mathsf{FS}[f](x)$ as a sequence of partial sums:

$$\mathsf{FS}[f](x) = \left(\frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{n} \left(\mathsf{a}_k[f] \cos(kx) + \mathsf{b}_k[f] \sin(kx) \right) \right).$$

Ideally we would like to have $\mathsf{FS}[f](x) = f(x)$, i.e.,

$$f(x) = \lim_{n \to \infty} \frac{\mathsf{a}_0[f]}{2} + \sum_{k=1}^{n} \left(\mathsf{a}_k[f] \cos(kx) + \mathsf{b}_k[f] \sin(kx) \right).$$

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., $f(-\pi) = f(\pi)$.

> **43.6 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: they are really the only possible choice. If there are sequences (a_k) and (b_k) such that

$$f(x) = \sum_{k=0}^{\infty} \left(a_k \cos(kx) + b_k \sin(kx) \right)$$

for each $x \in [-\pi, \pi]$, and if we can interchange integration and summation in the sense that

$$\langle f, g \rangle = \sum_{k=0}^{\infty} \left(a_k \langle \cos(k\cdot), g \rangle + b_k \langle \sin(k\cdot), g \rangle \right)$$

for any $g \in \mathcal{C}_{\mathrm{pw}}([-\pi, \pi])$, then just as with the trigonometric polynomial, we obtain

$$a_0 = 2\mathsf{a}_0[f], \qquad a_k = \mathsf{a}_k[f], \ k \geq 1, \qquad \text{and} \qquad b_k = \mathsf{b}_k[f], \ k \geq 1.$$

And here is a precise indication of why the Fourier coefficients are the right ones: they make $\mathsf{FS}[f]$ equal f, if f is sufficiently nice.

> **43.7 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.

> **43.8 Problem (!).** Suppose that u solves
>
> $$\begin{cases} u_t = u_{xx}, \ -\pi \leq x \leq \pi, \ t \in \mathbb{R} \\ u(-\pi, t) = u(\pi, t), \ u_x(-\pi, t) = u_x(\pi, t), \ t \in \mathbb{R} \\ u(x, 0) = f(x). \end{cases}$$

Prove that $f \in \mathcal{C}^2([-\pi, \pi])$ with

$$f(-\pi) = f(\pi) \quad \text{and} \quad f'(-\pi) = f'(\pi).$$

The hypotheses of Theorem 43.7 can be relaxed, and the result is that we still obtain convergence of the Fourier series to "something" resembling the original function.

We need a variant on the function space $\mathcal{C}^1_{\text{pw}}(\mathbb{R})$ from Definition 27.4. Recall that we use the notation

$$f(x^{\pm}) := \lim_{s \to x^+ \pm} f(s)$$

for a function f and a point x.

43.9 Definition. *Let $a, b \in \mathbb{R}$ with $a < b$. Let $f \in \mathcal{C}_{\text{pw}}([a, b])$ and suppose that*

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x^+)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(x+h) - f(x^-)}{h}$$

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

$$\lim_{h \to 0^+} \frac{f(a+h) - f(a^+)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(b+h) - f(b^-)}{h}$$

*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 $\mathcal{C}_{\text{pw}}([a, b])$.) Then we say that f is **PIECEWISE CONTINUOUSLY DIFFERENTIABLE** on $[a, b]$, and we denote the set of all such functions by $\mathcal{C}^1_{\text{pw}}([a, b])$.*

Here is the best that we can say about the convergence of Fourier series in general.

43.10 Theorem. *Let $f \in \mathcal{C}^1_{\text{pw}}([-\pi, \pi])$. Then*

$$\mathsf{FS}[f](x) = \begin{cases} \dfrac{f(x^+) + f(x^-)}{2}, & x \in (-\pi, \pi) \\[2mm] \dfrac{f(-\pi^+) + f(\pi^-)}{2}, & x = \pm\pi. \end{cases}$$

Except at the endpoints $\pm\pi$, this theorem is effectively the analogue of the Fourier inversion result in Theorem 27.5. In particular, $\mathsf{FS}[f](x) = f(x)$ at any $x \in (-\pi, \pi)$ at which f is continuous. 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.

43.11 Example. **(i)** We found the formal Fourier series of

$$f\colon [-\pi, \pi] \to \mathbb{R}\colon x \mapsto \begin{cases} 1, & |x| < \pi/2 \\ 0, & |x| > \pi/2 \end{cases}$$

in Example 43.5. The formula was not very helpful. However, now we can use the convergence theorem. By continuity, we have $\mathsf{FS}[f](x) = f(x)$ for $-\pi < x < \pi/2$, $-\pi/2 < x < \pi/2$, and $\pi/2 < x < \pi$. At the jump discontinuities at $x = \pm\pi/2$, we take averages of the left and right limits. At $x = \pm\pi$, it turns out that $\mathsf{FS}[f](\pi) = \mathsf{FS}[f](-\pi)$, which more generally is a consequence of the evenness of f. All together, we have

$$\mathsf{FS}[f](x) = \begin{cases} 0, & -\pi \le x < -\pi/2 \\ 1/2, & x = -\pi/2 \\ 1, & -\pi/2 < x < \pi/2 \\ 1/2, & x = \pi/2 \\ 0, & \pi/2 < x \le \pi. \end{cases}$$
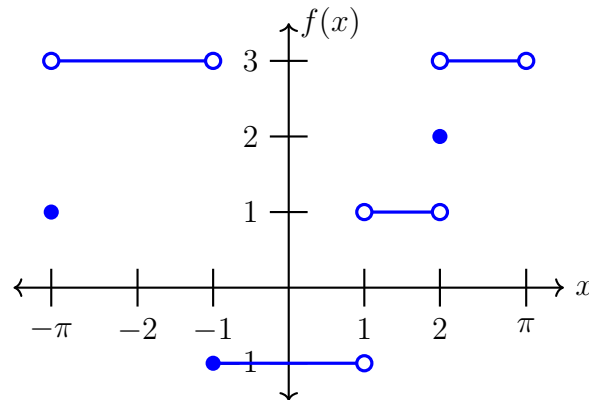
We graph f and $\mathsf{FS}[f]$ to compare them.
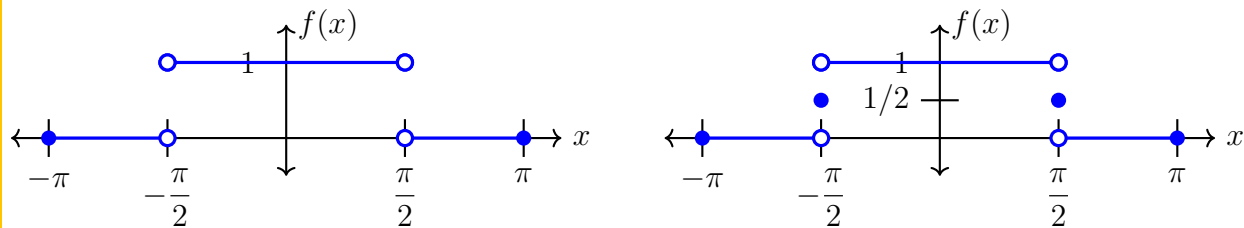
(ii) Define $f \in \mathcal{C}_{\mathrm{pw}}([-\pi, \pi])$ by

$$f(x) = \begin{cases} 1, & x = -\pi \\ 3, & -\pi < x < -1 \\ -1, & -1 \le x < 1 \\ 1, & 1 \le x < 2 \\ 2, & x = 2 \\ 3, & 2 < x < \pi. \end{cases}$$

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 FS[f].