MATH 4581
Classical Mathematical Methods
in Engineering
Fall Semester 2024

John McCuan

August 18, 2024

Contents

Preface

These notes accompany my lectures for MATH 4581 Classical Mathematical Methods of Engineering at Georgia Tech given in the fall semester of 2024.

Part I

Ordinary Differential Equations

Lecture 1

Spatial Ordinary Differential Equations

1.1 Introduction/Outline

This course is primarily intended as an introduction to the three main partial differential equations (PDE) of classical mathematical analysis, namely Laplace's equation, the heat equation, and the wave equation. This introduction is given in the context of separation of variables and Fourier series expansions. I will start with a short review of ordinary differential equations (ODE) and then discuss Fourier series expansion in isolation. Then we will consider various problems for the three equations above. The main text is *Applied Partial Differential Equations* (fourth edition) by Richard Haberman.

I will provide some notes for the first couple lectures. After that, most of the reading required to complete the assignments can be found in Haberman's book.

Assignments will be due approximately every week starting with the second week. Here is an outline:

Part 1 Spatial Ordinary Differential Equations

- **Lecture 1** Introduction:
  - **1** The Initial Value Problem
  - **2** The two point boundary value problem
- **Lecture 2** More on ODE:

1 Series solutions

2 Fourier series solutions

Part 2 Fourier Series

Lecture 3 Linear spaces of functions; norms

Lecture 4 Convergence of Fourier series

Lecture 5 More on Fourier series

Part 3 The Heat Equation

Lecture 6 Introduction/derivation

Lecture 7 Separation of variables

Lecture 8 More on the heat equation

Part 4 Laplace's Equation

Lecture 9 Introduction/derivation

Lecture 10 Properties: Mean value property and maximum principle

Lecture 11 More on Laplace's equation

Part 5 The Wave Equation

Lecture 12 Introduction:

Lecture 13

1.2 Initial Value Problems

Theorem 1. (general local existence and uniqueness) If

$$\mathbf{F} \in C^1(\mathbb{R}^n \times (a,b) \to \mathbb{R}^n),$$

then for any $\mathbf{p} \in \mathbb{R}^n$ and any $t_0 \in (a,b)$ there exists some $\epsilon > 0$ such that the initial value problem (IVP)

$$\begin{cases} \mathbf{x}' = \mathbf{F}(\mathbf{x}, t) & t_0 - \epsilon < t < t_0 + \epsilon \\ \mathbf{x}(t_0) = \mathbf{p} \end{cases} \tag{1.1}$$

has a unique solution.

This is called the **general existence and uniqueness** theorem for ODEs. There is no similar theorem for partial differential equations (PDE) which is one of the main things you should learn/understand by the end of this course. Of course, to appreciate an assertion such as this, you need to have some appreciation for this theorem.

1.2.1 Discussion

In principle courses in calculus, ordinary differential equations (ODE), and linear algebra are natural prerequisites for this course. If you didn't really learn these subjects yet, it may be difficult to get a good appreciation for partial differential equations. I've learned from experience, however, that many if not most students didn't really become experts in all the details of these prerequisite courses, so if you're not so familiar with some topics in calculus, ODE and linear algebra, don't worry too much. Maybe you will pick up the details along the way. I won't really review calculus and linear algebra explicitly, but obviously I'm planning to make some comments concerning ODEs. This is primarily with the idea of making some kind of comparison to PDE.

You may not have noticed (or thought too hard about) Theorem 1. Perhaps now is a good time. Take a look back at the ingredients, and I'll try to go over some of them.

(a,b) is an **open interval** in the real number line $\mathbb{R}$:

$$(a,b) = \{x \in \mathbb{R} : a < x < b\}.$$

The numbers a and b might be real numbers, but they might also be **extended real numbers**. That is,

$$a, b \in \mathbb{R} \cup \{\pm\infty\}.$$

The set $\mathbb{R} \cup \{\pm\infty\}$ has the nice property that it is ordered, so the condtion $a < b$ makes sense. Addition does not always make sense in this set of extended real numbers, however there are other sets of extended real numbers where addition does make sense, namely

$$\mathbb{R} \cup \{-\infty\} \qquad \text{and} \qquad \mathbb{R} \cup \{\infty\}.$$

If we add extended real numbers, it means we are restricting attention to one of these latter sets.

$\mathbb{R}^n$ is called **Euclidean space**:

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, x_2, \ldots, x_n) : x_j \in \mathbb{R},\ j = 1, 2, \ldots, n\}.$$

$\mathbb{R}^n$ (including $\mathbb{R}^1$ which we identify with just $\mathbb{R}$) is a **vector space** over $\mathbb{R}$. (Vector spaces are a big topic of discussion in linear algebra, and they will be of some importance for us too.)

1.2.2 $\mathbf{F} : \mathbb{R}^n \times (a, b) \to \mathbb{R}^n$

Notice that

$$\mathbb{R}^n \times (a, b) = \{(\mathbf{x}, t) : \mathbf{x} \in \mathbb{R}^n \text{ and } t \in (a, b)\}$$

may be thought of as a subset of $\mathbb{R}^{n+1}$. In fact, $\mathbb{R}^n \times (a, b)$ is essentially an **open set** in $\mathbb{R}^{n+1}$. When we write

$$\mathbf{F} : \mathbb{R}^n \times (a, b) \to \mathbb{R}^n$$

we mean the open set $\mathbb{R}^n \times (a, b)$ is the domain of a **vector valued function** $\mathbf{F}$ of several variables. That is,

$$\mathbf{F} = (f_1, f_2, \ldots, f_n)$$

where each $f_j = f_j(\mathbf{x}, t)$ for $j = 1, 2, \ldots, n$ is a **real valued function of several** (speciically $n + 1$) **variables**. Such functions should be somewhat familiar from multivariable calculus. In fact, the real valued function of several variables may be said to be the basic object of study in multivariable calculus. Question: What is the basic object of study in ordinary differential equations? Linear algebra?

Such a function has, or at least *may* have, **partial derivatives**. We will of course have a lot to say about partial derivatives when we consider **partial differential equations**. For now, here is a definition: Given $f : U \to \mathbb{R}$ where U is an open subset of $\mathbb{R}^n$ (for some natural number n)

$$\frac{\partial f}{\partial x_j} = \frac{\partial f}{\partial x_j}(\mathbf{p}) = \lim_{v \to 0} \frac{f(\mathbf{p} + v\mathbf{e}_j) - f(\mathbf{p})}{v} \tag{1.2}$$

for $j = 1, 2, \ldots, n$ where

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, 0, \ldots, 0) \in \mathbb{R}^n \\ \mathbf{e}_1 &= (0, 1, 0, \ldots, 0) \in \mathbb{R}^n \\ &\vdots \\ \mathbf{e}_n &= (0, \ldots, 0, 0, 1) \in \mathbb{R}^n. \end{aligned}$$

Question: How many (first) partial derivatives does a coordinate function f_j of the vector valued function $\mathbf{F}$ in Theorem 1 (probably) have? (Answer: $n + 1$.)

Given $f : U \to \mathbb{R}$ with U an open subset of $\mathbb{R}^n$,

$$\frac{\partial f}{\partial x_j}$$

is called the j-th **partial derivative** of f or the **derivative in the** x_j **direction** IF THE LIMIT IN (1.2) EXISTS.

The question of wether or not a derivative exists is an example of a **question of regularity**. In a certain sense a more fundamental starting place[1] for the discussion of regularity is continuity:

Definition 1. (continuity at a point) Given an open set $U \subset \mathbb{R}^n$ and a real valued function $f : U \to \mathbb{R}$ we say f is **continuous at $\mathbf{p} \in U$** if for any $\epsilon > 0$, there is some $\delta > 0$ so that

$$|f(\mathbf{x}) - f(\mathbf{p})| < \epsilon \tag{1.3}$$

whenever

$$|\mathbf{x} - \mathbf{p}| = \sqrt{\sum_{j=1}^{n} (x_j - p_j)^2} < \delta. \tag{1.4}$$

[1] than differentiability

The quantity $|\mathbf{x}-\mathbf{p}|$ appearing in (1.4) is not just simple absolute values like the quantity on the left in (1.3). This expression is called the **Euclidean distance** from $\mathbf{x}$ to $\mathbf{p}$. The Euclidean distance is constructed using the **Euclidean norm**

$$| \cdot | : \mathbb{R}^n \to [0,\infty) \qquad \text{by} \qquad |\mathbf{x}| = \sqrt{\sum_{j=1}^n x_j^2}.$$

The general notions of distance and norm will be important for us later.

Definition 2. (continuity on a set) Given an open set $U \subset \mathbb{R}^n$ and a real valued function $f : U \to \mathbb{R}$ we say f is **continuous on** U if f is continuous at each point $\mathbf{p} \in U$. In this case we write $f \in C^0(U)$.

Note carefully, $C^0(U)$ denotes the collection of all continuous real valued functions with domain U.

If $f, g \in C^0(U)$, then

(i) $f + g \in C^0(U)$ and

(ii) $cf \in C^0(U)$ for each $c \in \mathbb{R}$.

These are the two main properties that make $C^0(U)$ a vector space (which it is).

The subspace of $C^0(U)$ consisisting of functions $f : U \to \mathbb{R}$ each of whose partial derivatives

$$\frac{\partial f}{\partial x_j}, \quad j = 1, 2, \ldots, n$$

exist and satisfy

$$\frac{\partial f}{\partial x_j} \in C^0(U)$$

is denoted $C^1(U)$. The elements in $C^1(U)$ are called the **continuously differentiable functions**.

When you see $\mathbf{F} \in C^1(\mathbb{R}^n \times (a,b) \to \mathbb{R}^n)$ in the statement of Theorem 1 it means that each coordinate function f_j in

$$\mathbf{F} = (f_1, f_2, \ldots, f_n)$$

for $j = 1, 2, \ldots, n$ satisfies

$$f_j \in C^1(\mathbb{R}^n \times (a,b)).$$

Looking back at the statement of Theorem 1 we can observe at this point that quite a lot of information is packed into the hypothesis

$$\mathbf{F} \in C^1(\mathbb{R}^n \times (a,b) \to \mathbb{R}^n).$$

Quite a number of concepts are also required to understand exactly what this hypothesis means. If it is any consolation this is the only hopothesis in the theorem. If this one condition holds, then given any

$$(\mathbf{p}, t_0) \in \mathbb{R}^n \times (a,b)$$

there exists some $\epsilon > 0$ and a unique function

$$\mathbf{x} \in C^1((t_0 - \epsilon, t_0 + \epsilon) \to \mathbb{R}^n)$$

satisfying the IVP (1.1).

1.2.3 application(s)

Note carefully the use of $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ as an independent variable in the discussion of regularity above is quite different from the use of $\mathbf{x} = (x_1, x_2, \dots, x_n) \in C^1((t_0 - \epsilon, t_0 + \epsilon) \to \mathbb{R}^n)$ as a vector valued function in the conclusion of Theorem 1. In the latter case each coordinate x_j for $j = 1, 2, \dots, n$ is a real valued function of one real variable.

Example. Find $\mathbf{x} \in C^1(\mathbb{R} \to \mathbb{R}^2)$ for which

$$\begin{cases} \mathbf{x}' = \mathbf{x} + \begin{pmatrix} t \\ t^2 \end{pmatrix} & t \in \mathbb{R} \\ \mathbf{x}(0) = \begin{pmatrix} 3 \\ 4 \end{pmatrix}. \end{cases}$$

Question: If one is to apply Theorem 1 to this IVP, what is t_0?

Solution:

$$\mathbf{x}(t) = \begin{pmatrix} 3e^t + (1/2)t^2 \\ 4e^t + (1/3)t^2 \end{pmatrix}.$$

Question: What is $\mathbf{F}$ in this example?

Example.

$$\begin{cases} y'' = y^2 \\ y(t_0) = y_0. \end{cases}$$

From the point of view of Theorem 1 we should consider the problem as follows: Set $x_1 = y$ and $x_2 = x_1' = y'$. Then

$$\begin{cases} x_1' = x_2 \\ x_2' = x_1^2 \end{cases} \qquad \text{or} \qquad \mathbf{x}' = \mathbf{F}(\mathbf{x})$$

where

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1^2 \end{pmatrix}.$$

Or $f_1(x_1, x_2) = x_2$ and $f_2(x_{1,2}) = x_1^2$.

1.3 Linear ODE

There is an existence and uniqueness theorem for linear ODE with a stronger conclusion:

Theorem 2. (existence and uniqueness theorem for linear ODE) Let $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a < b$. If $a_{ij}, b_j \in C^0(a, b)$ for $i, j = 1, 2, \ldots, n$, then for every $(\mathbf{p}, t_0) \in \mathbb{R}^n \times (a, b)$ the IVP

$$\begin{cases} \mathbf{x}' = A\mathbf{x} + \mathbf{b}, \quad t \in (a, b) \\ \mathbf{x}(t_0) = \mathbf{p}, \end{cases} \tag{1.5}$$

where $A \in C^0((a, b) \to \mathbb{R}^{n \times n})$ is the $n \times n$ matrix malued function with the real valued function a_{ij} in the i-th row and j-th column and $\mathbf{b} \in C^0((a, b) \to \mathbb{R}^n)$ is the vector valued function with j-th component function b_j, has a unique solution $\mathbf{x} \in C^1((a, b) \to \mathbb{R}^n)$.

1.4 The two point boundary value problem

In the study of ODEs one usually focuses on the initial value problem for ODEs. This is largely due to the sweeping assertion of Theorem 1. In order to consider ODEs in a manner more comparable to the point of view taken in the study of PDE we now formulate a different kind of problem. This problem is called the two point boundary value problem (BVP): Let $a, b, x_0 \in \mathbb{R}$ be given with $a < x_0 < b$. Given $p, q, f \in C^0[a, b]$ and $y_a, y_b \in \mathbb{R}$, find $y \in C^1[a,b]$ satisfying

$$\begin{cases} y'' + p(x)\ y' + q(x)\ y = f(x), & x \in (a,b) \\ y(a) = y_a, \\ y(b) = y_b. \end{cases} \tag{1.6}$$

Here is a theorem on the existence and uniqueness of solutions for a two point BVP for ODEs:

Theorem 3. (existence and uniqueness theorem for a BVP) Given $L > 0$ and $g \in C^0[0,L]$, the problem

$$\begin{cases} u'' = g(x), & x \in (0,L) \\ u(0) = 0, \\ u(L) = 0. \end{cases} \tag{1.7}$$

has a unique solution $u \in C^1[0,L]$.

You can prove this theorem. It is interesting that not all two point boundary value problems have nonzero solutions and some do not have unique solutions: The ODE $y'' + y = 0$ has general solution $y = a\cos x + b\sin x$ where $a, b \in \mathbb{R}$. Thus, if we require $y(0) = 0$ we know $a = 0$. Then if we require also the second homogeneous boundary condition $y(L) = 0$ at some $L > 0$, then either $L = k\pi$ for some $k \in \mathbb{N}$ and $y = b\sin x$ is a solution for every $b \in \mathbb{R}$. If on the other hand $L \in \{x > 0 : x \neq k\pi,\ k = 1,2,3,\ldots\}$, then $\sin(L) \neq 0$, and we must have also $b = 0$, so $y \equiv 0$ is the only solution. Of course, it's nice to have a unique solution in this case, but it is not a very interesting solution.

1.5 Solving an Easy ODE

If one is interested in classical mathematical methods in engineering, then one is interested in partial differential equations, and one of the first partial differential equations he should consider is

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x)$$

which is a form of the heat equation. This equation is usually accompanied by an initial condition

$$u(x,0) = u_0(x)$$

where $u = u(x,t)$ is a function of two real variables, the first called a **spatial variable**, the latter variable t is used to model time, and $u(x,t)$ may be thought of as giving the temperature at time t at position x with $0 \le x \le L$ where the interval $[0, L]$ models a thin heat conducting rod. We will discuss the modeling and this equation in much more detail later, but a few simple observations now may be, shall we say, motivating.

First of all, the equation also usually comes along with **boundary conditions** which are of primary interest for the discussion to follow. One possibility is to have each end of the rod with a prescribed fixed temperature. These are called **fixed endpoint boundary conditions**. For example, if both endpoints are fixed at zero temperature, then the boundary conditions for this partial differential equations take the form

$$u(0,t) = 0 = u(L,t) \qquad \text{for} \qquad t \ge 0.$$

Putting all this together, we get an **initial/boundary value problem** of the form

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + f(x), & 0 < x < L,\ t > 0 \\ \\ u(x,0) = u_0(x), & 0 \le x \le L \\ u(0,t) = 0 = u(L,t), & t \ge 0. \end{cases} \tag{1.8}$$

The functions u_0 and f as well as the positive number L are given. As with all differential equations, the basic problem is to find the unknown function, in this case u. This can be a challenging task, but at the moment it might be interesting to know some special particular solutions. The number L, as just mentioned is positive. This means the function

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

is well-defined and satisfies the boundary condition in (1.8). If we take

$$u_0(x) = u(x,t) = \frac{L^2}{\pi^2} \sin\left(\frac{\pi}{L} x\right)$$

for $0 \le x \le L$, then the initial condition in (1.8) is satisfied as well. Finally, it will be observed that the function u is **independent of** t. Therefore,

$$\frac{\partial u}{\partial t} \equiv 0,$$

and if we take (or have)

$$f(x) = \sin\left(\frac{\pi}{L}x\right)$$

then we have a solution to a version of (1.8) in the form

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + \sin\left(\dfrac{\pi}{L}x\right), & 0 < x < L,\ t > 0 \\ u(x,0) = \dfrac{L^2}{\pi^2}\sin\left(\dfrac{\pi}{L}x\right), & 0 \le x \le L \\ u(0,t) = 0 = u(L,t), & t \ge 0. \end{cases} \tag{1.9}$$

A solution like this one, which is independent of t, is called an **equilibrium solution**.

In order to illustrate a behavior that you might guess about solutions of the heat equation (after thinking about it for a while) let's consider a modified version of (1.9):

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + \sin\left(\dfrac{\pi}{L}x\right), & 0 < x < L,\ t > 0 \\ u(x,0) = 0, & 0 \le x \le L \\ u(0,t) = 0 = u(L,t), & t \ge 0. \end{cases} \tag{1.10}$$

If we multiply the old solution

$$u_{\text{old}}(x) = \frac{L^2}{\pi^2}\sin\left(\frac{\pi}{L}x\right)$$

by a function $\phi = \phi(t)$ depending on ϕ alone, then we get the time dependent function

$$u(x,t) = \phi(t)\ u_{\text{old}}(x).$$

Notice the boundary condition in (1.10) is satisfied by this new independent function. Also, the initial condition will be satisfied if $\phi(0) = 0$. The partial derivatives are reasonably easy to calculate:

$$\frac{\partial u}{\partial t} = \frac{d\phi}{dt}\ u_{\text{old}}$$

and

$$\frac{\partial^2 u}{\partial x^2} = \phi\ \frac{d^2}{dx^2}u_{\text{old}} = \phi\ \left(-\frac{\pi^2}{L^2}\right)u_{\text{old}}.$$

Consequently, the PDE reduces to

$$\frac{d\phi}{dt}\, u_{\text{old}} = \phi \left(-\frac{\pi^2}{L^2}\right)\, u_{\text{old}} + \frac{\pi^2}{L^2}\, u_{\text{old}}.$$

This of course holds at any point (x,t) for which $u_{\text{old}}(x) = 0$. At points (x,t) where $u_{\text{old}}(x) \neq 0$, namely when $0 < x < L$, the PDE reduces to the condition

$$\frac{d\phi}{dt} = \left(-\frac{\pi^2}{L^2}\right)\, \phi + \frac{\pi^2}{L^2}$$

which is an ordinary differential equation (ODE). Since this lecture is supposed to be about ODEs, one might think we are making progress, and I think we are. Unfortunately, this is not really the ODE which is supposed to be the main star of the lecture. At any rate, I assume if you are taking this course, you've seen an ODE like this one before and can (at least with a little review) solve it. When you took a course on ODEs before, it is almost certainly the case that the emphasis was on something called the **initial value problem** (IVP) and that is precisely what we have here for the function ϕ:

$$\begin{cases} L^2\, \phi' + \pi^2\, \phi = \pi^2, & t \geq 0 \\ \phi(0) = 0. \end{cases} \tag{1.11}$$

Initial value problems are very natural to consider for ODEs of all sorts. I will discuss in more detail later about why this is the case. One consequence, however, is that all ODEs considered from this point of view are being considered as if the independent variable is some kind of (or something like) **time** even if the independent variable is not time at all. Thus, the preponderance of material in an elementary course on ODEs is about what might be called the **theory of temporal ODEs**.

The ODE in this initial value problem is **separable**, which means we can solve the problem as follows: Notice that for t close to 0 we know $\phi(t) \neq 1$. (Why?) Therefore, it makes sense to write the ODE as

$$\frac{1}{1-\phi}\, \phi' = \frac{\pi^2}{L^2}$$

and integrate both sides from $t = 0$ to some particular positive t. The result is

$$\int_0^t \frac{1}{1-\phi(\tau)}\, \phi'(\tau)\, d\tau = \int_0^t \frac{\pi^2}{L^2}\, d\tau = \frac{\pi^1}{L^2}\, t.$$

Assuming $\phi' \neq 0$, we can change variables setting $w = \phi(\tau)$ so that $dw = \phi'(\tau)\, d\tau$ and the integral on the left becomes

$$\int_0^\phi \frac{1}{1-w}\, dw = -\ln(1-\phi). \tag{1.12}$$

Thus, we find $\ln(1-\phi) = -\pi^2 t/L^2$ and

$$\phi(t) = 1 - e^{-\frac{\pi^2}{L^2}\, t}.$$

Notice that now, in retrospect, we have $\phi'(t) = -\pi^2 e^{-\pi^2 t/L^2} < 0$ for all t and $0 < \phi(t) < 1$ for all $t > 0$ as well. We have also solved the intial/boundary value problem (1.10) for the heat equation. This solution is more interesting.

There are other ways to solve the ODE for the time dependent factor $\phi = \phi(t)$. I can think of at least two other ways, and I've put them in the exercises and problems below for your reviewing pleasure. What is more important to me right now is that we have a solution

$$u(x,t) = \left(1 - e^{-\frac{\pi^2}{L^2}\, t}\right)\, \sin\left(\frac{\pi}{L}x\right),$$

and

$$\lim_{t \nearrow \infty} u(x,t) = u_{\text{old}}(x).$$

This kind of behavior should be typical for solutions of (1.8). Specifically, given a solution $u = u(x,t)$ of

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + f(x), & 0 < x < L,\ t > 0 \\ \\ u(x,0) = u_0(x), & 0 \le x \le L \\ u(0,t) = 0 = u(L,t), & t \ge 0 \end{cases} \tag{1.13}$$

we expect there exists some limiting temperature distribution

$$u_{\text{old}}(x) = \lim_{t \nearrow \infty} u(x,t).$$

The function u_{old} is independent of time. On the other hand, u_{old} is obtained as the limit of solutions, so we should expect whatever properties the solutions have should, as far as they can, carry over to the limiting function. In particular, though the initial condition doesn't make sense for or give any useful information about the limit, the boundary condition and the PDE should (or at least might) apply to u_{old}. If this is the case, it is natural to consider the **two point boundary value problem**

$$\begin{cases} u_{\text{old}}'' = -f, & 0 < x < L \\ u_{\text{old}}(0) = 0 = u_{\text{old}}(L). \end{cases} \tag{1.14}$$

This kind of problem is the star of today's lecture, and this should be something that is a bit different from anything you learned about in your first course on ODE. This is a **spatial ODE**.

The ODE $\phi'' = -f$ featured in (1.14) is a very simple ODE; hence the name of this section. The equation is a linear equation of second order, and it's better than linear; it is a **linear second order ODE with constant coefficients**. You may recall that this class of ODEs (which are the ones you probably learned to solve using Laplace transforms) involve second order linear ordinary differential operators of the form

$$F[\phi] = \phi'' + p\phi' + q\phi.$$

The equation here is so simple that the coefficients p and q are both constant zero. The equation is not homogeneous, but still it is very easy to solve (in general). Here is a first solution: Integrate both sides of the equation from 0 to x to obtain

$$\phi' - \phi'(0) = -\int_0^x f(\xi)\, d\xi.$$

That is,

$$\phi'(x) = \phi'(0) - \int_0^x f(\xi)\, d\xi. \tag{1.15}$$

The function

$$g(x) = \phi'(0) - \int_0^x f(\xi)\, d\xi$$

appearing on the right side here is a perfectly good differentiable function. In particular, the equation $\phi' = g(x)$ with which we are now faced is sort of the first order equivalent of the ODE with which we started; this is also a very simple ODE:

$$\phi(x) = \phi(0) + \int_0^x g(\tilde{\xi})\, d\tilde{\xi} = \phi(0) + \phi'(0)\, x - \int_0^x \int_0^{\tilde{\xi}} f(\xi)\, d\xi\, d\tilde{\xi}.$$

This looks a little complicated, but it is the general solution. You may remember various sorts of ODEs from when you took an introductory course. There were separable ODE and linear ODE. ODEs of the form $\phi'' = f$ and $\phi' = g$ where you simply have a derivitive specified by a given function are inhomogeneous linear ODE which I like to call FTC equations (after the fundamental theorem of calculus). You probably remember the version of the fundamental theorem of calculus which says

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

There is also a second version which says that given a function $f \in C^0(x_0 - r, x_0 + r)$ for some real numbers x_0 and r with $r > 0$, the function $g : (x_0 - r, x_0 + r) \to \mathbb{R}$ by

$$g(x) = \int_{x_0}^x f(\xi)\,d\xi$$

is continuously differentiable with $g'(x) = f(x)$. This second version of the fundamental theorem of calculus is what we are using here (to solve FTC equations).

In any case, if we want the boundary condition $\phi(0) = 0$ to be satisfied, then the general solution we have found simplifies to

$$\phi(x) = \phi'(0)\;x - \int_0^x \int_0^{\tilde{\xi}} f(\xi)\,d\xi\,d\tilde{\xi}, \tag{1.16}$$

and we still have one parameter $\phi'(0)$ at our disposal. Thus, if we want to have $\phi(L) = 0$, then we can take

$$\phi'(0) = \frac{1}{L}\int_0^L \int_0^{\tilde{\xi}} f(\xi)\,d\xi\,d\tilde{\xi}$$

and we have solved the two point boundary value problem (1.14):

$$\phi(x) = \frac{x}{L}\int_0^L \int_0^{\tilde{\xi}} f(\xi)\,d\xi\,d\tilde{\xi} - \int_0^x \int_0^{\tilde{\xi}} f(\xi)\,d\xi\,d\tilde{\xi}. \tag{1.17}$$

1.6 The Shooting Method

I'm now going to solve the simple two point boundary value problem in two more ways. The second method of solution is called the shooting method, and you will see that in this case the procedure is, for all practical purposes, not very different from what we did in the last section. The shooting method, however, is a kind of standard method for solving two point boundary value problems (or spatial ODEs) and it works for many more general two point boundary value problems. So, on the one hand, it's nice to have an opportunity to illustrate the general method in a simple case, and it's also nice for you to learn a new method for ODEs which you have perhaps not previously encountered. On the other hand, there are two point boundary value problems for which the shooting method does not work for the simple reason that there are two point boundary value problems (spatial ODEs) which do not have any solution.

This might seem somewhat strange to you after having a first introductory course in temporal ODEs, because the initial value problem (IVP) always has a good solution under some minimal assumptions. Before we actually get into the (simple) application of the shooting method, let me make an attempt to discuss some of the underlying theoretical distinctions between the initial value problem (temporal ODE) and the two point boundary value problem (spatial ODE).

One theorem you should have considered a highlight of your first course and a cornerstone of your understanding of ODE is the **local existence and uniqueness theorem**. There are various versions, but the following is a good place to start:

Theorem 4. (existence and uniqueness for temporal ODE) If a and b are real numbers with $a < b$ and $f : \mathbb{R} \times (a, b) \to \mathbb{R}$ is a real valued function which is continuously differentiable on the strip $\Sigma = \mathbb{R} \times (a, b)$, then for each $t_0 \in (a, b)$ and each $x_0 \in \mathbb{R}$, there exists some $r > 0$ so that the IVP

$$\begin{cases} x' = f(x, t), & t_0 - r < t < t_0 + r \\ x(t_0) = x_0 \end{cases} \tag{1.18}$$

has a unique solution $x \in C^1(t_0 - r, t_0 + r)$.

The set $C^1(t_0 - r, t_0 + r)$ consists of the continuously differentiable real valued functions defined on the open interval $(t_0 - r, t_0 + r)$. This set is a vector space, and Problem 1.3 introduces the concepts and notation associated with continuity, differentiability, and continuous differentiability. This is generally called **regularity**. The regularity requirements on the function f in the equation $x' = f(x, t)$ may be called **structural regularity**, and the regularity of a solution of the equation may be distinguished as the **regularity of solutions**.

When we say f is **continuously differentiable** on the strip Σ, we mean the partial derivatives

$$\frac{\partial f}{\partial x}(x_0, t_0) = \lim_{v \to 0} \frac{f(x_0 + v, t_0) - f(x_0, t_0)}{v}$$

and

$$\frac{\partial f}{\partial t}(x_0, t_0) = \lim_{v \to 0} \frac{f(x_0, t_0 + v) - f(x_0, t_0)}{v}$$

exist for every $(x_0, t_0) \in \Sigma$ and the functions

$$\frac{\partial f}{\partial x} \qquad \text{and} \qquad \frac{\partial f}{\partial t}$$

are continuous on Σ. The shorthand notation for the assumption of continuous differentiability is $f \in C^1(\Sigma)$.

For the moment, note the fact that the theorem gives no control or estimate on how large (or small) the positive number r may be. This is quite important, and it is good to understand how this plays out in some examples. Perhaps the main first example is

$$x' = x^2.$$

This is a nonlinear separable equation. It has one solution $x(t) \equiv 0$. This is the unique solution satisfying $x(t_0) = 0$ for any $t_0 \in \mathbb{R}$, and it happens to be defined on all of $\mathbb{R}$ and satisfy $x \in C^\infty(\mathbb{R})$.

If we consider a nonzero initial value $x(t_0) = x_0 \neq 0$, then the existence and uniqueness theorem, Theorem 4 applies to tell us there exists a unique C^1 solution defined on some interval $(t_0 - r, t_0 + r)$ with $r > 0$. For t close enough to t_0, furthermore, this solution will satisfy $x(t) \neq 0$. (Why?) As a consequence, we can find a formula for the solution as follows: We write

$$\frac{1}{x^2}\, x' = 1$$

and integrate from t_0 to t to obtain

$$\int_{b_0}^{t} \frac{1}{x(\tau)^2}\, x'(\tau)\, d\tau = t - t_0.$$

Changing variables in the integral on the left using $\xi = x(\tau)$ so that $d\xi = x'(\tau)\,d\tau$ we find

$$\int_{x_0}^{x} \frac{1}{\xi^2}\,d\xi = -\frac{1}{x} + \frac{1}{x_0} = t - t_0.$$

Thus,

$$x(t) = \frac{1}{(1/x_0) - (t - t_0)}.$$

Notice that this solution has a singularity (and a vertical asymptote) at $t_1 = t_0 + 1/x_0$. Though the (structural) regularity of the ODE ($f(x,t) = x^2$ with $f \in C^\infty(\mathbb{R}^2)$) is quite good, solutions can only be expected to be defined on a symmetric interval $(t_0 - r, t_0 + r)$ of half length $r = 1/|x_0|$. It is difficult to "see" the length $r = 1/|x_0|$ of the **interval of existence** for solutions just by looking at the ODE

$$x' = x^2.$$

Another good example which illustrates the structural regularity requirement in Theorem 4 and is good to know about is

$$x' = \sqrt{|x|}.$$

Notice that the zero solution $x(t) \equiv 0$ satisfies the initial value problem

$$\begin{cases} x' = \sqrt{|x|}, & x \in \mathbb{R} \\ x(0) = 0, \end{cases}$$

but this is not the only solution of this IVP. See Problem 1.8 below. In fact, if we only want existence (but maybe not uniqueness) then the structural regularity may be reduced to continuity:

Theorem 5. (Peano existence theorem) If a and b are real numbers with $a < b$ and $f : \mathbb{R} \times (a,b) \to \mathbb{R}$ is a real valued function which is continuous on the strip $\Sigma = \mathbb{R} \times (a,b)$, then for each $t_0 \in (a,b)$ and each $x_0 \in \mathbb{R}$, there exists some $r > 0$ so that the IVP

$$\begin{cases} x' = f(x,t), & t_0 - r < t < t_0 + r \\ x(t_0) = x_0 \end{cases} \tag{1.19}$$

has a solution $x \in C^1(t_0 - r, t_0 + r)$.

The two examples $x' = x^2$ and $x' = \sqrt{|x|}$ do not contain explicit dependence on the independent (time) variable t. That is, when we write down the structure function $f = f(x,t)$ we do not need the argument t, but can simply use $f = f(x)$. Such ODEs are called **autonomous**.

There are two or three other existence and uniqueness theorems about which it is nice to know. In particular, our star ODE appearing the the two point boundary value problem $x'' = f$ is a second order equation, and the theorems above only apply to a first order equation. You may (or may not) recall there was a curious way to get existence and uniqueness for **ODEs of higher order** in the temporal theory of ODEs. This was to consider **first order systems**.

1.7 Exercises and Problems

You should produce solutions of the problems in a form that may be submitted for feedback—if you want feedback and to get a grade of "A" or "B" in the course. When you see an exercise you will usually see/find a reference to some numbered equation in the notes above. You should usually look back at the text surrounding that numbered equation, review the notes, and figure out the context of the exercise from the discussion there.

Exercise 1.1. Recall the solution of the time dependent ODE $L^2\,\phi' = \pi^2(1-\phi)$ given above using the fact that the equation is separable and leading to (1.12).

(a) Why/how does one know there exists some $\delta > 0$ such that $\phi(t) \neq -1$ for $0 \le t < \delta$?

(b) Why is the lower limit of integration in (1.12) equal to zero?

(c) Why was I able to write $\ln(1-\phi)$ on the right in (1.12) instead of $\ln|1-\phi|$ which is what an integral table would suggest for that integral?

Problem 1.1. Recall the initial/boundary value problem

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + \sin\left(\dfrac{\pi}{L}x\right), & 0 < x < L,\ t > 0 \\ u(x,0) = 0, & 0 \le x \le L \\ u(0,t) = 0 = u(L,t), & t \ge 0 \end{cases}$$

considered in Lecture 1.

(a) Verify that the solution given above actually satisfies all the conditions required in (1.10).

(b) Take $L = 2$ and use mathematical software to plot the value $u(x, t_j)$ as a function of x for several positive times t_j, $j = 1, 2, 3, 4$.

(c) Take $L = 2$ and use mathematical software to animate the evolution of the temperature conduction in a thin rod modeled by (1.10) with the time t aas an animation parameter.

Exercise 1.2. Recall the solution of the time dependent ODE $L^2\ \phi' = \pi^2(1 - \phi)$ is also a first order linear equation which may be solved using an integrating factor. Solve the equation with initial value $\phi(0) = 0$ carefully using an integrating factor. "Carefully" means you should justify each step, carry out the integrations with explicit limits of integration, quote theorems when necessary, and so forth.

Problem 1.2. A linear first order ODE on an interval $(0, L)$ has the form $F[\phi] = g$ where $g = g(x)$ and $F : C^1[0, L] \to C^0[0, L]$ is a linear operator. Complete the following steps to give a different solution of the IVP

$$\begin{cases} L^2\ \phi' = \pi^2(1 - \phi), & t \geq 0 \\ \phi(0) = 0. \end{cases} \tag{1.20}$$

(a) When we say F is linear, we mean that for $c_1, c_2 \in \mathbb{R}$ and $f_1, f_2 \in C^1[0, L]$ there holds

$$F[c_1 f_1 + c_2 f_2] = c_1 F[f_1] + c_2 F[f_2]. \tag{1.21}$$

An expression like $c_1 f_1 + c_2 f_2$ is called a **linear combination** and the condition (1.21) for L is called the **linearity relation**.

Rewrite the ODE in the standard linear form $F[\phi] = g$; identify the linear operator and verify the linearity relation.

(b) Given any linear problem/equation of the form $F[\phi] = g$, the equation

$$F[\phi] = 0 \tag{1.22}$$

is called the **associated homogeneous equation**. In this context, the original equation is sometimes called the original **inhomogeneous equation**.

Show any linear combination of solutions of a homogeneous linear equation (1.22) is a solution of the same equation.

(c) Find **all** solutions of the homogeneous equation associated to the linear equation you identified in part **(a)** above. Hint: The equation is separable; you should find a one parameter family of solutions.

(d) Again for any linear problem/equation of the form $F[\phi] = g$ (show that) if you can find a **particular solution** ϕ_*, i.e., an element for which $F[\phi_*] = g$, then **every** (other) solution ϕ of the original equation has the form

$$\phi = \phi_0 + \phi_*$$

where ϕ_0 is a some solution of the associated homogeneous equation.

(e) Find a particular solution ϕ_* of the linear inhomogeneous equation from part **(a)** by guessing a particular form for ϕ_*. Hint: Take ϕ_* to be a constant function.

(f) At this point, you should be able use part **(d)** to find a one parameter family of solutions to the linear inhomogeneous equation from part **(a)**. Use the initial condition in (1.20) to determine a unique value of the parameter giving a solution of the IVP (1.20)/(1.11).

Exercise 1.3. Take the limit as $t \nearrow \infty$ in each condition of the initial/boundary value problem

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + f(x), & 0 < x < L,\ t > 0 \\ \\ u(x,0) = u_0(x), & 0 \le x \le L \\ u(0,t) = 0 = u(L,t), & t \ge 0 \end{cases} \tag{1.23}$$

under the assumption that

$$u_{\text{old}}(x) = \lim_{t \nearrow \infty} u(x,t) \tag{1.24}$$

exists as a function in $C^2[0, L]$.

(a) Write down the conditions on convergence of the partial derivatives appearing in the PDE required to obtain an ODE for u_{old}; what is this ODE?

(b) What happens when you take the limit of the initial condition?

(c) If (1.24) implies pointwise convergence, what two point boundary condition do you obtain for u_{old}?

Exercise 1.4. Find the derivative of the function $g \in C^1[0, L]$ with

$$g(x) = \phi'(0) - \int_0^x f(\xi)\, d\xi.$$

This function appears in (1.15).

Exercise 1.5. Use the fundamental theorem of calculus to check that the function $\phi \in C^2[0, L]$ with values given by

$$\phi(x) = \frac{x}{L}\int_0^L \int_0^{\tilde{\xi}} f(\xi)\, d\xi\, d\tilde{\xi} - \int_0^x \int_0^{\tilde{\xi}} f(\xi)\, d\xi\, d\tilde{\xi}$$

in (1.17) solves the two point boundary value problem

$$\begin{cases} u''_{\text{old}} = -f, & 0 < x < L \\ u_{\text{old}}(0) = 0 = u_{\text{old}}(L) \end{cases}$$

given in (1.14) as the **associated long-time equilibrium solution** for the forced heat equation in (1.13).

Problem 1.3. (continuity and continuous differentiability) For this problem, let x denote an independent variable on an open interval

$$I = (a, b) = \{x \in \mathbb{R} : a < x < b\}$$

where a and b are real numbers with $a < b$. A function $f : (a, b) \to \mathbb{R}$, meaning f assigns a real number $f(x)$ to each element x in the domain $I = (a, b)$, is **continuous** at $x_0 \in I$ if for each $\epsilon > 0$, there exists some $\delta > 0$ such that

$$|f(x) - f(x_0)| < \epsilon \qquad \text{whenever} \qquad |x - x_0| < \delta.$$

This is the definition of **continuity at a point** x_0. The same function f is said to be **continuous on all of** $I = (a, b)$ if f is continuous at each point $x_0 \in I$. This is the definition of **continuity on the open interval** $I = (a, b)$. In this latter case, we write $f \in C^0(a, b)$. That is, $C^0(a, b)$ denotes the collection of all continuous real valued functions on the interval (a, b).

(a) If f and g are continuous on (a, b), show $f+g \in C^0(a, b)$. Thus, $C^0(a, b)$ is **closed under addition**.

(b) If $c \in \mathbb{R}$ and $f \in C^1(a, b)$, then $cf \in C^0(a, b)$. Thus, $C^0(a, b)$ is closed under scaling.

(c) It was mentioned that $C^0(a, b)$ is a **vector space**. It is important that every vector space V contain an **additive identity element**, that is, an element $\mathbf{0} \in V$ such that $v+\mathbf{0} = \mathbf{0}+v = v$ for every $v \in V$. Identify the additive identity element in $C^0(a, b)$.

(d) A function $f : (a, b) \to \mathbb{R}$ is said to be **differentiable at** $x_0 \in (a, b)$ if

$$\lim_{v \to 0} \frac{f(x_0 + v) - f(x_0)}{v}$$

exists. This means there is some real number L (called the limit) for which given any $\epsilon > 0$, there exists some $\delta > 0$ such that

$$\left| \frac{f(x_0 + v) - f(x_0)}{v} - L \right| < \epsilon \qquad \text{whenever} \qquad |v| < \delta.$$

When this happens, the limit L is called the **derivative** of f at x_0 and is denoted by

$$f'(x_0) \text{ (Newton's notation)} \qquad \text{or} \qquad \frac{df}{dx}(x_0) \text{ (Leibniz' notation).}$$

If f is differentiable at every $x_0 \in (a, b)$, then we say f is **differentiable on the** (entire)**open interval** $I = (a, b)$. This is the definition of **differentiability on the open interval** $I = (a, b)$. There is no (standard) special notation for the collection of differentiable functions on an open interval, though this collection of functions does form a vector space.

A function f which is differentiable on (a, b) determines a second function $f' : (a, b) \to \mathbb{R}$, called the **derivative of** f on (a, b). If $f' \in C^0(a, b)$, then f is said to be **continuously differentiable** on the interval (a, b), and we write $f \in C^1(a, b)$. That is $C^1(a, b)$ is the collection of all continuously differentiable functions on the interval (a, b).

Show a function which is differentiable at $x_0 \in (a, b)$ is continuous at x_0. Conclude that $C^1(a, b) \subset C^0(a, b)$.

Problem 1.4. Find a function $f \in C^0(0,1) \backslash C^1(0,1)$.

Problem 1.5. Draw the strip $\Sigma = \mathbb{R} \times (a,b)$ and illustrate the existence and uniqueness assertion of Theorem 4. Hint(s): Your picture should have some appropriate interval labeled with endpoints a and b and some length(s) labeled $2r$ (or r).

Problem 1.6. Carefully use the existence and uniqueness theorem for ODEs to show the zero solution is the unique solution of the initial value problem

$$\begin{cases} x' = x^2, & t \in \mathbb{R} \\ x(0) = 0. \end{cases}$$

Be careful, it is not enough to apply the theorem at $t_0 = 0$. Why?

Problem 1.7. Consider the ODE $x' = x^2$ in the standard form $x' = f(x,t)$.

(a) Identify the function f determining the structure of this ODE.

(b) Determine the largest **natural domain** Σ on which f is defined.

(c) Determine the greatest **degree of regularity** of f on the natural domain you found in part **(b)**, for example, find the largest k for which $f \in C^k(\Sigma)$.

Problem 1.8. Consider the IVP

$$\begin{cases} x' = \sqrt{|x|}, & x \in \mathbb{R} \\ x(0) = 0. \end{cases}$$

(a) Find a solution $x \in C^1(\mathbb{R})$ with $x(1) > 0$.

(b) Explain why this situation may be expected in view of Theorems 4 and 5.

Lecture 2

Spatial Ordinary Differential Equations (continued)

2.1 Series Solutions

Consider the initial value problem

$$\begin{cases} y'' + y = 0, & x \in \mathbb{R} \\ y(0) = y_0, \\ y'(0) = y_0'. \end{cases} \tag{2.1}$$

You can solve this problem in various ways. Here is a way you may not have seen:

Assume

$$y(x) = \sum_{j=0}^{\infty} a_j x^j$$

is given as a power series with center of expansion $x_0 = 0$ with (unknown) coefficients $a_0, a_1, a_2, a_3, \ldots$. Then $y(0) = a_0$,

$$y'(x) = \sum_{j=1}^{\infty} j a_j x^{j-1},$$

and $y'(0) = a_1$. Therefore, we must have $a_0 = y_0$ and $a_1 = y_0'$. Differentiating the series expression again we can write

$$y''(x) = \sum_{j=2}^{\infty} j(j-1)a_j x^{j-2}$$
$$= \sum_{j=0}^{\infty} (j+2)(j+1)a_{j+2}x^j$$

where it may be noticed that we shifted the indices by two to obtain the last expression. Substituting the series expressions in the ODE we find

$$\sum_{j=0}^{\infty} \left[(j+2)(j+1)a_{j+2} + a_j\right]\ x^j = 0.$$

In order for the left side to be the power series expansion of the zero function, all the coefficients in this series must vanish. That is, we must have

$$a_{j+2} = -\frac{1}{(j+2)(j+1)}\ a_j \qquad \text{for} \qquad j = 0, 1, 2, 3, ldots.$$

It follows inductively that every coeffient with an even index is a multiple of a_0, and every coefficient with an odd index is a multiple of a_1. In fact, it is not too difficult to see that

$$a_{2k} = \frac{(-1)^k}{(2k)!}\ a_0 \qquad \text{and} \qquad a_{2k+1} = \frac{(-1)^k}{(2k+1)!}\ a_1 \qquad k = 0, 1, 2, 3, \ldots.$$

Rearranging the terms in the original assumed power series expansion we see

$$y = a_0\ \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\ x^{2k} + a_1\ \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\ x^{2k+1}$$
$$= a_0\ y_0(x) + a_1 y_1(x)$$

where $\{y_0, y_1\} \subset C^\omega(\mathbb{R})$ is a basis of solutions for the solution set

$$\Sigma = \{y \in C^\omega(\mathbb{R}) : y'' + y = 0\}.$$

In particular the solution space Σ is a two-dimensional vector space over $\mathbb{R}$.

2.2 Fourier series solutions

Recall that for our comparison to PDE we considered a two point boundary value problem for a relatively simple linear second order ODE. One such equation is $y'' = f(x)$. Given $L > 0$, I will now describe another approach to using a series to solve an IVP

$$\begin{cases} y'' + y = f, \quad x \in (0, L) \\ y(0) = 0, \\ y(L) = 0. \end{cases} \tag{2.2}$$

Here we assume the solution y, instead of a power series expansion, has a series expansion of the form

$$y = \sum_{j=1}^{\infty} a_j \sin\left(\frac{j\pi}{L}\, x\right). \tag{2.3}$$

As with the power series approach, the objective is to determine the coefficients $a_1, a_2, a_3, \ldots$. This may seem like something of an improbably approach, but it might be more plausible if it is possible to write the inhomogeneiety f as such a series as well:

$$f(x) = \sum_{j=1}^{\infty} b_j \sin\left(\frac{j\pi}{L}\, x\right). \tag{2.4}$$

In principle if $f \in C^0[0, L]$ we can only imagine this happens for functions f which are "compatible" in the sense that $f(0) = f(L) = 0$. Still there are some functions like that, and as we will see later even if f does not satisfy the homogeneous boundary conditions from the BVP, it can still make sense to consider such a series (which is a little bit amazing). Series like the ones in (2.3) and (2.4) are called **Fourier sine series**.

If we differentiate, more or less as we did with the power series, we find

$$y' = \sum_{j=1}^{\infty} a_j \left(\frac{j\pi}{L}\right) \cos\left(\frac{j\pi}{L}\, x\right)$$

and

$$y'' = -\sum_{j=1}^{\infty} a_j \left(\frac{j\pi}{L}\right)^2 \sin\left(\frac{j\pi}{L}\, x\right).$$

None of the terms disappear in this case and no shifting of indices is necessary.

Substituting in the ODE $y'' + y = f$ we get

$$\sum_{j=1}^{\infty} a_j \left[-\left(\frac{j\pi}{L}\right)^2 + 1 \right] \sin\left(\frac{j\pi}{L}\, x\right) = \sum_{j=1}^{\infty} b_j \sin\left(\frac{j\pi}{L}\, x\right).$$

Equating the coefficients yeilds the interesting relation

$$a_j = \frac{1}{-\left(\frac{j\pi}{L}\right)^2 + 1}\, b_j \qquad \text{for} \qquad j = 1, 2, 3, \ldots.$$

You may notice that this formula is going to be problematic if $L = j\pi$. So in this case, we want to consider intervals $(0, L)$ with $L \neq j\pi$ for $j = 1, 2, 3, \ldots$. Again, there are certainly some interval lengths L for which this condition holds. If

$$L \notin \{j\pi : j = 1, 2, 3, \ldots\}$$

we say L is not in the **spectrum** of the operator $Ly = y'' + y$.

In summary, we are left with a series which looks like it very well might represent some kind of solution of the BVP:

$$y = \sum_{j=1}^{\infty} \frac{1}{-\left(\frac{j\pi}{L}\right)^2 + 1}\, b_j \ \sin\left(\frac{j\pi}{L}\, x\right). \tag{2.5}$$

We are also left with a number of questions about this kind of approach:

1. Which functions f are reasonably represented by Fourier series as in (2.4)?
2. What kind of representation does a Fourier (sine) series really give?
3. Assuming everything goes well and all the various assumptions are satisfied, does the series for the solution given in (2.5) "converge" in some sense?

In a certain sense the last question promises to have a "nice" answer. Generally, as with power series, convergence will work better when and if the coefficients are smaller, and notice that for j large the numbers

$$a_j = \frac{1}{-\left(\frac{j\pi}{L}\right)^2 + 1}\, b_j$$

are certainly at least as small as the coefficients b_j. Thus, one expects that if the inhomogeneity f has a reasonable Fourier sine series expansion, then the solution should as well.

Part II

Fourier Series

Lecture 3

Linear spaces of functions; norms

3.1 C^k spaces

We have discussed the linear spaces $C^0(a,b)$ consisting of the continuous real valued functions with domain an open interval $(a,b) \subset \mathbb{R}$ with $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a < b$ and $C^1(a,b)$ consisting of functions in $C^0(a,b)$ with (at least) one derivative in $C^0(a,b)$. These are called the **space of continuous functions** and the **space of continuously differentiable functions** respectively. These two spaces provide prototypes for the spaces $C^0(U)$ of continuous functions $u : U \to \mathbb{R}$ where U is an open subset of $\mathbb{R}^n$ for $n \geq 1$ and continuously differentiable functions in $C^0(U)$ with continuous first order partial derivatives. The latter space is denoted by $C^1(U)$.

There can be some complications defining $C^0(A)$ and $C^1(A)$ consisting of functions with domain $A \subset \mathbb{R}^n$ (for some $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$) when A is not an open set. The situation with $C^0(A)$ is relatively easy to handle, but there are in fact different choices for the definition of $C^1(A)$ especially when the dimension n is greater than $n = 1$. Ignoring these difficulties for the moment the C^0 (continuity) spaces and the C^1 (continuous differentiability) spaces can be used as a foundation for defining a natural heirarchy of greater degrees of differentiability. Roughly speaking one may proceed inductively to define

$$C^k(A) = \{u \in C^{k-1}(A) : D^\beta u \in C^1(A) \text{ for } |\beta| = k - 1\} \tag{3.1}$$

so that for example

$$C^2(A) = \left\{ u \in C^1(A) : \frac{\partial u}{\partial x_j} \in C^1(A), j = 1, 2, \ldots, n \right\},$$

and one usually wishes to have $C^1(A) \supset C^1(A) \supset C^2(A) \supset \cdots$. Thus, one can define

$$C^\infty(A) = \bigcap_{k=0}^{\infty} C^k(A)$$

the linear space of infinitely differentiable functions.

3.1.1 multiindices

The notation for higher order partial derivatives used in (3.1) may be unfamiliar. In the notation $D^\beta u$, the symbol β represents a **multiindex**, that is, an element of $\mathbb{N}_0^n$ where $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ is the natural numbers with zero. That is $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ where each β_j is a nonnegative integer. There are, first of all, various convenient quantities associated with a multiindex β:

norm $|\beta| = \beta_1 + \beta_2 + \cdots + \beta_n$.

power $\mathbf{x}^\beta = x_1^{\beta_1} x_2^{\beta_2} \cdots x_n^{\beta_n}$ for $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

factorial $\beta! = \beta_1!\ \beta_2!\ \cdots\ \beta_n!$.

Each of these is a generalization of familiar quantities from arithmetic. Notice the norm, however, is not the Euclidean norm, but rather the sum of the entries. You can determine for yourself if this quantity deserves to be called a norm.

The notation $D^\beta u$ is a very efficient way to express the partial derivative given in classical notation by

$$D^\beta u = \frac{\partial^{|\beta|} u}{\partial x_1^{\beta_1} \partial x_2^{\beta_2} \cdots \partial x_n^{\beta_n}}.$$

3.1.2 power series

Given $f \in C^\infty(a, b)$ and $x_0 \in (a, b)$ one has derviatives

$$f^{(j)}(x_0) = \frac{d^j f}{dx^j} \qquad \text{for} \qquad j = 0, 1, 2, 3, \ldots.$$

Furthermore the **Taylor coefficients** are given by

$$\frac{f^{(j)}(x_0)}{j!}$$

in terms of these derivatives, and one has a **Taylor series expansion**

$$\sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!}\,(x-x_0)^j \tag{3.2}$$

with **center of expansion** x_0. Hopefully, all this is familiar, or at least you have seen it before. Some interesting facts of which you may or may not be aware are the following:

(i) It is **not always true** that

$$f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!}\,(x-x_0)^j \tag{3.3}$$

for a given $x \in (a,b)$. In fact, it is not always true that the series for a given $x \in (a,b)$ always represents any specific real number in any reasonable sense.

(ii) It is always true that

$$f(x_0) = \sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!}\,(x-x_0)^j \qquad \text{for} \qquad x = x_0.$$

Question: Do you see why?

(iii) It is **always true** that

$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(x_0)}{j!}\,(x-x_0)^j + \frac{f^{(k+1)}(x_*)}{(k+1)!}\,(x-x_0)^{k+1}$$

for $k = 0, 1, 2, 3, \ldots$ and some x_* between x and x_0. This is called the **Taylor approximation theorem**.

Basically, the reason the approximation of **(iii)** does not imply the equality in (3.3) is because sometimes the **Taylor remainder/error term**

$$\left| f(x) - \sum_{j=0}^{k} \frac{f^{(j)}(x_0)}{j!}\,(x-x_0)^j \right| = \left| \frac{f^{(k+1)}(x_*)}{(k+1)!}\,(x-x_0)^{k+1} \right|$$

does not tend to zero when k tends to ∞.

The linear space $C^\omega(a,b)$ is defined as follows:

$$\begin{aligned} C^\omega(a,b) = \{ f \in C^\infty(a,b) : &\text{ for each } x_0 \in (a,b) \\ &\text{there is some } \epsilon > 0 \text{ such that} \\ &\quad \text{the equality (3.3) holds for} \\ &\quad\quad \text{all } x \text{ with } x_0 - \epsilon < x < x_0 + \epsilon) \}\,. \end{aligned}$$

A function in $C^\omega(a,b)$ is locally represented by a power series and is said to be **real analytic**. Thus, we may extend our inclusion relations:

$$C^0(a,b) \supset C^1(a,b) \supset C^k(a,b) \supset C^\infty(a,b) \supset C^\omega(a,b), \qquad k = 2,3,4,\ldots.$$

Most "nice" functions you know are in $C^\omega(a,b)$.

This is a convenient time to mention power series expansion in several variables and the related function space $C^\omega(U)$. You may not have seen this topic before and it is not a main tool in this course, but sometimes the multivariable version of Taylor's approximation theorem can be just as useful as the one variable version. Also, we happen to have the appropriate notion at our disposal.

Given $u \in C^\infty(U)$, the Taylor coefficents of u at $\mathbf{p}$ are given by

$$\frac{D^\beta u(\mathbf{p})}{\beta!},$$

and the Taylor series for u with center of expansion $\mathbf{p}$ is

$$\sum_{|\beta|=0}^{\infty} \frac{D^\beta u(\mathbf{p})}{\beta!}\,(\mathbf{x}-\mathbf{p})^\beta.$$

If $\overline{B_r(\mathbf{p})} \subset U$ it always holds that

$$\left| u(\mathbf{x}) - \sum_{|\beta|=0}^{k} \frac{D^\beta u(\mathbf{p})}{\beta!}\,(\mathbf{x}-\mathbf{p})^\beta \right| \le C\, \frac{\displaystyle\sum_{|\beta|=k+1} \max_{\xi \in \overline{B_r(\mathbf{p})}} |D^\beta u(\xi)|}{\beta!}\; |(\mathbf{x}-\mathbf{p})^\beta|$$

for $\mathbf{x} \in B_r(\mathbf{p})$. If for each $\mathbf{p} \in U$ there exists some $r > 0$ for which $B_r(\mathbf{p}) \subset U$ and

$$u(bx) = \sum_{|\beta|=0}^{\infty} \frac{D^\beta u(\mathbf{p})}{\beta!} (\mathbf{x} - \mathbf{p})^\beta \qquad \text{for} \qquad \mathbf{x} \in B_r(\mathbf{p})$$

in the sense that

$$\lim_{k\to\infty} \left| u(\mathbf{x}) - \sum_{|\beta|=0}^{k} \frac{D^\beta u(\mathbf{p})}{\beta!} (\mathbf{x} - \mathbf{p})^\beta \right| = 0,$$

then we say u is **real analytic** and write $u \in C^\omega(U)$.

3.1.3 The spaces $C^k(\overline{U})$ and $L^p(U)$

As mentioned above, it is a somewhat delicate matter to define a linear space $C^k(A)$ when A is not an open set. In fact such a linear space may have different definitions in different contexts. Without getting into the details we will consider at least to some extent one special case of this somewhat delicate situation. This is the situation when the domain $A \subset \mathbb{R}^n$ of the functions $u : A \to \mathbb{R}$ under consideration is the closure of an open subset of $\mathbb{R}^n$. Thus we wish to consider for example $C^k(\overline{U})$ where $U \subset \mathbb{R}^n$ is an open set. The main advantage $C^k(\overline{U})$ has over $C^k(U)$ is that $C^k(\overline{U})$ can be equipped with a **norm** by means of which the distance from one function to another can be measured and the notion of convergence of functions can be properly considered.

Definition 3. (norm) If X is a real linear space a **norm** on X is a function $\| \cdot \| : X \to [0, \infty)$ satisfying the following properties

N1 $\|v\| = 0$ if and only if $v = \mathbf{0}$ is the zero vector in X.

N2 $\|cv\| = |c| \, \|v\|$ for every $c \in \mathbb{R}$ and $v \in X$.

N3 $\|v + w\| \le \|v\| + \|w\|$ whenever $v, w \in X$.

A linear space equipped with a norm is called a **normed linear space** or a **vector space**.[1]

[1]You may have thought of a "vector space" as simply what we are referring to as a "linear space," i.e., a set with operations of addition and scalar multiplication satisfying

Property **N1** is said to express that the norm is **positive definite** and property **N2** that the norm is **nonnegative homogeneous**. Property **N3** is called the **triangle inequality** for a norm.

The norm on $C^0(\overline{U})$ is called the C^0 norm (read "C-zero norm") and is given by

$$\|u\|_{C^0} = \max_{\mathbf{x}\in\overline{U}} |u(\mathbf{x})|. \tag{3.4}$$

You can see from this definition why there is no norm (or at least no simple natural norm) on $C^0(U)$ where $U \subset \mathbb{R}^n$ is open.

Question: Can you see from (3.4) why there is no norm on $C^0(U)$?

The simplest kind of convergence in $C^0(U)$ or $C^0(\overline{U})$ is **pointwise convergence**. One says a sequence of functions

$$\{u_j\}_{j=1}^{\infty} \subset C^0(U)$$

converges to a function $u \in C^0(U)$ if for each $\mathbf{p} \in U$

$$\lim_{j\to\infty} u_j(\mathbf{p}) = u(\mathbf{p}).$$

That is, given a fixed point $\mathbf{p} \in U$ and some $\epsilon > 0$, there exists some $N > 0$ so that

$$j > N \qquad \text{implies} \qquad |u_j(\mathbf{p}) - u(\mathbf{p})| < \epsilon. \tag{3.5}$$

The condition (3.5) is often expressed in symbols as

$$j > N \qquad \Longrightarrow \qquad |u_j(\mathbf{p}) - u(\mathbf{p})| < \epsilon.$$

Pointwise convergence seems like a natural notion of convergence, but this kind of convergence allows some complicated situations which are often difficult to deal with mathematically. In particular, it is not easy to express

certain algebraic properties. This is the usual terminology in elementary linear algebra courses. In such courses the main space under consideration is the Euclidean space $\mathbb{R}^n$, and $\mathbb{R}^n$ is almost invariably considered with the Euclidean norm. On the other hand, I may offer a caution that the terminology I am using is not universal by any means. In fact, I discovered at least one author who uses exactly the opposite terminology referring to a set in which elements can be added and scaled by real numbers as a "vector space" and reserving the term "linear space" to denote such a set equipped with a a norm.

pointwise convergence in terms of any kind of reasonable distance between functions.

In contrast, we say a sequence $\{u_j\}_{j=1}^{\infty} \subset C^0(U)$ **converges uniformly** to $u \in C^0(U)$ if given $\epsilon > 0$ there is some N for which

$$j > N \qquad \text{implies} \qquad |u_j(\mathbf{x}) - u(\mathbf{x})| < \epsilon \qquad \text{for every } \mathbf{x} \in U.$$

This may seem like a condition that is not very different from pointwise convergence, but it really is very different. Furthermore, if we carry over the same definition to $C^0(\overline{U})$ where we have a norm, then this condition becomes

$$j > N \qquad \Longrightarrow \qquad \|u_j - u\|_{C^0} < \epsilon, \tag{3.6}$$

and we say

$$\lim_{j\to\infty} u_j = u \qquad \text{with respect to the } C^0 \text{ norm.} \tag{3.7}$$

It turns out to be really nice that a condition like (3.6) can always be interpreted as a natural limiting condition involving the distance between two functions whenever the functions are in a normed space. The norm may change giving a different kind of convergence, but once you have the norm, then you can say very precisely and conveniently what you mean by that particular kind of convergence.

There is a nice norm on each C^k space for $k = 1, 2, 3, \ldots$. The C^1 norm on $C^1(\overline{U})$ is often expressed in terms of the C^1 *seminorm*

$$[\,\cdot\,]_{C^1} : C^1(\overline{U}) \to [0, \infty) \qquad \text{by} \qquad [u]_{C^1} = \max_{|\beta|=1, \mathbf{x}\in\overline{U}} |D^\beta u(\mathbf{x})| = \max_{|\beta|=1} \|D^\beta u\|_{C^0}.$$

Question: Can you see why the C^1 seminorm is not a norm? If you can, then you can guess the definition of the mathematical object called a **seminorm** of which the C^1 seminorm is an example.

The C^1 norm is given by

$$\|u\|_{C^1} = \|u\|_{C_0} + [u]_{C^1}.$$

It turns out that there is another C^1 norm that is in common use. This "other" C^1 norm is defined by

$$\|u\|_{C^1} = \sum_{|\beta|\le 1} \max_{\mathbf{x}\in\overline{U}} |D^\beta u(\mathbf{x})| = \|u\|_{C^0} + \sum_{|\beta|=1} \|D^\beta u\|_{C^0}.$$

It turns out that in most instances one doesn't need to worry too much about which norm is being used because these two norms are **equivalent**.

Definition 4. (equivalent norms) Given two norms $\| \cdot \|_1 : X \to [0, \infty)$ and $\| \cdot \|_2 : X \to [0, \infty)$ on the same linear space X, we say $\| \cdot \|_1$ and $\| \cdot \|_2$ are **equivalent** norms if there exist positive constants m and M such that

$$m\|v\|_1 \le \|v\|_2 \le M\|v\|_1 \qquad \text{for every } v \in X.$$

You may have noticed there are actually two different definitions for a seminorm on $C^1(\overline{U})$ here. I guess one could define what it means to have equivalent seminorms, but I don't remember ever seeing that.

Similarly, there are equivalent norms on $C^k(\overline{U})$ given by

$$\|u\|_{C^k} = \max_{|\beta|\le k} \|D^\beta u\|_{C^0}$$

or

$$\|u\|_{C^k} = \sum_{|\beta|\le k} \|D^\beta u\|_{C^0}.$$

The C^k spaces for $k \in \mathbb{N}_0 \cup \{\infty\}$ are well-suited for many purposes. For example, $C^\infty(a, b)$ is a convenient domain on which to consider an ordinary differential operator $N : C^\infty(a, b) \to C^\infty(a, b)$. On the other hand, it is often natural to consider an n-th order ordinary differential operator on the larger space $C^n(a, b)$. In this case, there is a minor complication: One cannot expect the operator to have the same domain and co-domain, so that in this case the natural operator has

$$N : C^n(a, b) \to C^0(a, b).$$

There are from some perspective a number of shortcomings of the C^k spaces. Three of these are the following:

1. These spaces are not the appropriate spaces on/in which to consider Fourier series.

2. These spaces do not contain many functions of interest to engineers for applications.

3. $C^k(U)$ is not a normed space for U and open subset of $\mathbb{R}^n$

As examples of the second shortcoming observe that the **absolute value function** $f : \mathbb{R} \to [0, \infty)$ by

$$f(x) = \begin{cases} -x, & x \le 0 \\ x, & x \ge 0 \end{cases}$$

satisfies $f \in C^0(\mathbb{R}) \backslash C^k(\mathbb{R})$ for $k = 1, 2, 3, \ldots$. Also, the **Heaviside function** $h : \mathbb{R} \to \{0, 1\}$ by

$$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

and the **standard unit pulse** $\sigma : \mathbb{R} \to \{0, 1\}$ by

$$\sigma(x) = h(x) - h(x - 1)$$

are not in any C^k space over an interval containing $x = 0$.

More generally it can be said that power series are most compatible with or heavily dependent on differentiability whicle functions of interest in applications are often better considered in terms of Fourier series which are fundamentally compatible with integration.

Lecture 4

Fourier Series

In Chapter 3 of his book *Applied Partial Differential Equations with Fourier Series and Boundary Value Problems* Richard Haberman asks several questions about series of the form

$$a_0 + \sum_{j=1}^{\infty} a_j \;\cos\left(\frac{j\pi}{L}\; x\right) + \sum_{j=1}^{\infty} b_j \;\sin\left(\frac{j\pi}{L}\; x\right)$$

where $L > 0$ is given. Perhaps a main question is the following: Given a function $f : [-L, L] \to \mathbb{R}$ is there a choice of coefficients $a_0, a_1, a_2, \ldots, b_1, b_2, b_3, \ldots$ such that (in some rigorous sense)

$$f(x) = a_0 + \sum_{j=1}^{\infty} a_j \;\cos\left(\frac{j\pi}{L}\; x\right) + \sum_{j=1}^{\infty} b_j \;\sin\left(\frac{j\pi}{L}\; x\right) \qquad ? \tag{4.1}$$

As we have seen the function f may be taken to be in many different function spaces, and the answer to the main question above may be expected to depend on that kind of property of the function f. For example, if $f \in C^0[-L, L]$, then a very natural way for the equality in (4.1) to make sense is in the C^0 norm, that is (4.1) can take on the rigorous meaning that the limit of partial sums

$$f_n = a_0 + \sum_{j=1}^{n} a_j \;\cos\left(\frac{j\pi}{L}\; x\right) + \sum_{j=1}^{n} b_j \;\sin\left(\frac{j\pi}{L}\; x\right)$$

satisfying[1] $\{f_n\}_{n=1}^{\infty} \subset C^{\infty}[-L, L]$ also satisfies

$$\lim_{n\to\infty} \|f_n - f\|_{C^0} = 0, \tag{4.2}$$

[1] In fact $f_n \in C^{\infty}(\mathbb{R})$.

which we know is equivalent to uniform convergence. We have pointed out that one may very well wish to consider other functions which are not continuous like the Heaviside function or pulse functions. For these more general functions we could consider the question of simple pointwise convergence, and we will to a certain extent. It is much less complicated and rather more convenient in most instances, however, to consider notions of convergence which can be considered in a manner analogous to (4.2) but possibly with respect to some different norm or **distance** in a function space (see below).

In a certain sense the main question may be "unpacked" along the following lines:

1. What conditions must f satisfy for (4.1)?

2. In what sense can (4.1) be expected to hold, i.e., in what norm to the partial sums converge to f? Note that this is really preliminary to question **1**.

3. Given f how should the coefficients $a_0, a_1, a_2, \ldots, b_1, b_2, b_3, \ldots$ be chosen? Is there only one way? Note that this is really preliminary to question **2**.

So we should perhaps start with question **3**.

Haberman also points out that the answers to these (kinds of) questions are generally complicated. We have at least set the rough context of function spaces and norms in which the main answer can be given. Let us pause for a moment and describe one additional structure which may be helpful in setting the context and, at any rate, is a structure with which it can be useful to be familiar.

Definition 5. (abstract distance) Given any set X, a function $d : X \times X \to [0, \infty)$ satisfying

M1 $d(x, y) \geq 0$ with equality if and only if $x = y$.

M2 $d(x, y) = d(y, x)$ for any $x, y \in X$.

M3 $d(x, z) \leq d(x, y) + d(y, z)$ for any $x, y, z \in X$.

The function d is variously called a **distance function**, a **metric distance**, or simply a **metric**. A set X equipped with a distance function is sometimes called a *distance space* but much more often a **metric space**. Property **M1** is said to express that the metric is **positive definite**. Property **M2** is **symmetry** of the distance function, and property **M3** is called the **trinagle inequality** for metric distance.

Theorem 6. Every normed space is a metric space with $d(x, y) = \|x - y\|$.

You can prove this result; this particular distance function given by $d(x, y) = \|x - y\|$ is called the metric distance **induced** by the norm or the **norm induced distance**.

4.0.1 Fourier coefficients

Given a function $f \in \mathfrak{L}^1(-L, L)$, which means

$$\int_{-L}^{L} |f(x)|\, dx = \int_{(-L,L)} |f| \tag{4.3}$$

makes sense and is finite, it is always possible to find a formula for the Fourier coefficients $a_0, a_1, a_2, \ldots, b_1, b_2, b_3, \ldots$ in (4.1). We need to be a little careful with this statement. What is intended is that there exist unique numbers for which **some version** of (4.1) **might** hold. That may seem a little strange, but hopefully the significance of this strange statement will become clear(er) as we proceed further to answer the main question.

What should be clear (if you think about it a bit) is that we have definitely defined a function space $\mathfrak{L}^1(-L, L)$ called the first Lebesgue space, the "Lebesgue integrable functions," or just simply "ell-one." The expressions on each side of (4.3) are really just different notations for the integral of the absolute value of the function f. There are already a few technicalities I'm sweeping under the rug here, but let me give some superficial pseudo-explanation of the important underlying mathematical reality I'm glossing over. The notation on the left of (4.3) should look familiar. This is the usual notation for the integration introduced by Isaac Newton and Wilhelm Leibniz, and this notation is still used in calculus courses today where that integration is discussed and used. All the good theorems in those courses require the functions to which they apply to be in the nice C^k spaces we've discussed a little bit above. You may not have noticed this, but it's true. Here are five examples:

Theorem 7. (intermediate value theorem) If $f \in C^0[a, b]$ with $f(a) \neq f(b)$, then for each c between $f(a)$ and $f(b)$ there is some $x_* \in (a, b)$ with $f(x_*) = c$.

The conclusion of this result does not hold for the Heaviside function h on $[-1, 1]$ essentially because $h \notin C^0[-1, 1]$.

Theorem 8. (mean value theorem) If $f \in C^0[a, b] \cap C^1(a, b)$, then there exists some $x_* \in (a, b)$ with

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

Theorem 9. (fundamental theorems of calculus)

1. If $f \in C^1[a, b]$, then

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

2. If $f \in C^0[a, b]$, then $F : [a, b] \to \mathbb{R}$ by

$$F(x) = \int_a^x f(\xi)\, d\xi$$

satisfies $F \in C^1[a, b]$ and

$$F'(x) = f(x) \qquad \text{for} \qquad x \in [a, b].$$

Theorem 10. (Green's theorem) If $\mathbf{v} = (v_1, v_2) \in C^1(\overline{U} \to \mathbb{R}^2)$ where U is a bounded open subset of $\mathbb{R}^2$ with ∂U a simple closed (piecewise) C^1 curve with counterclockwise tangent vector field T, then

$$\int_U \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) = \int_{\partial U} \mathbf{v} \cdot T.$$

Theorem 11. (Taylor's theorem) If $u \in C^{k+1}(U)$ and $\mathbf{p} \in U$, then there is some $r > 0$ such that for each $\mathbf{x} \in B_r(\mathbf{p})$ there exists some points $\mathbf{x}_\beta \in B_r(\mathbf{p})$ for $|\beta| = k + 1$ such that

$$u(\mathbf{x}) = \sum_{|\beta| \le k} \frac{D^\beta u(\mathbf{p})}{\beta!} (\mathbf{x} - \mathbf{p})^\beta + \sum_{|\beta| = k+1} \frac{D^\beta u(\mathbf{x}_\beta)}{\beta!} (\mathbf{x} - \mathbf{p})^\beta.$$

You may not have seen these theorems in exactly these forms, and you may not have seen some of them at all, but they do illustrate how the natural setting for the main theorems in calculus is the C^k spaces. In fact, the integral used in calculus is what is traditionally called the **Riemann integral**[2] and the main theorem on the existence of the Riemann integral is that such an integral is well-defined for functions that are continuous on a closed interval, i.e., functions in $C^0[a, b]$. This fundamental fact of existence is especially evident in the statement of the fundamental theorem of calculus.

On the other hand, even Newton and Leibniz were well aware that it made good sense to integrate some functions that are not continuous. For example there is no problem integrating the Heaviside function. Also with some special cases called "improper integrals" it is pretty easy to make good sense of quantities like

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx \qquad \text{or} \qquad \int_1^\infty \frac{1}{x^2}\,dx.$$

In short, the Riemann integral makes fine sense for some discontinuous functions. However, there was no good existence theorem for Riemann integration applying to any reasonable large class of discontinuous functions. It was also known by the time the existence theorem for Riemann integrals of continuous functions was carefully verified that there are some functions like $\chi_{\mathbb{R}\setminus\mathbb{Q}} : [0, 1] \to \mathbb{R}$ by

$$\chi_{\mathbb{R}\setminus\mathbb{Q}}(x) = \begin{cases} 0, & \text{if } x \text{ is a rational number} \\ 1, & \text{if } x \text{ is not a rational number} \end{cases}$$

that were definitely **not** Riemann integrable. It was far from clear however where to draw the line between functions for which some kind of integration made sense and those for which no kind of integration made sense. Actually, the latter notion, that there can be functions which are so badly behaved that no notion of integration makes sense, was not really something most people expected to be the case. Basically, the overall situation was sorted out by Lebesgue who defined a notion of integration, now known as Lebesgue integration, and defined a function space of those functions for which integration makes sense. This is the Lebesgue space mentioned above and the

[2]The notation for the integral on the left in (4.3) is associated with the Riemann integral, and that is why this expression is familiar from calculus.

notation for integrals of this more general sort is on the right in (4.3). The function $\chi_{\mathbb{R}\setminus\mathbb{Q}}$ is in $\mathfrak{L}^1(0,1)$ and

$$\int_{(0,1)} \chi_{\mathbb{R}\setminus\mathbb{Q}} = 1.$$

In fact, using the new notation one can refer to the integral over a very large class of sets A (though not quite all sets) and the functions f in $\mathfrak{L}^1(A)$ by writing

$$\int_A f.$$

Thus,

$$\int_{[0,1]} \chi_{\mathbb{R}\setminus\mathbb{Q}}$$

makes perfectly good sense too, though one of the consequences of Lebesgue's identification of the sets upon which one can integrate and the functions one can integrate on those sets is that single points generally will not effect the value of an integral. In honor of this deep revelation we usually, when referring to an integral over any interval or a Lebesgue space, write the integral as if it is over the corresponding open interval (leaving out the two endpoints which do not make a difference).

There is one minor addition to the updated "Lebesgue notation" for integrals which is convenient to use and draws inspiration from Riemann's notation. Technically, we should always distinguish, at least in our minds, between **a function** which may have a name like f **and the values of the function** $f(x)$ at a given x in the domain of the function f. In practice we often say inaccurate things like "we have a function $f = f(x)$." At any rate it is sometimes useful to talk about a function and also have a name for and refer to the argument of that function in an integral. For example, say we have a function f. If $f : (a,b) \to \mathbb{R}$ is continuous, we can happily write

$$\int_a^b f(x) \sin x \, dx.$$

There is, of course, a name for the function determined by the values $\sin x$. You may not be surprised to know the name of this function is something like sine. Thus, perhaps the proper way to express the integral above when f is more complicated and we may need to use a Lebesgue integral is

$$\int_{(a,b)} f \text{ sine}.$$

We never do this however. It is just not really done. Among the things that people actually write, mostly when trying to write things "correctly," is

$$\int_{x\in(a,b)} f(x)\sin(x).$$

This is especially helpful if we are integrating with respect to the dependence in one variable, and there is another variable involved. This is the case, for example in

$$\int_{x\in(a,b)} f(x)\sin(x+y)$$

which is defining a function of y. The "general form" is

$$\int_{x\in A} f(x) \tag{4.4}$$

which indicates Lebesgue integration over a set A of a function f, precisely the same as the cleaner

$$\int_A f.$$

The notation (4.4) is used if for some reason we wish to indicate a/the variable of integration. This, as suggested above, comes about for primarily two reasons: Either there is a function whose name is unusual and disturbing to see in an integral, e.g.,

$$\int_A \text{sine},$$

or there are other variables involved, e.g.,

$$\int_{x\in\mathbb{Q}} \chi_{\mathbb{R}\setminus\mathbb{Q}}(x-y).$$

4.0.2 main answer

Haberman gives a theorem for the term by term integration of a Fourier series. It is certain that a much more general assertion is valid, but such results are not commonly stated because many if not most functions useful for applications have only a finite number of discontinuities. I believe the following is probably correct though I can't seem to find an explicit statement anywhere.

Conjecture 1. (termwise integration) If $f \in \mathfrak{L}^1(-L, L)$ and

$$f(x) = a_0 + \sum_{j=1}^{\infty} a_j \cos\left(\frac{j\pi}{L}x\right) + \sum_{j=1}^{\infty} b_j \sin\left(\frac{j\pi}{L}x\right)$$

in the sense of $\mathfrak{L}^1$, then for each $x \in (-L, L)$

$$\int_{(-L,x)} f = a_0(x+L) + \sum_{j=1}^{\infty} \frac{La_j}{j\pi} \sin\left(\frac{j\pi}{L}x\right) + \sum_{j=1}^{\infty} \frac{Lb_j}{j\pi}\left[\cos\left(\frac{j\pi}{L}x\right) - (-1)^j\right].$$

Assuming this result is valid we have (by evalution of the assertion at $x = L$)

$$\int_{(-L,L)} f = 2L\, a_0.$$

Notice this gives a formula for a_0:

$$a_0 = \frac{1}{2L}\int_{(-L,L)} f.$$

Applying the result to the function $g \in \mathfrak{L}^1(-L, L)$ given by

$$g(x) = f(x) \cos\left(\frac{k\pi}{L}x\right)$$

where $k = 1, 2, 3, \ldots$ and evaluating at $x = L$, we obtain something else interesting. To see this first observe that

$$\begin{aligned} g(x) = a_0 \cos\left(\frac{k\pi}{L}x\right) &+ \sum_{j=1}^{\infty} a_j \cos\left(\frac{j\pi}{L}x\right)\cos\left(\frac{k\pi}{L}x\right) \\ &+ \sum_{j=1}^{\infty} b_j \sin\left(\frac{j\pi}{L}x\right)\cos\left(\frac{k\pi}{L}x\right). \end{aligned}$$

This is not immediately of the form of a Fourier series though the first term would qualify. For the terms in the first summation recall the trigonometric identity

$$\begin{aligned} \cos\left(\frac{j\pi}{L}x\right)\cos\left(\frac{k\pi}{L}x\right) &= \cos\left(\frac{(j+k)\pi}{L}x\right) + \sin\left(\frac{j\pi}{L}x\right)\sin\left(\frac{k\pi}{L}x\right) \\ &= \frac{1}{2}\left[\cos\left(\frac{(j+k)\pi}{L}x\right) + \cos\left(\frac{(j-k)\pi}{L}x\right)\right] \end{aligned}$$

which follows more or less as indicated from the cosine addition formula. Upon substitution then we can write

$$\sum_{j=1}^{\infty} a_j \cos\left(\frac{j\pi}{L}\, x\right) \cos\left(\frac{k\pi}{L}\, x\right)$$
$$= \sum_{j=1}^{\infty} \frac{a_j}{2} \cos\left(\frac{(j+k)\pi}{L}\, x\right) + \sum_{j=1}^{\infty} \frac{a_j}{2} \cos\left(\frac{(j-k)\pi}{L}\, x\right).$$

The first term/summation on the right is a sum of fourier modes, and we can rewrite it

$$\sum_{m=k+1}^{\infty} \frac{a_{m-k}}{2} \cos\left(\frac{m\pi}{L}\, x\right)$$

just by shifting indices. In the second summation $j-k$ takes infinitely many positive integer values but also takes the value 0 when $j = k$ and if $k > 1$ will also take finitely many negative integer values. Overall, we can write

$$\sum_{j=1}^{\infty} \frac{a_j}{2} \cos\left(\frac{(j-k)\pi}{L}\, x\right)$$
$$= \sum_{j=1}^{k-1} \frac{a_j}{2} \cos\left(\frac{(j-k)\pi}{L}\, x\right) + \frac{a_k}{2}$$
$$+ \sum_{j=k+1}^{\infty} \frac{a_j}{2} \cos\left(\frac{(j-k)\pi}{L}\, x\right)$$
$$= \sum_{m=1}^{k-1} \frac{a_{k-m}}{2} \cos\left(\frac{(-m)\pi}{L}\, x\right) + \frac{a_k}{2}$$
$$+ \sum_{m=1}^{\infty} \frac{a_{m+k}}{2} \cos\left(\frac{m\pi}{L}\, x\right)$$
$$= \frac{a_k}{2} + \sum_{m=1}^{k-1} \left(-\frac{a_{k-m}}{2} + \frac{a_{m+k}}{2}\right) \cos\left(\frac{m\pi}{L}\, x\right)$$
$$+ \sum_{m=k}^{\infty} \frac{a_{m+k}}{2} \cos\left(\frac{m\pi}{L}\, x\right).$$

Again, this has the form of a Fourier series in $\mathfrak{L}^1(-L, L)$. Applying the same approach to the sum

$$\sum_{j=1}^{\infty} b_j \ \sin\left(\frac{j\pi}{L}\ x\right) \cos\left(\frac{k\pi}{L}\ x\right)$$

we note first the trigonometric identity

$$\begin{aligned}\sin\left(\frac{j\pi}{L}\ x\right) \cos\left(\frac{k\pi}{L}\ x\right) &= \sin\left(\frac{(j+k)\pi}{L}\ x\right) - \cos\left(\frac{j\pi}{L}\ x\right) \sin\left(\frac{k\pi}{L}\ x\right) \\ &= \frac{1}{2}\left[\sin\left(\frac{(j+k)\pi}{L}\ x\right) + \sin\left(\frac{(j-k)\pi}{L}\ x\right)\right]\end{aligned}$$

which gives

$$\begin{aligned}&\sum_{j=1}^{\infty} b_j \ \sin\left(\frac{j\pi}{L}\ x\right) \cos\left(\frac{k\pi}{L}\ x\right) \\ &\qquad = \sum_{j=1}^{\infty} \frac{b_j}{2} \sin\left(\frac{(j+k)\pi}{L}\ x\right) + \sum_{j=1}^{\infty} \frac{b_j}{2} \sin\left(\frac{(j-k)\pi}{L}\ x\right) \\ &\qquad = \sum_{m=k+1}^{\infty} \frac{b_{m-k}}{2} \sin\left(\frac{m\pi}{L}\ x\right) + \sum_{j=1}^{k-1} \frac{b_j}{2} \sin\left(\frac{(j-k)\pi}{L}\ x\right) \\ &\qquad\qquad + \sum_{j=k+1}^{\infty} \frac{b_j}{2} \sin\left(\frac{(j-k)\pi}{L}\ x\right).\end{aligned}$$

Notice here that the constant term drops out because when $j = k$

$$\sin\left(\frac{(j-k)\pi}{L}\ x\right) = \sin(0) = 0.$$

Consequently, this expression becomes

$$\begin{aligned}\sum_{j=1}^{\infty} b_j \ \sin\left(\frac{j\pi}{L}\ x\right)\cos\left(\frac{k\pi}{L}\ x\right)& \\ &= \sum_{m=k+1}^{\infty} \frac{b_{m-k}}{2} \sin\left(\frac{m\pi}{L}\ x\right) - \sum_{m=1}^{k-1} \frac{b_{k-m}}{2} \sin\left(\frac{m\pi}{L}\ x\right) \\ &\qquad + \sum_{m=1}^{\infty} \frac{b_{m+k}}{2} \sin\left(\frac{m\pi}{L}\ x\right) \\ &= \sum_{m=1}^{k-1} \frac{b_{m+k} - b_{k-m}}{2} \sin\left(\frac{m\pi}{L}\ x\right) + \frac{b_{2k}}{2} \sin\left(\frac{2k\pi}{L}\ x\right) \\ &\qquad + \sum_{m=k+1}^{\infty} \frac{b_{m-k} + b_{m+k}}{2} \sin\left(\frac{m\pi}{L}\ x\right).\end{aligned}$$

Putting all these calculations together (with some more combining of terms) we see the function

$$g(x) = f(x)\cos\left(\frac{k\pi}{L}\ x\right)$$

is indeed expressible as a Fourier series and a Fourier series with constant term $a_k/2$. Integrating termwise from $-L$ to L, assuming this is justified and gives correct information, we see

$$\int_{x\in(-L,L)} f(x)\cos\left(\frac{k\pi}{L}\ x\right) = La_k$$

or

$$a_k = \frac{1}{L}\int_{x\in(-L,L)} f(x)\cos\left(\frac{k\pi}{L}\ x\right)$$

for $k = 1, 2, 3, \ldots$, which is the value Haberman simply defines as one of the Fourier coefficients in his boxed equations (3.2.2) on page 91. You can check that Haberman gives a value of a_0 which agrees with the one calculated above, and you can calculate to see he gives a nice formula for b_j when $j = 1, 2, 3, \ldots$.

4.0.3 The space $\mathfrak{L}^2$

The real problem with the foregoing discussion is that not every function in $\mathfrak{L}^1$ is represented by its Fourier series. There is the smaller set of functions in $\mathfrak{L}^1$ which do happen to be represented by their Fourier series in $\mathfrak{L}^1$, and we know what the coefficients have to be in that case, but this set is not very nicely tied to the norm and does not provide a very satisfactory answer to our original question.

There is a better answer. Given an interval (a, b) with $a, b \in \mathbb{R}$ satifying as always $a < b$, there is for each $p > 1$ a Lebesgue space consisting of those functions f in $\mathfrak{L}^1(a, b)$ for which the integral

$$\int_{(a,b)} |f|^p < \infty.$$

Among these spaces the choice $p = 2$ is special. These are precisely the functions that are represented by their Fourier series, and the natural convergence of Fourier series is with respect to the $\mathfrak{L}^2$ norm which looks like this:

$$\|u\|_{\mathfrak{L}^2} = \left(\int u^2\right)^{1/2}.$$

It's also true that $\mathfrak{L}^2(a, b) \subset \mathfrak{L}^1(a, b)$, so if there are going to be Fourier coefficients, they have to be the ones we found for $\mathfrak{L}^1$. Specifically, there is a thing called the **Hölder inequality** which says that if $p, q > 1$ with

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then given $f \in \mathfrak{L}^p$ and $g \in \mathfrak{L}^q$ the product fg is in $\mathfrak{L}^1$ with

$$\|fg\|_{\mathfrak{L}^1} = \int |fg| \le \|f\|_{\mathfrak{L}^p} \, \|g\|_{\mathfrak{L}^q} \tag{4.5}$$

where

$$\|f\|_{\mathfrak{L}^p} = \left(\int |f|^p\right)^{1/p} \qquad \text{and} \qquad \|g\|_{\mathfrak{L}^q} = \left(\int |g|^q\right)^{1/q}.$$

Hence you can just take $p = q = 2$ and and $g = 1$ to see

$$\|f\|_{\mathfrak{L}^1} \le (b - a)^{1/2} \, \|f\|_{\mathfrak{L}^2}.$$

so $\mathfrak{L}^2(a, b) \subset \mathfrak{L}^1(a, b)$. In fact, in the special case $p = q = 2$ both the functions f and g in (4.5) are in the same space, and the left side of (4.5) is often recognized to define or at least involve the values of a function

$$\langle \,\cdot\, , \,\cdot\, \rangle : \mathfrak{L}^2(a, b) \times \mathfrak{L}^2(a, b) \to \mathbb{R}$$

satisfying

$$\langle f, g \rangle = \int_{(a,b)} fg.$$

This function is called an **inner product** on $\mathfrak{L}^2$ and may be viewed as a generalization of the usual dot product on $\mathbb{R}^n$. As with the distance function and the norm there are well-defined properties satisfied by an inner product.

Definition 6. (inner product) Given a linear space X, a function

$$\langle \cdot , \cdot \rangle : X \times X \to \mathbb{R}$$

satisfying

IP1 $\langle x, x \rangle \geq 0$ with equality only if $x = \mathbf{0}$ is the zero vector in X.

IP2 $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in X$.

IP3 $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$ and $\langle z, ax + by \rangle = a\langle z, x \rangle + b\langle z, y \rangle$ for all $x, y, z \in X$.

IP1 expresses that the inner product is **positive definite**, **IP2** is **symmetry**, and **IP3** is called **bilinearity**. Notice that if one only assumes linearity in the "first slot," then linearity in the "second slot" necessarily follows by symmetry.

A linear space X equipped with such a function is called an **inner product space**.

Oddly enough a linear space of functions with an inner product admits a notion of **angle** between pairs of nonzero functions with the angle θ between f and g satisfying $0 \leq \theta \leq \pi$ given by

$$\theta = \cos^{-1}\left(\frac{\langle f, g \rangle}{\|f\| \ \|g\|}\right). \tag{4.6}$$

In particular, two functions in an inner product space[3] are said to be **orthogonal** or **perpendicular** if $\langle f, g \rangle = 0$. When Haberman says the Fourier coefficients are derived using "certain orthogonality integrals," this is what he means, though he never introduces the notion of an (abstract) inner product space or the space $\mathfrak{L}^2(-L, L)$ of square integrable functions.

[3]Or more generally two elements in any inner product space.

Implicit in (4.6) is the fact that a linear space with an inner product is a normed space with a specific norm induced by the inner product; specifically in (4.6)

$$\|f\| = \sqrt{\langle f, f\rangle}. \tag{4.7}$$

The main interesting part in showing the formula (4.7) for the inner product norm actually defines a norm, specifically showing the triangle inequality for norms holds for the function $\| \cdot \| : X \to [0, \infty)$ defined in (4.7), is usually executed using an inequality called the **Cauchy-Schwarz inequality**.

Theorem 12. (Cauchy-Schwarz inequality) In any inner product space X one has

$$|\langle x, y\rangle| \le \|x\| \, \|y\| \qquad \text{for all } x, y \in X.$$

Of course this same inequality is expressed without the suggestive use of the norm notation by

$$|\langle x, y\rangle|^2 \le \langle x, x\rangle \, \langle y, y\rangle \qquad \text{or} \qquad |\langle x, y\rangle| \le \sqrt{\langle x, x\rangle} \, \sqrt{\langle y, y\rangle}.$$

Technically, there is a slight problem in that $\mathfrak{L}^p$ is not a normed space as it stands and in particular $\mathfrak{L}^2$ is not quite an inner product space. There are, however, in all cases closely related normed spaces L^p, and the spaces $\mathfrak{L}^p$ we have introduced for $p > 1$ make a pretty good stand-in for the nicer L^p spaces. I won't get into the details of this technical glitch, but I'll leave it to you to figure out the possible shortcomings of $\mathfrak{L}^p$.

In view of the foregoing discussion, when we look at a Fourier expansion

$$f(x) = a_0 + \sum_{j=1}^{\infty} a_j \, \cos\left(\frac{j\pi}{L}\, x\right) + \sum_{j=1}^{\infty} b_j \, \sin\left(\frac{j\pi}{L}\, x\right)$$

as in (4.1) it is most natural to view the right side as a kind of infinite linear combination orthogonal basis vectors in $\mathfrak{L}^2(-L, L)$ and focus not only on the coefficients, but the functions comprising the basis as well. These functions are v_0 the constant function with $v_0(x) \equiv 1$,

$$v_j = \cos\left(\frac{j\pi}{L}\, x\right) \qquad \text{for} \qquad j = 1, 2, 3, \ldots$$

and

$$w_j = \sin\left(\frac{j\pi}{L}\, x\right) \qquad \text{for} \qquad j = 1, 2, 3, \ldots.$$

These functions form an orthogonal basis for $\mathfrak{L}^2(-L, L)$ in the sense that

$$\int v_0 v_j = \int (1) v_j = 0 \quad \text{and} \quad \int v_0 w_j \int (1) w_j = 0 \text{ for } j = 1, 2, 3, \ldots,$$

$$\int v_j v_k = 0 = \int w_j w_k \qquad \text{for} \qquad j, k = 1, 2, 3, \ldots, \qquad \text{with} \qquad j \neq k$$

and

$$\int v_j w_k = 0 \qquad \text{for all } j, k = 1, 2, 3, \ldots.$$

The other interesting numbers are the norms associated with these basis vectors/functions $v_0, v_1, v_2, v_3, \ldots, w_1, w_2, w_3, \ldots$. The squares of those numbers are

$$\|v_0\|^2 = \int v_0^2 = \int 1 = 2L,$$

$$\begin{aligned} \|v_j\|^2 &= \int_{x \in (-L,L)} \cos^2\left(\frac{j\pi}{L}\, x\right) \\ &= \int_{x \in (-L,L)} \frac{1}{2}\left[1 + \cos\left(\frac{2j\pi}{L}\, x\right)\right] \\ &= L \end{aligned}$$

for $j = 1, 2, 3, \ldots$, and

$$\|w_j\|^2 = L$$

for $j = 1, 2, 3, \ldots$ as well.

Let's see how this all works. Take $L = 1$ and consider the function $f \in C^\omega[-1, 1]$ with values $f(x) = x - x^3$. Of course, $C^\omega[-1, 1] \subset \mathfrak{L}(-1, 1)$ so we can expect a Fourier series expansion

$$f = a_0 + \sum_{j=1}^{\infty} a_j v_j + \sum_{j=1}^{\infty} b_j w_j.$$

This particular function has a property which will be of interest to us in general, specifically, f is **odd** meaning that

$$f(-x) = -f(x) \qquad \text{for} \qquad x \in [-1, 1].$$

Generally, we say a function f is odd if $f(-x) = -f(x)$ for all x in the domain of f (and of course both x and $-x$ are in the domain of f whenever

x is). Also, a function f is **even** if $f(-x) = f(x)$ for all x. These two properties will come up again. But this particular function $f(x) = x - x^3$ is odd. As a consequence

$$\begin{aligned}\int_{(-L,L)} f &= \int_{(-L,0)} f + \int_{(0,L)} f \\ &= \int_{x\in(-L,0)} f(x) + \int_{(0,L)} f \\ &= \int_{x\in(0,L)} f(-x) + \int_{(0,L)} f \\ &= \int_{x\in(0,L)} [-f(x)] + \int_{(0,L)} f \\ &= -\int_{x\in(0,L)} f(x) + \int_{(0,L)} f \\ &= 0.\end{aligned} \tag{4.8}$$

In fact, whenever an odd function is integrated on a symmetric interval the result is always zero. Incidentally, the change of variable in (4.8) may have struck you as lacking a minus sign. It's actually okay because when you integrate on sets, as one does with Lebesgue integration, the usual convention of orientation for Riemann integrals do not apply. To be very specific on this, so you can get it straight, let me state a result:

Theorem 13. (change of variables) If $f \in \mathfrak{L}^1(a,b)$ and $\psi : (c,d) \to (a,b)$ is a differentiable monotone change of variables, then

$$\int_{(a,b)} f = \int_{(c,d)} f \circ \psi \; |\psi'|. \tag{4.9}$$

Notice the absolute value of ψ' in the scaling factor. This is in contrast to the familiar Riemann integral change of variable formula which would (or might) look like

$$\int_a^b f(x)\, dx = \int_{\psi^{-1}(a)}^{\psi^{-1}(b)} f \circ \psi(\xi)\; \psi'(\xi)\, d\xi.$$

The point is that maybe $\psi^{-1}(a) > \psi^{-1}(b)$, so the definition of

$$\int_{\psi^{-1}(a)}^{\psi^{-1}(b)} f \circ \psi(\xi)\; \psi'(\xi)\, d\xi$$

in this case is

$$\int_{\psi^{-1}(a)}^{\psi^{-1}(b)} f \circ \psi(\xi)\ \psi'(\xi)\, d\xi = -\int_{\psi^{-1}(b)}^{\psi^{-1}(a)} f \circ \psi(\xi)\ \psi'(\xi)\, d\xi$$
$$= \int_{\xi \in (\psi^{-1}(b), \psi^{-1}(a))} f \circ \psi(\xi)\ \psi'(\xi).$$

This is the Riemann integral orientation convention. You can check that the signs line up correctly so that the absolute value in (4.9) gives the correct answer when one is integrating directly on intervals as sets.

Returning to $f(x) = x - x^3$, we have shown that the coefficient of the basis function $v_0 \equiv 1$ is

$$a_0 = \frac{1}{L} \int_{(-L,L)} f = 0.$$

More generally, the product of an odd function, like f, with an even function like v_j for $j = 1, 2, 3, \ldots$ is odd, so

$$a_j = \frac{1}{2L} \int_{x \in (-L,L)} f(x)\ \cos\left(\frac{j\pi}{L}\, x\right) = 0.$$

That is, for an odd function on $(-L, L)$, all the coefficients of even basis functions vanish.

Now the real work begins. For the nonzero coefficients we need to calculate

$$b_j = \frac{1}{L} \int_{x \in (-L,L)} f(x)\ \sin\left(\frac{j\pi}{L}\, x\right)$$
$$= \frac{1}{L} \left[\int_{x \in (-L,L)} x\ \sin\left(\frac{j\pi}{L}\, x\right) - \int_{x \in (-L,L)} x^3\ \sin\left(\frac{j\pi}{L}\, x\right) \right]. \tag{4.10}$$

There are various ways to proceed. I think they all pretty much involve integration by parts, so one needs to get used to that. Also, using mathematical software with an integration facility can be handy at least to double check answers. It's easy to make errors. Taking first the integral

$$\int_{x \in (-L,L)} x^3\ \sin\left(\frac{j\pi}{L}\, x\right),$$

I'll take $u = x^3$ and

$$dv = \sin\left(\frac{j\pi}{L}\ x\right)\ dx$$

so that $du = 3x^2\,dx$ and

$$v = -\frac{L}{j\pi}\ \cos\left(\frac{j\pi}{L}\ x\right).$$

The integration by parts formla $\int u\,dv = uv - \int v\,du$ then gives

$$\begin{aligned}\int_{x\in(-L,L)} x^3\ \sin\left(\frac{j\pi}{L}\ x\right) &= -x^3\ \frac{L}{j\pi}\ \cos\left(\frac{j\pi}{L}\ x\right)\Big|_{x=-L}^{L} \\ &\quad + 3\frac{L}{j\pi}\int_{x\in(-L,L)} x^2\ \cos\left(\frac{j\pi}{L}\ x\right). \\ &= -\frac{2L^4}{j\pi}\ (-1)^j + \frac{3L}{j\pi}\int_{x\in(-L,L)} x^2\ \cos\left(\frac{j\pi}{L}\ x\right).\end{aligned}$$

This suggests to me the separate consideration of

$$\begin{aligned}\int_{x\in(-L,L)} x^2\ \cos\left(\frac{j\pi}{L}\ x\right) &= x^2\ \frac{L}{j\pi}\ \sin\left(\frac{j\pi}{L}\ x\right)\Big|_{x=-L}^{L} \\ &\quad - 2\frac{L}{j\pi}\int_{x\in(-L,L)} x\ \sin\left(\frac{j\pi}{L}\ x\right) \\ &= -\frac{2L}{j\pi}\int_{x\in(-L,L)} x\ \sin\left(\frac{j\pi}{L}\ x\right),\end{aligned}$$

and finally

$$\begin{aligned}\int_{x\in(-L,L)} x\ \sin\left(\frac{j\pi}{L}\ x\right) &= -x\ \frac{L}{j\pi}\ \cos\left(\frac{j\pi}{L}\ x\right)\Big|_{x=-L}^{L} \\ &\quad + \frac{L}{j\pi}\int_{x\in(-L,L)} \cos\left(\frac{j\pi}{L}\ x\right) \\ &= -\frac{2L^2}{j\pi}\ (-1)^j + \left(\frac{L}{j\pi}\right)^2 \sin\left(\frac{j\pi}{L}\ x\right)\Big|_{x=-L}^{L} \\ &= -\frac{2L^2}{j\pi}\ (-1)^j.\end{aligned}$$

Substituting back, checking with Mathematica, and watching the signs, hopefully the following should be correct:

$$\int_{(-L,L)} x \ \sin\left(\frac{j\pi}{L}\ x\right) = -\frac{2L^2}{j\pi}\ (-1)^j,$$

$$\int_{x\in(-L,L)} x^2 \ \cos\left(\frac{j\pi}{L}\ x\right) = \frac{4L^3}{(j\pi)^2}\ (-1)^j,$$

and

$$\int_{x\in(-L,L)} x^3 \ \sin\left(\frac{j\pi}{L}\ x\right) = \frac{2L^4}{j\pi}\ (-1)^j \left[-1+\frac{6}{(j\pi)^2}\right].$$

Returning to (4.10) we see

$$\begin{aligned} b_j &= \frac{2}{L}\left[-\frac{L^2}{j\pi}+\frac{L^4}{j\pi} \ -\frac{6L^4}{(j\pi)^3}\right](-1)^j \\ &= \frac{L}{j\pi}\left[-1+L^2-\frac{6L^2}{(j\pi)^2}\right](-1)^j. \end{aligned}$$

Finally, taking account of the fact that $L=1$ we get

$$b_j = \frac{12}{(j\pi)^3}\ (-1)^{j+1},$$

and we can expect

$$\lim_{n\to\infty} \|f_n - f\|_{\mathfrak{L}^2} = 0$$

where

$$f_n(x) = \frac{12}{\pi^3}\sum_{j=1}^{n} \frac{(-1)^{j+1}}{j^3}\ \sin\left(\frac{j\pi}{L}\ x\right)$$

and

$$\|f_n - f\|_{\mathfrak{L}^2} = \left(\int (f_n - f)^2\right)^{1/2}$$

as usual. Of course in this case $f \in C^\omega[-1,1]$ and we can expect pointwise convergence and even uniform convergence of any fixed number of derivatives. In Figure 4.1 I have plotted $|f_n - f|$ for $n=2$, $n=3$ and $n=10$. This gives a pretty good indication that I've gotten the coefficients correct.

You may ask why one would turn in a perfectly good function like $x - x^3$ for a complicated series of sines. One answer is that these sinusoidal functions

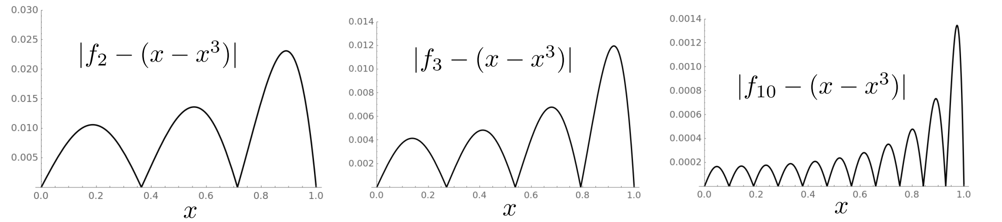

Figure 4.1: The C^0 or $\mathfrak{L}^\infty$ error $|f_n - f|$ of the Fourier series for $f(x) = x - x^3$ for partial sums with 2 terms (left), 3 terms (middle), and 10 terms (right).

have some special properties with respect to derivatives. Surely you noticed when you took/learned ODEs that

$$v_j'' = -\left(\frac{j\pi}{L}\right)^2 v_j \qquad \text{and} \qquad w_j'' = -\left(\frac{j\pi}{L}\right)^2 w_j$$

for $j = 1, 2, 3, \ldots$. This observation has some interesting consequences for the three partial differential equations we want to consider. For example, if we look at the function $u : [0, L] \times [0, M] \to \mathbb{R}$ of two variables given by

$$u(x, y) = \sin\left(\frac{j\pi}{L}\, x\right) \sin\left(\frac{j\pi}{M}\, y\right), \tag{4.11}$$

then

$$\Delta u = \frac{\partial u}{\partial x^2} + \frac{\partial u}{\partial y^2} = -(j\pi)^2 \left(\frac{1}{L^2} + \frac{1}{M^2}\right)\, u.$$

That is, the product of Fourier basis functions given in (4.11) is a kind of **eigenfunction** for the Laplace operator $\Delta : C^\infty([0, L] \times [0, M])$.

Looking at this another way, we have found an explicit solution of the Poisson partial differential equation

$$\Delta u = -(j\pi)^2 \left(\frac{1}{L^2} + \frac{1}{M^2}\right)\, \sin\left(\frac{j\pi}{L}\, x\right) \sin\left(\frac{j\pi}{M}\, y\right)$$

satisfying $u(x, y) \equiv 0$ for (x, y) in the boundary of the rectangle $U = [0, L] \times [0, M]$. This may seem kind of special, but solutions of partial differential equations are not so easy to come by, so any explicit solutions you happen to know about may turn out to be of use.

It turns out that if you have zero boundary conditions like this, say on the rectangle $U = [0, L] \times [0, M]$, then there are no interesting solutions to Laplace's equation $\Delta u = 0$. The only solution is just $u \equiv 0$. You can find interesting solutions of Laplace's equation however if you take, for example,

$$u(x, y) = \sin\left(\frac{j\pi}{L}\, x\right) \sinh\left(\frac{j\pi}{L}\, y\right).$$

Again this kind of explicit solution turns out to be useful in solving more general problems due to the connection with Fourier series.

Another PDE of interest in the **heat equation**

$$u_t = u_{xx} \qquad \text{or} \qquad \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

Taking

$$u(x, t) = e^{at}\ \sin\left(\frac{j\pi}{L}\, x\right)$$

it is possible to pick the constant a so that $u = u(x, t)$ is a solution of the heat equation. Try it.

Finally,

$$u(x, t) = \cos(\omega t)\ \sin\left(\frac{j\pi}{L}\, x\right)$$

will be an explicit solution of the wave equation

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

if the constant ω is chosen correctly.

In short, the Fourier basis functions appear as "pieces" of solutions of the partial differential equations we are supposed to learn something about, and ultimately this is why we are interested in them. Though it's also very cool to be able to express or approximate a given function in terms of the Fourier basis functions, and this has other applications inside and outside PDE.

Part III

The Heat Equation

Part IV

Laplace's Equation

Part V

The Wave Equation