MATH 4310: PARTIAL DIFFERENTIAL EQUATIONS

Daily Log for Lectures and Readings

Timothy E. Faver

February 21, 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 $c_1u + c_2v$ is also a solution.

1.1 Problem (!). Prove that.

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

$u_t + uu_x = 0$	Burgers's equation
$u_t + u_{xxx} + uu_x = 0$	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 $c_1u + c_2v$ 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:

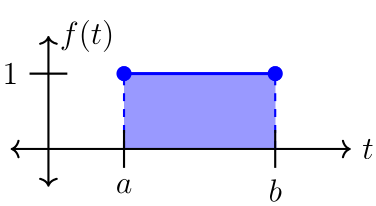
$$\int_a^b f = \int_a^b f(x)\ dx = \int_a^b f(u)\ du = \int_a^b f(s)\ ds = \int_a^b f(t)\ dt = \int_a^b f(\tau)\ d\tau.$$

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

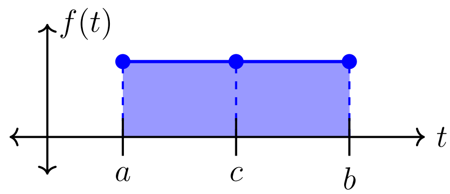
$$\int_a^b 1\ dt = b - a.$$



($\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

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

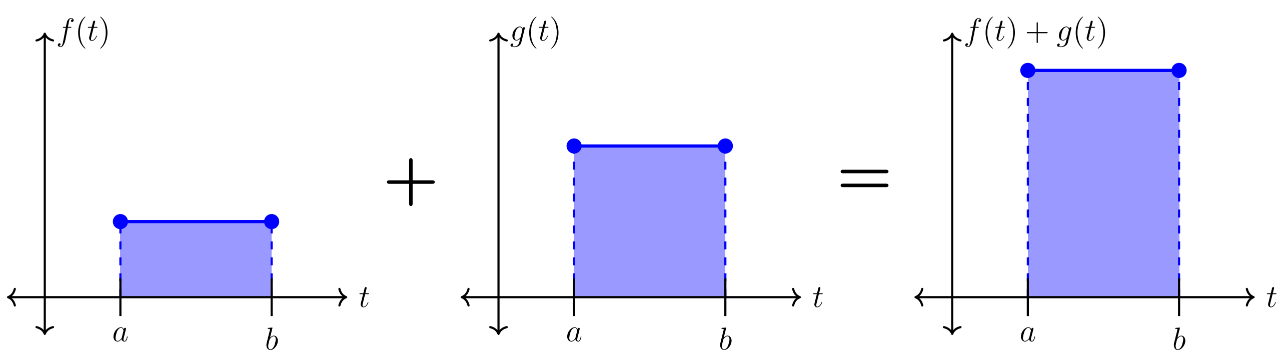


($\int$3) If f is nonnegative, the net area of the region between the graph of f and the interval $[a, b]$ should be the genuine area of the region between the graph of f and the interval $[a, b]$, and this should be a positive quantity. So, we expect that if $0 \leq f(t)$ on $[a, b]$, then

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

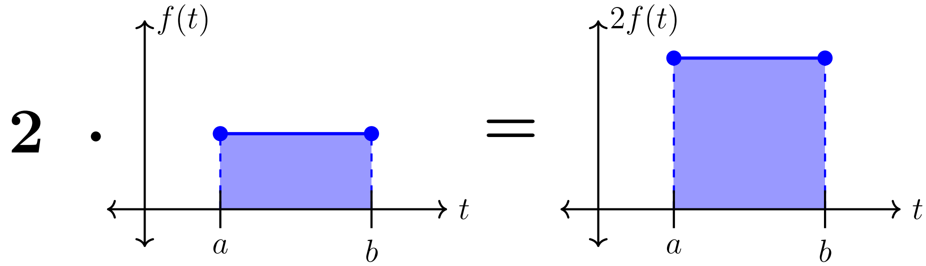
($\int$4) Adding two functions $f, g\colon [a, b] \subseteq \mathbb{R} \to \mathbb{R}$ should "stack" the graphs of f and g on top of each other. Then the region between the graph of f and the interval $[a, b]$ gets "stacked" on top of region between the graph of g and the interval $[a, b]$. Consequently, 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 $a \leq b$ and $0 \leq f(t)$ for all $t \in [a,b]$, then*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

2.3 Problem (+). 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 $a \le b$. If $f(t) \le g(t)$ for all $t \in [a, b]$, show that

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

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

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

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

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

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

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

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

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

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

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

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

Then F is an antiderivative of f on I.

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

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

equivalently,

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

We first compute

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

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

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

It therefore suffices to show that

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

and we do that in the following lemma. ■

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

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

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

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

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

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

for any $t \in I$.

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

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

We now need to show that

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

This will involve the definition of continuity.

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

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

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

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

thus (2.5) holds. ■

2.8 Problem ($\star$). 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*

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

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

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

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

Substitution implies

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

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

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

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

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

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

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

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

(ii) First explain why

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

Then substitute $u = t - p$ to show

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

and use the p-periodicity of f.

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

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g. \tag{2.7}$$

[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 (⋆). 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

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

Show that

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

By considering the special case $f(x) = x^2$, explain why we might call this a "difference of squares" estimate. [Hint: *use Example 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 $t \in I$. 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

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

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

$$\begin{cases} y'(t) \le ay(t), \ t \in \mathbb{R} \\ y(0) = 00 \le y(t), \ t \in \mathbb{R} \end{cases}$$

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

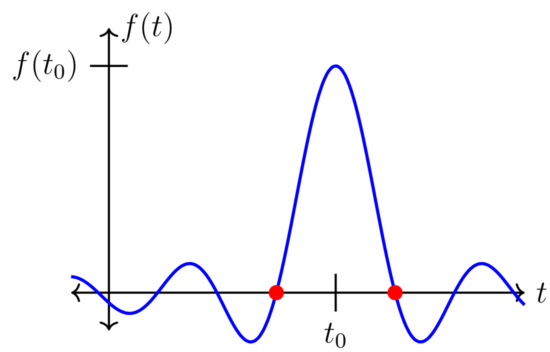
$$|f(t)| \le M$$

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 $y_0 \neq 0$. That is, $y(0) \neq 0$. 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 $y(t) \neq 0$. 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 δ, $\epsilon > 0$ 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 $u_t = u_x = 0$.

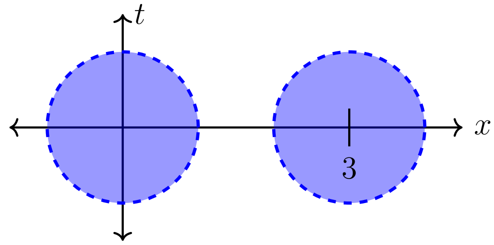
(i) First suppose that we want to solve this PDE on $\mathbb{R}^2$. We expect that since u is constant in both x and t, it is simply constant. Here is why. Fix $x_0 \in \mathbb{R}$ and set $v(t) = u(x_0, t)$. By the way, in the future we might refer to this function v as $u(x_0, \cdot)$; this notation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

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

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 γ_1, $\gamma_2 \in \mathcal{C}^1([0,1])$ such that

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

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

$$u(x,t) \approx \frac{\text{the amount of the substance between points } x-h \text{ and } x+h \text{ on the path at time } t}{2h} \tag{7.1}$$

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 $k = 0, \ldots, n-1$ 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).

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

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

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

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

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

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

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

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

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

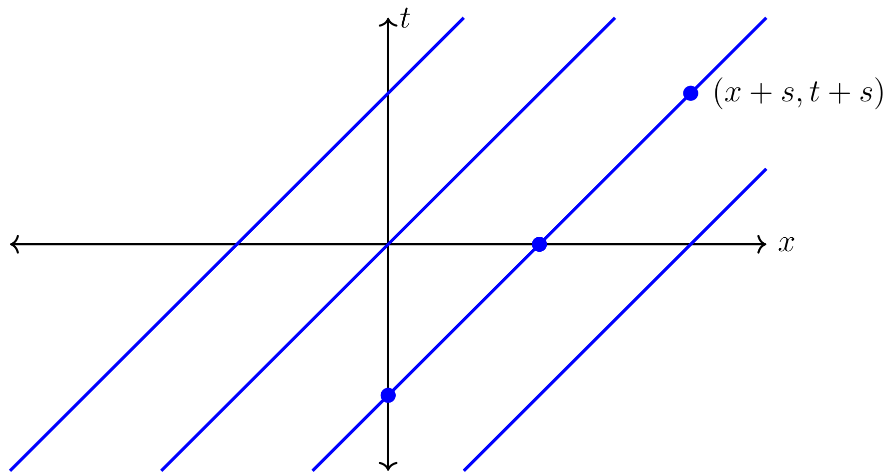
$$\nabla u \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0.$$

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

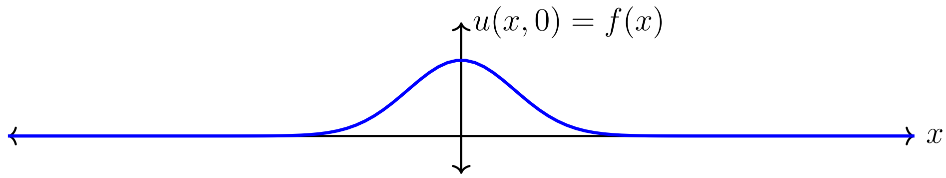
$$u(x,t) = f(x-t).$$

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

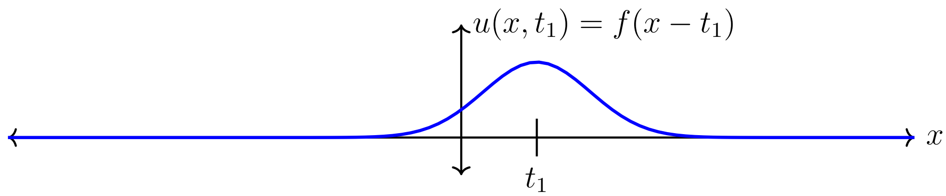
8.2 Problem (!). Check that. Be sure to explain the importance of the regularity requirement $f \in \mathcal{C}^1(\mathbb{R})$.

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.

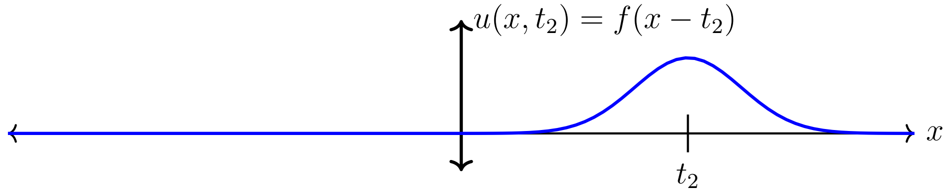
Here is a graph for the initial data f.



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.



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*

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

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

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

$S \in \mathbb{R}$. Then $h \in \mathcal{C}^1(\mathbb{R})$ and

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

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.

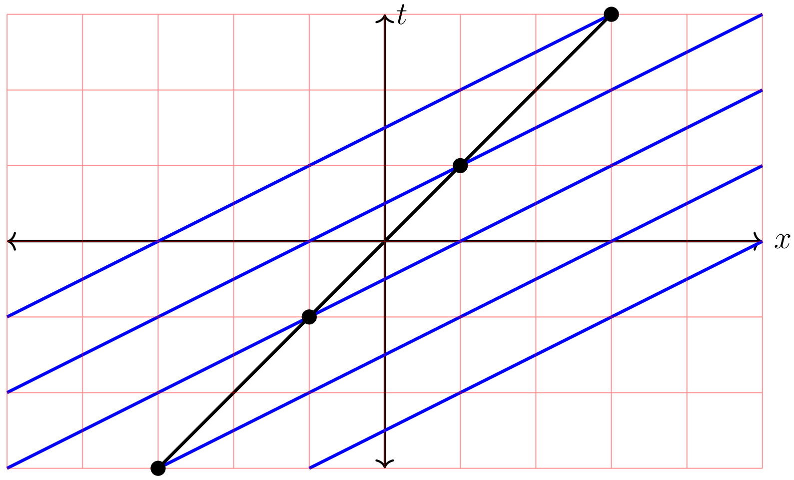
Our work with side conditions has been strictly algebraic; now we consider the interaction of the side condition curve with the geometry of the PDE.

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

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

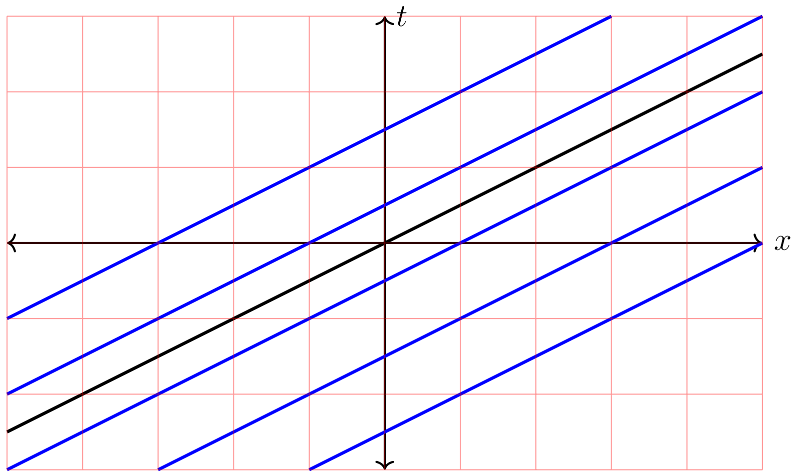
solutions u are constant on lines parallel to $(3,1)$, i.e., lines with slope $1/3$.

(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

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

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

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

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

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

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

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

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

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

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

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

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, α, β, $\gamma \in \mathbb{R}$ 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 $\partial_t\big[\int_0^t f(t,s)\ ds\big]$. 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

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

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

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

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 $a \leq b$. Let $\mathcal{D} := \{(x, s) \in \mathbb{R}^2 \mid x \in I,\ a \leq s \leq b\}$. Suppose that $f \in \mathcal{C}(\mathcal{D})$ and that f_x exists on $\mathcal{D}$ with $f_x \in \mathcal{C}(\mathcal{D})$. Then the map*

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

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

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

(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 $|f_{xx}(x, s)| \leq M$ 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

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

by FTC1 and

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

by Leibniz's rule.

We have proved the following.

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

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

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

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

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

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

$$\mathcal{J}(t) = \int_{x_0-ct_0}^{x-ct_1} f(\tau) \ d\tau.$$

(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 α, $\beta \in \mathbb{R}$. Choose α and β so that U solves $U_{TT} = U_{XX}$ and use the work above (updating the initial conditions as needed) to find a formula for U. From that, develop the formula (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.

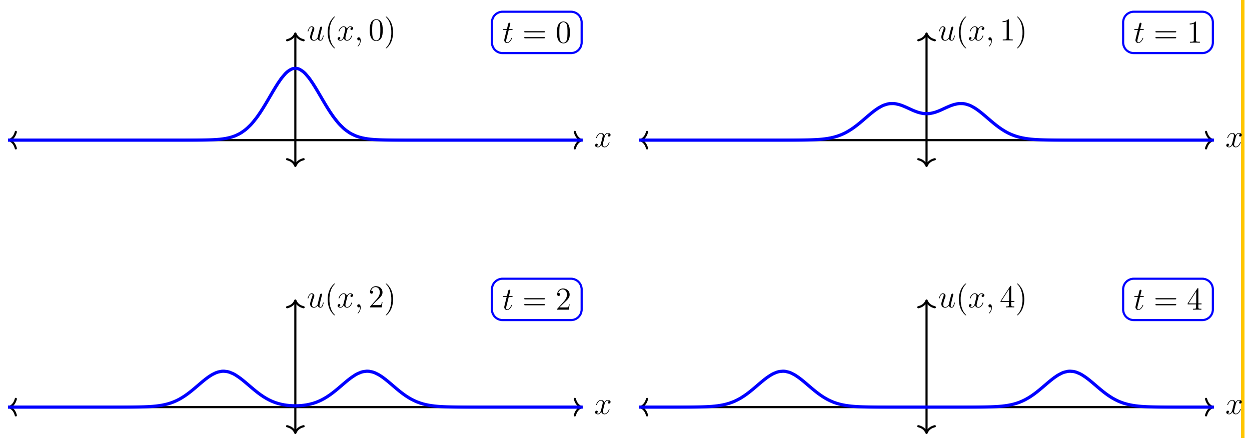
(i) Take

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

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.

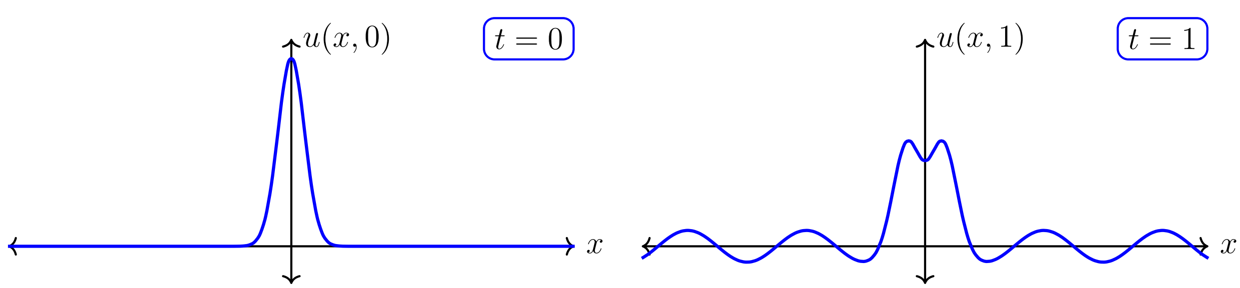
(ii) Take

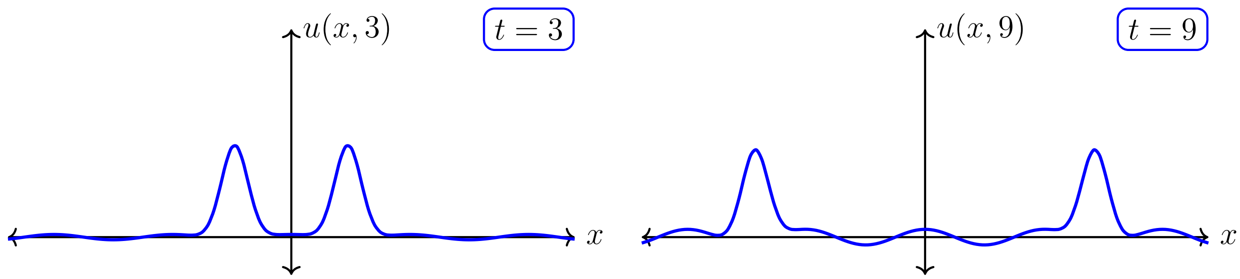
$$f(x) = 10e^{-x^2} \quad \text{and} \quad g(x) = \cos(x).$$

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

$$u(x,t) = \frac{10e^{-(x+t)^2} + 10e^{-(x-t)^2}}{2} + \frac{1}{2}\int_{x-t}^{x+t} \cos(s)\ ds$$
$$= 5\big(e^{-(x+t)^2} + e^{-(x-t)^2}\big) + \frac{\big(\sin(x+t) - \sin(x-t)\big)}{2}.$$

Here are some graphs.





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

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

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

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.

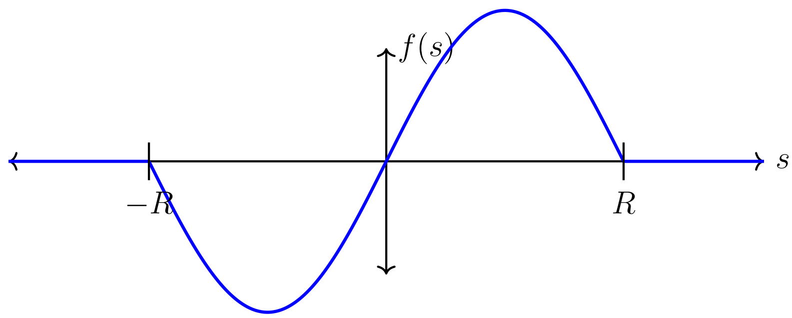
We will continue making traveling wave ansatzes for other PDE that we meet and interpreting those solutions physically and mathematically in the broader context of those equations. Now we return to the wave equation and tease out more properties from D'Alembert's formula.

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

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

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

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

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

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

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

Here is what we have proved.

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:

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

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

Of course, the solution to

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

is

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

and the solution to

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

is

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

The key to everything is rewriting this second solution:

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

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

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

so

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

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

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

2. The nonhomogeneous transport IVP similarly "splits":

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

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

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

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

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

is given by D'Alembert's formula:

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

This morally resembles the first two terms of the solution (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 ($\star$). 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

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

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,

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

and

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

Thus

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

Now we work on the time derivative. We have

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

We compute

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

and, for τ fixed,

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

Then

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

so

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

Certainly

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

while

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

All together,

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

after comparison to (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.,

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

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 ($\star$). 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 ($\star$). 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 $x \in \mathbb{R}$.*

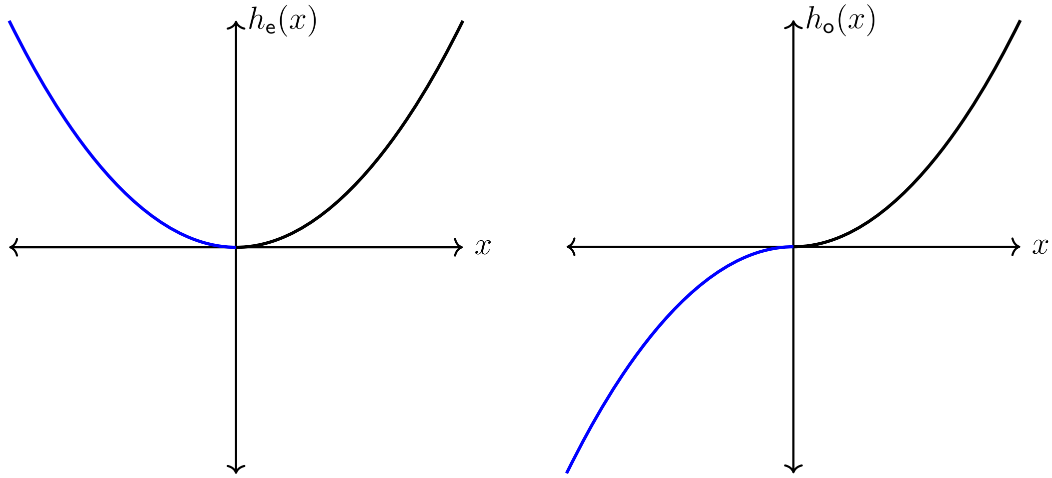
(ii) *The* **EVEN EXTENSION (REFLECTION)** *of a function $h\colon [0,\infty) \to \mathbb{R}$ is the function*

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

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

$$\widetilde{u}(x,t) = \begin{cases} u(x,t), & x \ge 0, \ t \in \mathbb{R} \\ -u(-x,t), & x < 0, \ t \in \mathbb{R}, \end{cases} \qquad f_{\mathsf{o}}(x) = \begin{cases} f(x), & x \ge 0 \\ -f(-x), & x < 0, \end{cases}$$

$$\text{and} \qquad g_{\mathsf{o}}(x) = \begin{cases} g(x), & x \ge 0 \\ -g(-x), & x < 0, \end{cases}$$

show that $f_{\mathsf{o}} \in \mathcal{C}^2([0,\infty))$ and $g_{\mathsf{o}} \in \mathcal{C}^1([0,\infty))$, show that $\widetilde{u} \in \mathcal{C}^2(\mathbb{R}^2)$, and show that

$$\begin{cases} \widetilde{u}_{tt} = \widetilde{u}_{xx}, & -\infty < x, \ t < \infty \\ \widetilde{u}(x,0) = f_{\mathsf{o}}(x), & -\infty < x < \infty \\ \widetilde{u}_t(x,0) = g_{\mathsf{o}}(x), & -\infty < x < \infty. \end{cases}$$

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 $f_{\circ}$ and $g_{\circ}$, 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 $u_{\circ}$, 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* $g_{\circ} \in \mathcal{C}^1(\mathbb{R})$.

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

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

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

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

exist and are equal.

For $h > 0$, we have

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

and so

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

Next, for $h < 0$,

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

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

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

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 $g_\circ$ is differentiable on all of $\mathbb{R}$ with

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

Now we check continuity at $x = 0$:

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

while

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

This finishes the proof of continuity of $g_\circ'$. ■

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 $f \in \mathcal{C}^2([0,\infty))$ with $f(0) = f''(0) = 0$, then $f_\circ \in \mathcal{C}^2(\mathbb{R})$. [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 $\widetilde{u} \in \mathcal{C}^2(\mathbb{R}^2)$ 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

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

with $x \ge 0$ and $t \in \mathbb{R}$. Since $f_\circ \in \mathcal{C}^2(\mathbb{R})$ and $g_\circ \in \mathcal{C}^1(\mathbb{R})$, 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 $0 \le x < \infty$ and $t \in \mathbb{R}$. Rather, why does this formula meet the boundary condition $u(0,t) = 0$ for all t?

We compute

$$f_\circ(0+t) + f_\circ(0-t) = f_\circ(t) + f_\circ(-t) = f_\circ(t) - f_\circ(t) = 0$$

by the oddness of $f_\circ$. Next, we want $\int_{-t}^{t} g_\circ(s)\ ds = 0$. 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).

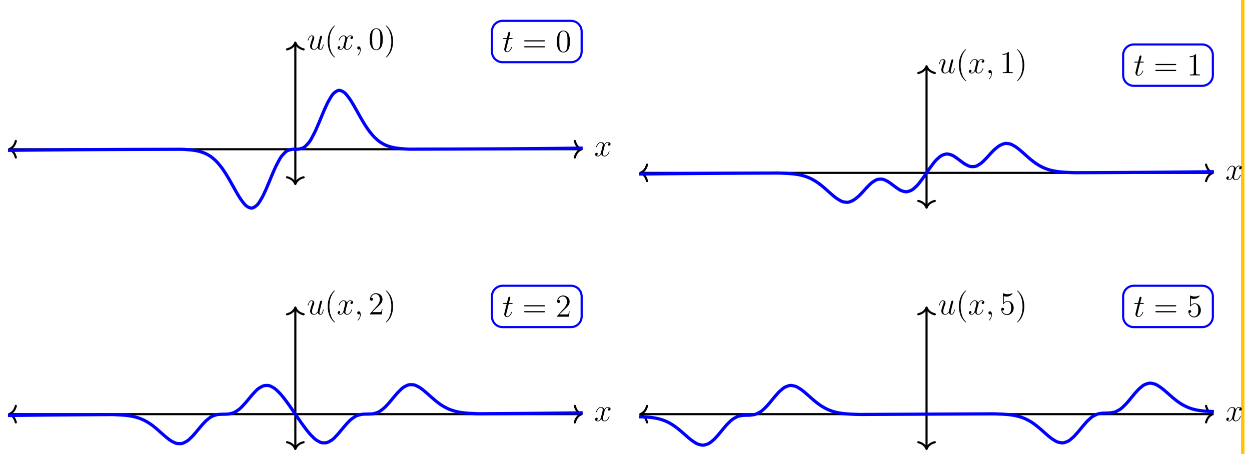
14.9 Example. We solve

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

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 $f_{\circ}^{\text{per}}$ and $g_{\circ}^{\text{per}}$ 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 $\{(x,t) \in \mathbb{R}^2 \mid 0 \le x \le L,\ t \in \mathbb{R}\}$. This forces $f = u(\cdot, 0) \in \mathcal{C}^2([0,L])$ and the same for g.)

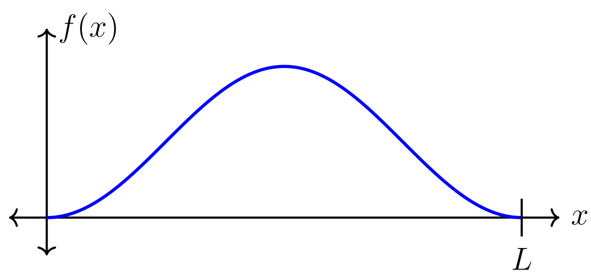
Our success with the semi-infinite string suggests that we extend f and g *carefully* to $\mathbb{R}$ and use D'Alembert's formula. By "carefully," we mean that the extensions should be sufficiently differentiable, and our experience with the semi-infinite problem suggests that this regularity will rely also on knowing some exact values of f and g at 0, and maybe L.

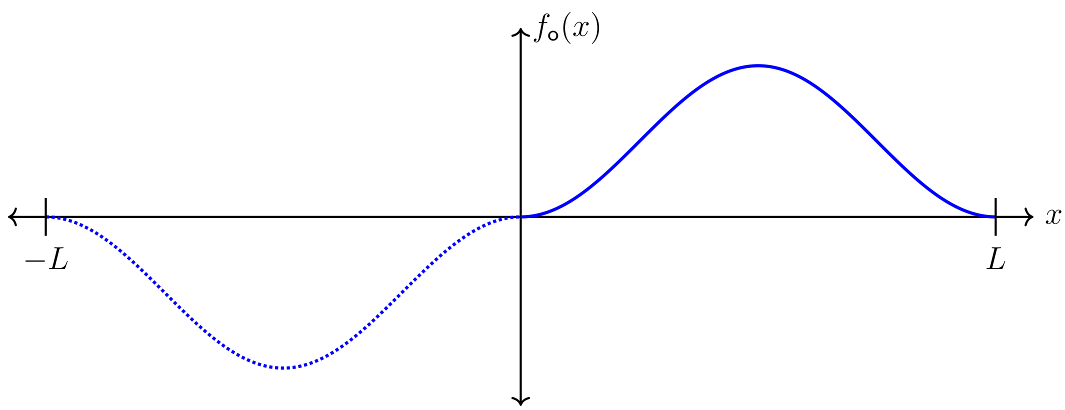
15.1 Problem (!). Prove that if (15.1) has a solution, then the "compatibility conditions"

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

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 $[-L, L]$.

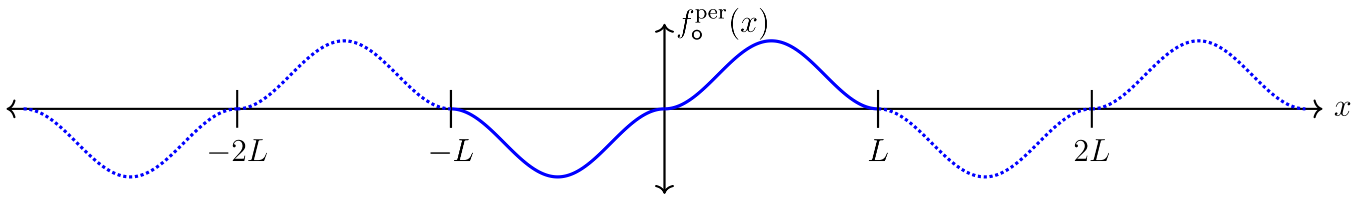




Here is the new idea. We are going to extend the odd extensions from $[-L, L]$ to $(-\infty, \infty)$ *periodically*. Informally, we "copy and paste" the graphs from $[-L, L]$ to the intervals $[(2k+1)L, (2k+3)L]$ for $k \in \mathbb{Z}$. For the initial data, call these periodic extensions $f_\circ^{\text{per}}$ and $g_\circ^{\text{per}}$; we require them to satisfy

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

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 $\widetilde{u}^{\text{per}} \in \mathcal{C}^2(\mathbb{R}^2)$, $f_\circ^{\text{per}} \in \mathcal{C}^2(\mathbb{R})$, and $g_\circ^{\text{per}} \in \mathcal{C}^1(\mathbb{R})$, and if

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

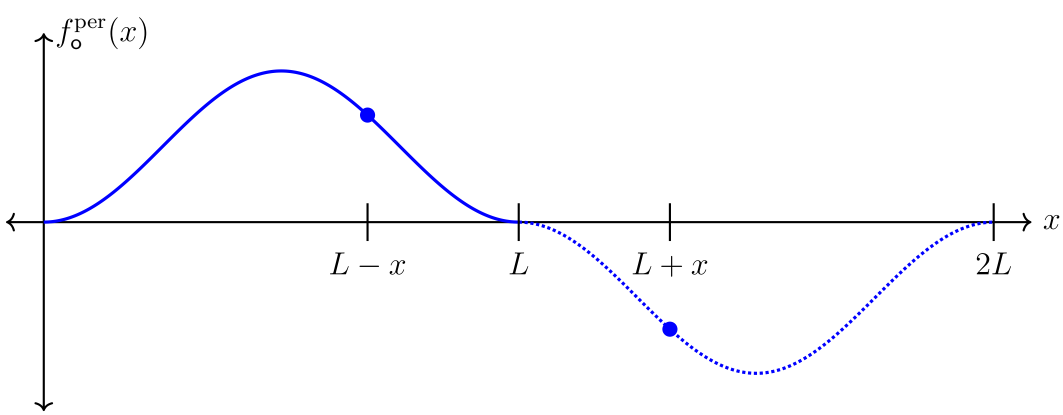
then we can invoke D'Alembert's formula. Conversely, we would need to check that the formula satisfies (15.1), in particular the boundary conditions.

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 $f_\circ^{\text{per}}(L - x) = -f_\circ^{\text{per}}(L + x)$ for $0 \leq x\ leL$. (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 $f(x) \ge 0$ 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

$$\begin{aligned} E'(t) = \partial_t\left[\int_0^L \left[w_t(x,t)^2 + w_x(x,t)^2\right] \, dx\right] &= \int_0^L \partial_t\left[w_t(x,t)^2 + w_x(x,t)^2\right] \, dx \\ &= 2\int_0^L \left[w_t(x,t)w_{tt}(x,t) + w_x(x,t)w_{xt}(x,t)\right] \, dx. \end{aligned}$$

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

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

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.

15.5 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

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

Show that $E_2'(t) = 2\big(u_t(a,t)^2 - u_t(b,t)^2\big)$.

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

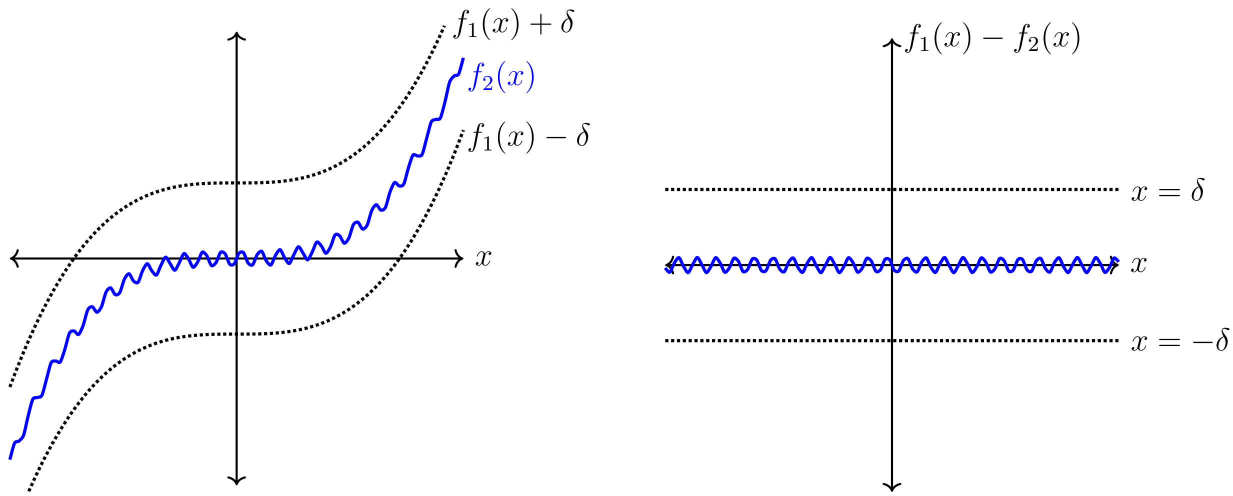
$$\begin{cases} u_{tt} = u_{xx}, \ -\infty < x, \ t < \infty \\ u(x,0) = f_1(x), \ -\infty < x < \infty \\ u_t(x,0) = g_1(x), \ -\infty < x < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, \ -\infty < x, \ t < \infty \\ v(x,0) = f_2(x), \ -\infty < x < \infty \\ v_t(x,0) = g_2(x),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

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 (⋆). Here is a somewhat silly example of how this could occur. Let δ, $\epsilon > 0$. Take $f_1 = g_1 = 0$ and $f_2(x) = \delta/2$ and $g_2(x) = \epsilon/2$. Show that if u and v solve (16.2), then

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

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

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

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.

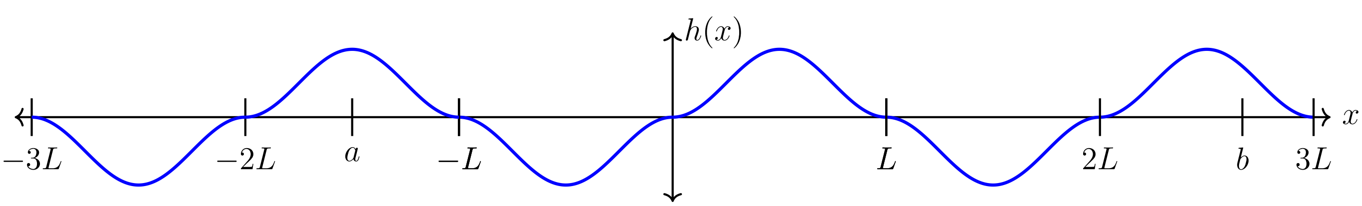
To cut down on writing, we let $h \in \mathcal{C}(\mathbb{R})$ be odd and $2L$-periodic with $|h(x)| < \epsilon$ for all x. We claim that

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

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

$$\begin{aligned}\left|\int_a^b h\right| = \left|\int_a^{(2j+3)L} h + \int_{(2k+3)L}^b h\right| &\le \left|\int_a^{(2j+3)L} h\right| + \left|\int_{(2k+3)L}^b h\right| \le \int_a^{(2j+3)L} |h| + \int_{(2k+3)L}^b |h| \\ &< \int_a^{(2j+3)L} \epsilon + \int_{(2k+3)L}^b \epsilon = \epsilon((2j+3)L - a) + \epsilon(b - (2k+3)L).\end{aligned}$$

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 $\delta,\ \epsilon > 0$ with*

$$|f_1(x) - f_2(x)| < \delta \quad \text{and} \quad |g_1(x) - g_2(x)| < \epsilon$$

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

$$\begin{cases} u_{tt} = u_{xx}, & -\infty < x,\ t < \infty \\ u(x,0) = f_1(x), & -\infty < x < \infty \\ u_t(x,0) = g_1(x), & -\infty < x < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, & -\infty < x,\ t < \infty \\ v(x,0) = f_2(x), & -\infty < x < \infty \\ v_t(x,0) = g_2(x), & -\infty < x < \infty. \end{cases} \tag{18.6}$$

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 $\delta,\ \epsilon > 0$ with

$$|f_1(x) - f_2(x)| < \delta \quad \text{and} \quad |g_1(x) - g_2(x)| < \epsilon$$

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

$$\begin{cases} u_{tt} = u_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ u(x,0) = f_1(x), \ 0 \le x \le L \\ u_t(x,0) = g_1(x), \ 0 \le x \le L \\ u(0,t) = u(L,t) = 0, \ -\infty < t < \infty \end{cases} \quad \text{and} \quad \begin{cases} v_{tt} = v_{xx}, \ 0 \le x \le L, \ -\infty < t < \infty \\ v(x,0) = f_2(x), \ 0 \le x \le L \\ v_t(x,0) = g_2(x), \ 0 \le x \le L \\ v(0,t) = v(L,t) = 0, \ -\infty < t < \infty. \end{cases} \tag{18.7}$$

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 $t \ge 0$, 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 $\operatorname{Re}(z) := x$; the* **IMAGINARY PART** *of z is $\operatorname{Im}(z) := y$; and the* **MODULUS** *of z is $|z| := \sqrt{\operatorname{Re}(z)^2 + \operatorname{Im}(z)^2}$. 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 $\operatorname{Re}(z) = \operatorname{Re}(w)$ and $\operatorname{Im}(z) = \operatorname{Im}(w)$.

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

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

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*

$$f'(t) := f_1'(t) + if_2'(t);$$

(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 $e^{2\pi ik} = 1$ 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*

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

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

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

so $\int_{-\infty}^0 f = \int_0^\infty f = 1$. Thus f is integrable and

$$\int_{-\infty}^\infty f = 1 + 1 = 2.$$

It is often both difficult to establish that f is integrable and unnecessary to calculate $\int_{-\infty}^\infty f$ exactly. Instead, the following tests usually suffice. 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:*

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

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

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

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 $e^{ax-ax^2} \le C$ for $0 \le x \le 1$. Then argue that $e^{-ax^2} \le e^{-ax}$ for $x \ge 1$. Put these estimates together to show $e^{-ax^2} \le (C+1)e^{-ax}$ for $x \ge 0$.*]

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 ($g(x_1) \le g(x_2)$ for $0 \le x_1 \le x_2$), and bounded above (there is $M > 0$ such that $0 \le g(x) \le M$ for all $x \ge 0$), then $\lim_{x\to\infty} g(x)$ exists. The proof of this result depends on the completeness of the real numbers, but drawing a picture probably suggests why it is true. Draw such a picture. Then use this result to show that if $f\colon \mathbb{R} \to [0,\infty)$ is continuous, and if there is $M > 0$ such that $\left|\int_{-R}^{R} f\right| \le M$ for all $R \ge 0$, then f is integrable. [Hint: *apply the result to the functions* $R \mapsto \int_0^R f$ *and* $R \mapsto \int_{-R}^0 f$.]

(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:

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

To avoid confusion, we will not write this as $\widehat{u}_t(k,t)$. All together, we expect that a "nice" solution u to the heat equation with "nice" initial data f will satisfy

$$\begin{cases} \partial_t[\widehat{u}](k,t) = -k^2\widehat{u}(k,t) \\ \widehat{u}(k,0) = \widehat{f}(k). \end{cases}$$

This is really a family of IVP at the ODE level parametrized in $k \in \mathbb{R}$. (We posed the heat equation only for $t > 0$, but we can solve this IVP for all t, so we might as well consider all t here.) The notation may be burdensome, but all this is asking us to do is solve

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

for each $k \in \mathbb{R}$. Certainly we know how to do that: $y(t) = \widehat{f}(k)e^{-k^2 t}$. And so $\widehat{u}$ should satisfy

$$\widehat{u}(k,t) = \widehat{f}(k)e^{-k^2 t}.$$

Now we can recover u from $\widehat{u}$ by Untheorem 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 $|\widehat{f}(k)| \le M$ 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

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

and recover the expected, beloved formula $u(x,t) = f(x-t)$. [Hint: *apply the Fourier transform to u in x and get an ODE-type IVP for $\widehat{u}$. Solve it. Then recover u from its Fourier transform via Untheorem 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 $t \le 0$? This is one mathematical reason to take $t > 0$ in our statement of the heat equation—we do it because it leads to a problem that we can solve!)

We first replace $\widehat{f}(k)$ in (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.

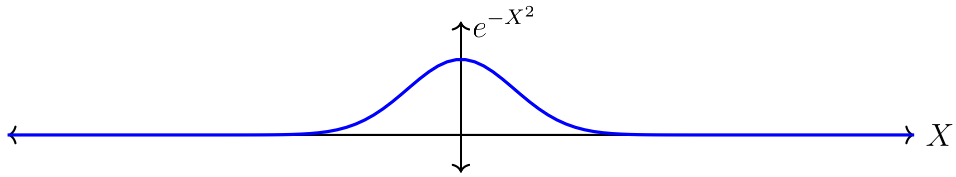
We start by cleverly rewriting H:

$$H(s,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-k^2 t} e^{iks} \, dk = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(k\sqrt{t})^2} e^{-i(-s)k} \, dk \right).$$

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

affects the limits of integration.]

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

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

is integrable with

$$\int_{-\infty}^{\infty} g(x+d)\ dx = \int_{-\infty}^{\infty} g(u)\ du \quad \text{and} \quad \widehat{S^d g}(k) = e^{ikd}\widehat{g}(k).$$

[Hint: *for integrability, it may be easier to prove that the limits in Definition 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 $\widehat{\mathcal{G}}$?