Fourier Analysis & Methods
The mathematics of heat, light, waves, and sound

Julie Rowlett

March 12, 2021

Contents

Chapter 1

Introduction: every equation requires a sound check!

1.1 Every concert, and every partial differential equation, requires a sound check.

Have you ever noticed that at a concert, even if the band has played numerous times in the exact same venue, they *always* do a sound check? The venue is the same, so why don't they just have the same settings every time they play that venue? The mathematical equation of sound is the wave equation, and it's the same mathematical equation every time the band plays. If we could develop the equation and solve it wouldn't sound checks be obsolete? The complexity of sound is a challenging matter. This unfortunate condition is the reason bands do sound checks to create some certainty that they sound their best. The complexity of sound is similarly confronted when we try to solve the mathematical equation of sound.

The mathematical equation of sound is not easy to solve, because although the equation may remain the same, its solution changes depending on many different factors. Similarly, the way a band sounds depends on many different factors. How many people are at the venue? Where are they standing or sitting? What are they wearing? As I was waiting for my friend's first performance in the heavy metal band, Sodom, watching them do their sound check, I realized that performing sound checks at concerts is similar to solving problems in Fourier analysis! You can see my friend rocking the stage in Figure 1.1.

1.1.1 The definition of ordinary and partial differential equations.

The mathematical equation of sound is the wave equation and is an example of a *partial differential equation.*

Definition 1 (Ordinary and partial differential equations)**.** An ordinary differential equation is an equation for an unknown function that depends on *one* independent real variable. Writing u for the unknown function, an ordinary differential equation for u is an equation that u must satisfy and that contains u together with one or more derivatives of u. The ordinary differential equation may also contain other, *specified* functions. A partial differential equation is an equation for an unknown function that depends on *two or more* independent real variables. Writing U for the unknown function, a partial differential equation for U is an equation that U must satisfy and that contains U together with one or more partial derivatives of U. The partial differential equation may also contain other, *specified* functions.

Example 2. *Here is an ordinary differential equation: $u' - u = 0$. It is common to omit the independent variable, but if we choose to include it and call it x, then the equation looks like $u'(x) - u(x) = 0$. Here is a partial differential equation: $U_{xx} - 9U_{tt} = 0$. We have again omitted the independent variables, but we can guess from the notation that the function U depends on two independent variables, named x and t. We could just as well write the equation $\frac{\partial^2}{\partial_x^2}U(x,t) - 9\frac{\partial^2}{\partial_t^2}U(x,t) = 0$.*

Figure 1.1: On March 3, 2019, I went to see my friend play guitar in the band, Sodom, in Gothenburg. They not only rocked, but also inspired me with the 'sound check analogy' comparing band sound checks to Fourier analysis and methods. Photograph copyright Moritz 'Mumpi' Künster.

The wave equation, and indeed nearly all partial differential equations are emphatically HARD to solve. There is no single unifying theory or user manual that we can consistently follow to solve partial differential equations. It's like playing in a band: we have to do a sound check for each and every concert. There is no magic pre-set we can use for all our concerts. Similarly, we have to carefully investigate each and every partial differential equation. Many of them cannot be solved, but our focus will be on those that can be solved with a collection of techniques known as *Fourier analysis and methods.*

1.2 How to read this text

What kind of music do you like? I enjoy many different genres including heavy metal and industrial. Now, just because I have watched and listen to awesome guitarists, does this mean I can play the guitar? Hardly! To play awesome music, one must practice! Mathematics is quite similar. One must do mathematics to learn it. The mathematical analogue of practicing guitar is *active learning.* With this text, I am asking you to read it *actively.* The goal is not for me to explain a bunch of math to you while you sit and read and absorb it like a sponge. On the contrary, I invite you to think critically as you read. Be skeptical. Verify my claims on your own terms, make sure you understand how one line in a proof follows from the hypotheses and the preceding calculations. Most importantly: do the exercises!

1.3 Help me illustrate this text!

Do they also have the saying in Sweden that 'a picture is worth a thousand words?' I have included some pictures in this text, but due to time limitations together with the fact that I suck at illustrating, it is quite likely that this text could be improved by including further pictures and illustrations. Would *you* like to help? Do *you* have an idea for an illustration that could help to convey a certain concept in the text? If so, then just like that old United States Army poster in Figure 1.2, "we need you!" If you contribute a picture, figure, graph, illustration, gif, or any other visual, your contribution will be explicitly mentioned with your name to live in print forever. Unless you prefer to be anonymous. You may be able to contribute to both your class as well as future generations and the international community of math students ability to understand this subject! So, please keep this in mind as you read the text.

Acknowledgements

Thank you to everyone who contributed to the creation of this text! Gottfrid Olsson made Figures 2.4, 3.9, 3.5, 4.7, **??**, the pink paw print that signifies the end of a proof. Ebba Grönfors and Alva Brycke invented the forget-me-not-fluffballs, known as Katodanodes, that help us remember key concepts in Figures 2.6, 3.8, and 4.8. Eric Lindgren and his cat Berlioz contributed the figures in §4.2.1 and §4.8.2 created using the python script in §B.2. Anton Rosén contributed Figure 4.6 and the Matlab code used to create it in §B.1. Edvin Martinson contributed Figure 4.12. Joao Pedro Paulos, Erik Broback, Waldemar Hult, Anton Rosén, Ludvig Nordqvist, Petronella Taube, Martin Jirlow, Gottfrid Olsson, Edvin Martinson, Ronja Hedberg, Erik Levin, Sofia Wartenberg, Carl-Joar Karlsson, Samuel Winqvist, Adam Johansson, Isak Drevander, and more people? identified typos in an early version of the text.

Figure 1.2: Uncle Sam (note the initials U.S.) is a national personification of the United States federal government that according to legend was named for Samuel Wilson and introduced in the War of 1812. The famous 1917 poster by J. M. Flagg was used to recruit soldiers for the first and second world wars. Here, I am recruiting you to help me by contributing visual aids like pictures, figures, graphs, illustrations, gifs, or any other visualization of the mathematical content! Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`.

Chapter 2

Separation of variables makes the problem ordinary; superposition turns the ordinary extraordinary!

The first technique for solving partial differential equations is a method that enables us to turn partial differential equations for functions that depend on several variables into ordinary differential equations for functions that depend on only one variable. This technique is known as *separation of variables.* As in the description of this chapter, we will use this technique to change the problem of solving a partial differential equation into the problem of solving several ordinary differential equations, thereby making the problem *ordinary.* Once we have obtained all solutions of the ordinary problem, we will use the second technique, known as *superposition* to build a 'supersolution,' thereby transforming the ordinary solutions into an extraordinary solution of the original problem. We will introduce these techniques with the help of two fundamental examples: the wave equation that describes a vibrating string, like on that of a guitar, and the heat equation that describes the propagation of heat.

2.1 The sound of guitars can be understood by solving the wave equation

Without guitars, a heavy metal concert would just be an angry guy (or gal) shouting at drums. The rhythm, lead, and bass guitars create sounds that are crucial to the music. The vibrations of a guitar string and the sounds they make can be understood using Fourier analysis and methods. Consider a vibrating string, like the guitar or bass strings in our metal band. The ends of the string are held fixed, so they're not moving. You know this if you play or watch people play guitar. Let's mathematicize the string, by identifying it with the interval $[0, \ell] \subset \mathbb{R}$. The string length is ℓ. Let's define

$$u(x,t) := \text{ the height of the string at the point } x \in [0,\ell] \text{ at time } t \in [0,\infty[.$$

Then, let's just define the sitting-still height to be height 0. So, the fact that ends are sitting still means that

$$u(0,t) = u(\ell,t) = 0 \quad \forall t.$$

A positive height means above the sitting-still height, whereas a negative height means under the sitting-still height. The wave equation (I'm not going to derive it, but maybe you clever physics students can do that?) says that:

$$u_{tt} = \mathfrak{c}^2 u_{xx}.$$

The constant $\mathfrak{c}$ depends on how fast the string vibrates. The string is in some initial position with some initial velocity at the starting time:

$$u(x,0) = f(x), \quad u_t(x,0) = g(x).$$

The functions f and g are called *initial data.* In this method, we will handle the initial data *at the very end.* One way to remember this is with the acronym TIDGLAS: The Initial Data Goes LASt. The reason we do this is because the entire method works in exactly the same way, up until the very end, for *any* initial data. So, it is necessary to be patient, like waiting for an hourglass as shown in Figure 2.1, or in Swedish, timglas. Perhaps a more suitable name for an hourglass would be a timeglass. The Swedish translation would then be exactly tidglas.

Question 3. *Is this equation a PDE or an ODE?*[1]

2.1.1 The first technique for your toolbox: separation of variables

Separation of Variables starts like this: we *assume* that

$$u(x,t) = X(x)T(t),$$

that is a product of two functions, each of which depends only on *one* variable. Why can we do this? Who knows, maybe it is rubbish! Maybe u is not of this form. It is a bit like the sound check: we first make a guess for the sound levels and then play a bit to see if it sounds good. This is the same general idea.

Assuming that u is of this form, we substitute this into the PDE:

$$u_{tt} = \mathfrak{c}^2 u_{xx} \iff X(x)T''(t) = \mathfrak{c}^2 X''(x)T(t).$$

Now, we would like to *separate variables* by getting everything dependent on x to one side of the equation and everything dependent on t to the other side. To achieve this, we divide both sides by $X(x)T(t)$:

$$\mathfrak{c}^2 \frac{X''}{X}(x) = \frac{T''}{T}(t). \tag{2.1.1}$$

Stop. Think. The left side depends only on x, whereas the right side depends only on t.

Exercise 4. *Explain in your own words why if one side of an equation depends on x and the other side depends on t, then both sides must be constant.*

What should we solve for first? X or T? With this method, we will always solve for X first. The reason we do this is because of the boundary conditions. The boundary conditions for X is that the ends of the string do not move, so we must have

$$u(0,t) = X(0)T(t) = 0 \text{ for all times, and } u(L,t) = X(L)T(t) = 0 \text{ for all times.}$$

Since T does not depend on x at all, the only way to guarantee this is to demand that

$$X(0) = X(\ell) = 0.$$

On the other hand, the function T depends on time $t \geq 0$. Time can run off to infinity, so the only boundary for the t variable is at $t = 0$. The shape and velocity of the string at $t = 0$ is called the initial data. Our mantra TIDGLAS reminds us that the initial data goes last. We therefore look at equation (2.1.1) and consider what it means for X first. Both sides are equal to a constant, so

$$\frac{X''}{X}(x) = \text{ constant },$$

[1] Answer: it's a PDE because the function depends on two independent variables: position on the string x and time t.

Figure 2.1: An hourglass is one way to measure time by waiting until all of the sand passes from the top to the bottom. Then it is reset by flipping it over. In Swedish an hourglass is called a 'timglas.' Since an hourglass does not necessarily measure one hour, but could be made to measure many different lengths of time, why not call it a timeglass, or TIDGLAS in Swedish instead? Then this fits with our mantra to be patient, wait for the TIDGLAS and deal with the initial data last. Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`

and the boundary condition demands that

$$X(0) = X(\ell) = 0.$$

Let's give the constant a name. Call it λ. Then write

$$X''(x) = \lambda X(x), \quad X(0) = X(\ell) = 0.$$

Does this problem look familiar? It should have been covered in previous calculus (analysis) courses. There are three cases to consider:

Case 1: $\lambda = 0$. This means $X''(x) = 0$. One way to solve for X is to just think about physics. The second derivative is acceleration. This is zero. That means constant velocity. The only functions with constant velocity are linear. So, the graph of X looks like a straight line, and we also know that $X(0) = 0 = X(L)$. So this straight line is horizontal, that is $X = 0$. Or, if you prefer to solve for X, we can integrate twice and arrive at the same conclusion. Integrating once gives $X'(x) = \text{constant} = m$. Integrating a second time gives $X(x) = mx + b$. Requiring $X(0) = X(\ell) = 0$, first makes $b = 0$, and the second makes $m = 0$. So, the solution is $X(x) = 0$. The 0 solution. The waveless wave. Not too interesting.

Case 2: $\lambda > 0$. The solution here will be of the form

$$X(x) = ae^{\sqrt{\lambda}x} + be^{-\sqrt{\lambda}x}.$$

Exercise 5. *Show that it is equivalent to write the solution as $A\cosh(\sqrt{\lambda}x) + B\sinh(\sqrt{\lambda}x)$, for two constants A and B. Determine the relationship between A and B and a and b. Show that in order to guarantee that $X(0) = X(\ell) = 0$ you need $a = A = B = b = 0$.*

Thus, with our teamwork, (me providing hints and you doing the actual work by solving the exercise) we have gotten the 0 solution again. The waveless wave. No fun there.

Case 3: $\lambda < 0$. Finally, we have a solution of the form

$$a\cos(\sqrt{|\lambda|}x) + b\sin(\sqrt{|\lambda|}x).$$

To make $X(0) = 0$, we need $a = 0$. Uh oh... are we going to get that stupid 0 solution again? Well, let's see what we need to make $X(\ell) = 0$. For that we just need

$$b\sin(\sqrt{|\lambda|}\ell) = 0.$$

That will be true if

$$|\lambda| = \frac{k^2\pi^2}{\ell^2}, \quad k \in \mathbb{Z}.$$

Super! We still don't know what b ought to be, but at least we've found all the possible X's, up to constant factors.

Just to clarify the fact that we've now found *all* solutions, we recall here a theorem from multivariable calculus.

Theorem 6 (Second order ODEs)**.** *Consider the second order linear homogeneous ODE,*

$$au'' + bu' + cu = 0, \quad a \neq 0.$$

If $b = c = 0$, then a basis of solutions is given by

$$\{x, 1\},$$

so that all solutions can be expressed as a linear combination of x and 1, hence for some constants A and B

$$u(x) = Ax + B.$$

If $c = 0$, then a basis of solutions is $\{e^{-b/ax}, 1\}$ so that all solutions can be expressed as a linear combination of these two functions, so we can find constants A and B such that

$$u(x) = Ae^{-bx/a} + B.$$

If $c \neq 0$, then a basis of solutions is one of the following depending on whether or not $b^2 = 4ac$:

1. $\{e^{r_1 x}, e^{r_2 x}\}$ *if* $b^2 \neq 4ac$, *where*

$$r_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad r_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

2. $\{e^{rx}, xe^{rx}\}$ *if* $b^2 = 4ac$, *with* $r = -\frac{b}{2a}$.

Exercise 7. *The equation we solved using separation of variables is*

$$X'' = \lambda X \iff X'' - \lambda X = 0, \quad X(0) = X(\ell) = 0. \tag{2.1.2}$$

So, in the language of Theorem 6, $a = 1$, $b = 0$, *and* $c = -\lambda$. *Apply the theorem to show that up to constant factors, all solutions to* (2.1.2) *are the functions*

$$X_k(x) = \sin\left(\frac{k\pi x}{\ell}\right), \quad \lambda_k = -\frac{k^2\pi^2}{\ell^2}, \quad k \in \mathbb{Z}.$$

Now that we have determined the constants in (2.1.1), we can solve for the T functions as well! The equation, for each k reads:

$$\mathfrak{c}^2 \frac{X_k''}{X_k} = \lambda_k = -\frac{k^2\pi^2}{\ell^2} = \frac{T''}{T}(t).$$

This is almost the same equation we had before. Here we have, re-arranging:

$$T_k'' = -\mathfrak{c}^2 \frac{k^2\pi^2}{\mathfrak{c}^2 \ell^2} T_k. \tag{2.1.3}$$

Exercise 8. *Use Theorem 6 to show that a basis of solutions is given by*

$$\left\{ e^{\frac{ik\pi \mathfrak{c} t}{\ell}}, e^{-\frac{ik\pi \mathfrak{c} t}{\ell}} \right\}.$$

Show that it is equivalent to use

$$\left\{ \cos\left(\frac{k\pi t\mathfrak{c}}{\ell}\right), \sin\left(\frac{k\pi t\mathfrak{c}}{\ell}\right) \right\}$$

as a basis. Hint: remember $e^{i\theta} = \cos\theta + i\sin\theta$ *for* $i = \sqrt{-1}$ *for any* $\theta \in \mathbb{R}$.

We have therefore found *all* functions of the form XT that satisfy the partial differential equation and the boundary conditions. There are infinitely many of them, and they can be enumerated as

$$u_k(x,t) = X_k(x)T_k(t) = A_k \sin\left(\frac{k\pi x}{\ell}\right)\left(B_k \cos\left(\frac{k\pi t\mathfrak{c}}{\ell}\right) + C_k \sin\left(\frac{k\pi t\mathfrak{c}}{\ell}\right)\right), \quad k \in \mathbb{Z}.$$

We do not know the values of the constant factors A_k, B_k, and C_k at this point in time. The reason is that these will be determined by the initial conditions, and our mantra is TIDGLAS, the initial conditions go last.

2.1.2 Superposition: building a super solution!

The next step in the method of variable separation is called superposition. The functions u_k that we have found all satisfy a *homogeneous, linear, partial differential equation.* To see that the equation is homogeneous, we bring all the terms in the partial differential equation that contain the unknown function to one side:

$$u_{tt} - \mathfrak{c}^2 u_{xx} = 0.$$

On the other side of the equation is zero, and it is precisely for that reason that this equation is homogeneous. The equation is called linear because:

$$(u+v)_{tt} - \mathfrak{c}^2(u+v)_{xx} = u_{tt} - \mathfrak{c}^2 u_{xx} + v_{tt} - \mathfrak{c}^2 v_{xx}.$$

In other words, if a function v also satisfies the equation:

$$v_{tt} - \mathfrak{c}^2 v_{xx} = 0 \implies (u+v)_{tt} - \mathfrak{c}^2 (u+v)_{xx} = 0.$$

Another important observation is that the problem we are solving has *homogeneous boundary conditions.* The solution should satisfy

$$u(0,t) = u(\ell,t) = 0.$$

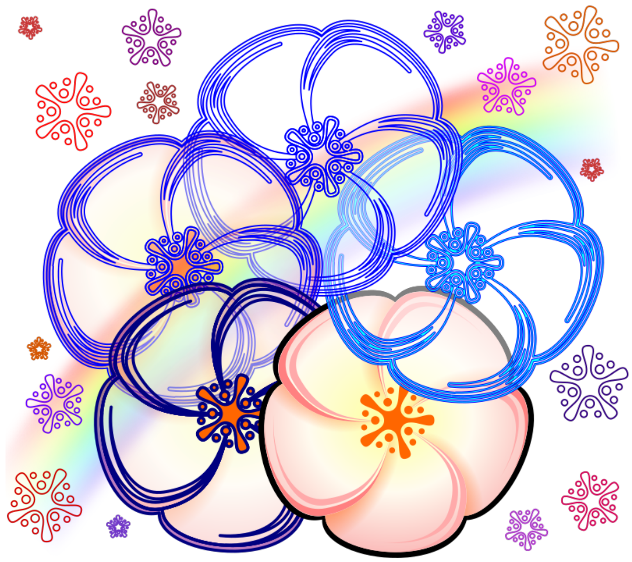


Figure 2.2: The process of superposition is similar to taking flowers that are the same shape but different sizes and colors and organizing them together. Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`

When we used separation of variables, we guaranteed this will hold for each of the functions

$$X_k(x) = \sin(k\pi x/\ell) \implies X_k(0) = X_k(\ell) = 0, \text{ and } u_k(0,t) = u_k(\ell,t) = 0.$$

So, if we add two of these together, the boundary condition is still satisfied:

$$u_j(0,t) + u_k(0,t) = 0 = u_j(\ell,t) + u_k(\ell,t).$$

Whenever one has solutions to a partial differential equation that has these properties, then the sum of any two solutions to the PDE is again a solution to the PDE, and it will again satisfy the nice homogeneous boundary condition. Adding solutions together is called superposition. One of the interesting differences between solving partial differential equations compared to solving ordinary differential equations is that we often find a lot more solutions - like infinitely many - to the PDE. When solving ODEs, the standard procedure is to find a basis of 'general solutions' and then use them to build a particular solution depending on the initial conditions. Now, solving PDEs, we will do the same, except now we use infinitely many 'general solutions.' We can visualize these solutions as having something in common, namely that they solve the PDE and the boundary conditions, but also something different, because they are all different. This is similar to the flowers in Figure 2.2 that are the same shape but different colors and sizes.

Consequently, we define

$$u(x,t) = \sum_{k \in \mathbb{Z}} u_k(x,t).$$

For any finite sum,

$$\left(\sum_{k=-N}^{M} u_k(x,t) \right)_{tt} - \mathfrak{c}^2 \left(\sum_{k=-N}^{M} u_k(x,t) \right)_{xx} = \sum_{k=-N}^{M} \partial_{tt} u_k(x,t) - \mathfrak{c}^2 \partial_{xx} u_k(x,t) = 0.$$

So, if we are lucky, then letting N and M go to infinity, the full sum will converge to define a function $u(x,t)$ that also satisfies the partial differential equation. The final pieces of the puzzle that we will need to find are the unknown coefficients, since

$$u_k(x,t) = X_k(x)T_k(t) = A_k \sin\left(\frac{k\pi x}{\ell}\right)\left(B_k \cos\left(\frac{k\pi t \mathfrak{c}}{\ell}\right) + C_k \sin\left(\frac{k\pi t \mathfrak{c}}{\ell}\right)\right), \quad k \in \mathbb{Z}.$$

These coefficients will be determined by the two initial conditions, the position of the string at time zero and its velocity

$$u(x,0) = f(x), \quad u_t(x,0) = g(x).$$

As per our mantra of the day, TIDGLAS, these initial conditions will be dealt with last. Our focus here is to hone the two skills: separation of variables and superposition. We will require additional techniques to complete the last step. Bruce Lee once said that it is more effective to practice one technique a hundred times than it is to practice a hundred techniques one time each. With this philosophy, we will remain focused for this chapter on practicing these two techniques and postpone learning the additional techniques required for the last step to the next chapter.

2.2 Heat flow in a circular rod

Consider a circular shaped rod, like a rod that's been bent into a circle. Let's mathematicize it! To specify points on the rod, we just need to know the angle at the point. For this reason, we use the real variable x for the position, where x gives us the angle at the point on the rod. We use the variable $t \geq 0$ for time. The function $u(x,t)$ is the temperature on the rod at position x at time t. We could imagine that this circular rod is like the wheel of a hot rod like in Figure 2.3 that has been racing around town and now is sitting in the garage.

The laws of physics can be used to prove that the temperature function satisfies the heat equation, that is the partial differential equation

$$u_t = k u_{xx},$$

for some constant $k > 0$. The temperatures of the rod at the initial time are given by some function,

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

Since there is no source adding heat to the rod nor sink taking heat away from the rod, the temperatures at all later times are determined by the temperatures at time $t = 0$ together with the fact that u must satisfy the heat equation. Although this may not be so obvious mathematically, it should make sense to physicists. Now, the fact that the rod is circular means that we can define the function $f(x)$ for all real x by simply demanding that

$$f(x + 2\pi) = f(x).$$

This is because the angles x and $x + 2\pi$ correspond to the exact same point on the circular rod. Similarly, the same is also true for the function u, that is

$$u(x + 2\pi, t) = u(x,t),$$

for all x and all $t > 0$.

We will use our two techniques, separation of variables and superposition, and see what solutions we obtain to this partial differential equations. We first use separation of variables. This starts by assuming that

$$u(x,t) = X(x)T(t).$$

Next, we put this into the heat equation:

$$T'(t)X(x) = kX''(x)T(t).$$

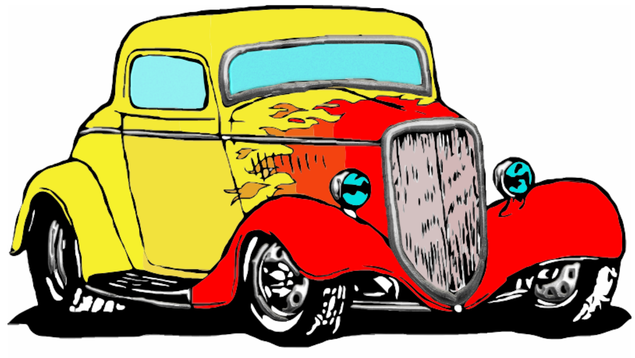


Figure 2.3: A *hot rod* like the one in this picture is a colloquial term for an American car that has been rebuilt or modified for more speed and acceleration. Often they are painted with flames. Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org

We want to *separate variables,* so we want all the t-dependent bits on the left say, and all the x-dependent bits on the right. This can be achieved by dividing both sides by $X(x)T(t)$,

$$\frac{T'(t)}{T(t)} = k\frac{X''(x)}{X(x)}.$$

We now know that both sides must be constant. Let us call the constant λ, so that

$$\frac{T'}{T} = \lambda = k\frac{X''}{X}.$$

Exercise 9. *In your own words, explain why both sides of the equation must be constant.*

As with the wave equation, we will solve for the function X first. The equation for X is:

$$X''(x) = \frac{\lambda}{k}X(x)$$

for a constant λ. Moreover, since we must have

$$u(x+2\pi,t) = X(x+2\pi)T(t) = u(x,t) = X(x)T(t) \quad \text{for all } x \text{ and } t,$$

this demands that

$$X(x+2\pi) = X(x) \text{ for all } x.$$

So, we have the additional information that X is a periodic function of period 2π. We will solve for all functions X and all constants λ that satisfy both the differential equation and the periodicity condition by consider the three possible cases for λ.

Exercise 10. ***Case 1:*** *Show that if $\lambda = 0$, there is no solution to $X''(x) = 0$ which is 2π periodic, other than the constant solutions.*

Case 2: If $\lambda > 0$, then a basis of solutions is,

$$\{e^{\sqrt{\lambda}x/\sqrt{k}}, e^{-\sqrt{\lambda}x/\sqrt{k}}\}.$$

So, we can write

$$X(x) = ae^{\sqrt{\lambda}x/\sqrt{k}} + be^{-\sqrt{\lambda}x/\sqrt{k}}.$$

For the 2π periodicity to hold, we need

$$X(0) = X(2\pi) \implies a+b = ae^{\sqrt{\lambda}2\pi/\sqrt{k}} + be^{-\sqrt{\lambda}2\pi/\sqrt{k}} \implies a(e^{\sqrt{\lambda}2\pi/\sqrt{k}} - 1) = b(1 - e^{-\sqrt{\lambda}2\pi/\sqrt{k}})$$

$$\implies a = b\frac{(1 - e^{-\sqrt{\lambda}2\pi/\sqrt{k}})}{e^{\sqrt{\lambda}2\pi/\sqrt{k}} - 1}.$$

We also need

$$X(-2\pi) = X(0) \implies a + b = ae^{-\sqrt{\lambda}2\pi/\sqrt{k}} + be^{\sqrt{\lambda}2\pi/\sqrt{k}} \implies a(e^{-\sqrt{\lambda}2\pi/\sqrt{k}} - 1) = b(1 - e^{\sqrt{\lambda}2\pi/\sqrt{k}})$$

$$\implies a = b\frac{1 - e^{\sqrt{\lambda}2\pi/\sqrt{k}}}{e^{-\sqrt{\lambda}2\pi/\sqrt{k}} - 1}.$$

So, we have two equations for a, therefore they should be equal:

$$a = b\frac{1 - e^{-\sqrt{\lambda}2\pi/\sqrt{k}}}{e^{\sqrt{\lambda}2\pi/\sqrt{k}} - 1} = b\frac{1 - e^{\sqrt{\lambda}2\pi/\sqrt{k}}}{e^{-\sqrt{\lambda}2\pi/\sqrt{k}} - 1}.$$

If $b = 0$ then $a = 0$ so the whole solution is the zero solution. If $b \neq 0$ then we must have

$$\frac{1 - e^{-\sqrt{\lambda}2\pi/\sqrt{k}}}{e^{\sqrt{\lambda}2\pi/\sqrt{k}} - 1} = \frac{1 - e^{\sqrt{\lambda}2\pi/\sqrt{k}}}{e^{-\sqrt{\lambda}2\pi/\sqrt{k}} - 1}.$$

Changing the sign of the top and bottom on the right side, this is equivalent to:

$$\frac{1 - e^{-\sqrt{\lambda}2\pi/\sqrt{k}}}{e^{\sqrt{\lambda}2\pi/\sqrt{k}} - 1} = \frac{e^{\sqrt{\lambda}2\pi/\sqrt{k}} - 1}{1 - e^{-\sqrt{\lambda}2\pi/\sqrt{k}}}.$$

Call the left side $\star$. Then the right side is $\frac{1}{\star}$. So the equation is

$$\star = \frac{1}{\star} \implies \star^2 = 1 \implies \star = \pm 1.$$

Exercise 11. *Show that* $\star > 0$*.*

If

$$\star = 1 \implies 1 - e^{-\sqrt{\lambda}2\pi/\sqrt{k}} = e^{\sqrt{\lambda}2\pi/\sqrt{k}} - 1 \implies 2 = e^{\sqrt{\lambda}2\pi/\sqrt{k}} + e^{-\sqrt{\lambda}2\pi/\sqrt{k}}.$$

I don't like the negative exponent thing (it is really a fraction), so I am going to multiply by $e^{\sqrt{\lambda}2\pi/\sqrt{k}}$. Also, doing this turns it into a quadratic equation:

$$2e^{\sqrt{\lambda}2\pi/\sqrt{k}} = e^{4\pi\sqrt{\lambda}/\sqrt{k}} + 1 \iff e^{4\pi\sqrt{\lambda}/\sqrt{k}} - 2e^{2\pi\sqrt{\lambda}/\sqrt{k}} + 1 = 0$$

Now we can factor this equation because the left side is

$$(e^{2\pi\sqrt{\lambda}/\sqrt{k}} - 1)^2 = 0 \implies e^{2\pi\sqrt{\lambda}/\sqrt{k}} = 1 \iff 2\pi\sqrt{\lambda}/\sqrt{k} = 0 \lightning.$$

That $\lightning$ indicates a contradiction. Therefore, in the case where $\lambda > 0$, the only solution which is 2π periodic is the zero solution.

Hence, we are left with **Case 3:** $\lambda < 0$. Then, a basis of solutions is

$$\{\sin(\sqrt{|\lambda|}x/\sqrt{k}), \cos(\sqrt{|\lambda|}x/\sqrt{k})\}.$$

We need these solutions to be 2π periodic. They will be as long as $\sqrt{|\lambda|}/\sqrt{k}$ is an integer. So we need

$$\lambda < 0, \quad \frac{\sqrt{|\lambda|}}{\sqrt{k}} = n \in \mathbb{Z} \implies \lambda_n = -n^2 k.$$

Hence, we can list all solutions as

$$X_n(x) = a_n \cos(nx) + b_n \sin(nx), \quad n \in \mathbb{Z}.$$

Exercise 12. *Show that allowing complex coefficients, it is equivalent to use a basis of solutions*

$$\{e^{inx}\}_{n\in\mathbb{Z}}.$$

Find A_n and B_n in terms of a_n and b_n so that

$$X_n(x) = A_n e^{inx} + B_n e^{-inx}.$$

Now, we can solve for the partner functions, $T_n(t)$, because $u_n(x,t) = X_n(x)T_n(t)$. The equation for T_n is

$$\frac{T_n'(t)}{T_n(t)} = \lambda_n = -n^2 k,$$

that re-arranges to

$$T_n'(t) = -n^2 k T_n(t).$$

Consequently, from calculus, the only solutions to this equation are

$$T_n(t) = e^{-n^2kt} \text{ times a constant}$$

Putting this together, we have

$$u_n(x,t) = X_n(x)T_n(t) = e^{-n^2kt}(a_n\cos(nx) + b_n\sin(nx)), \quad \text{for some constants } a_n \text{ and } b_n.$$

These solutions satisfy the heat equation

$$\partial_t u_n - k\partial_{xx}u_n = 0.$$

This is a linear and homogeneous partial differential equation. Consequently, we can use the superposition principle to smash all these solutions we have found into a super solution:

$$u(x,t) = \sum_{n\in\mathbb{Z}} u_n(x,t) = \sum_{n\in\mathbb{Z}} e^{-n^2tk}(a_n\cos(nx) + b_n\sin(nx)).$$

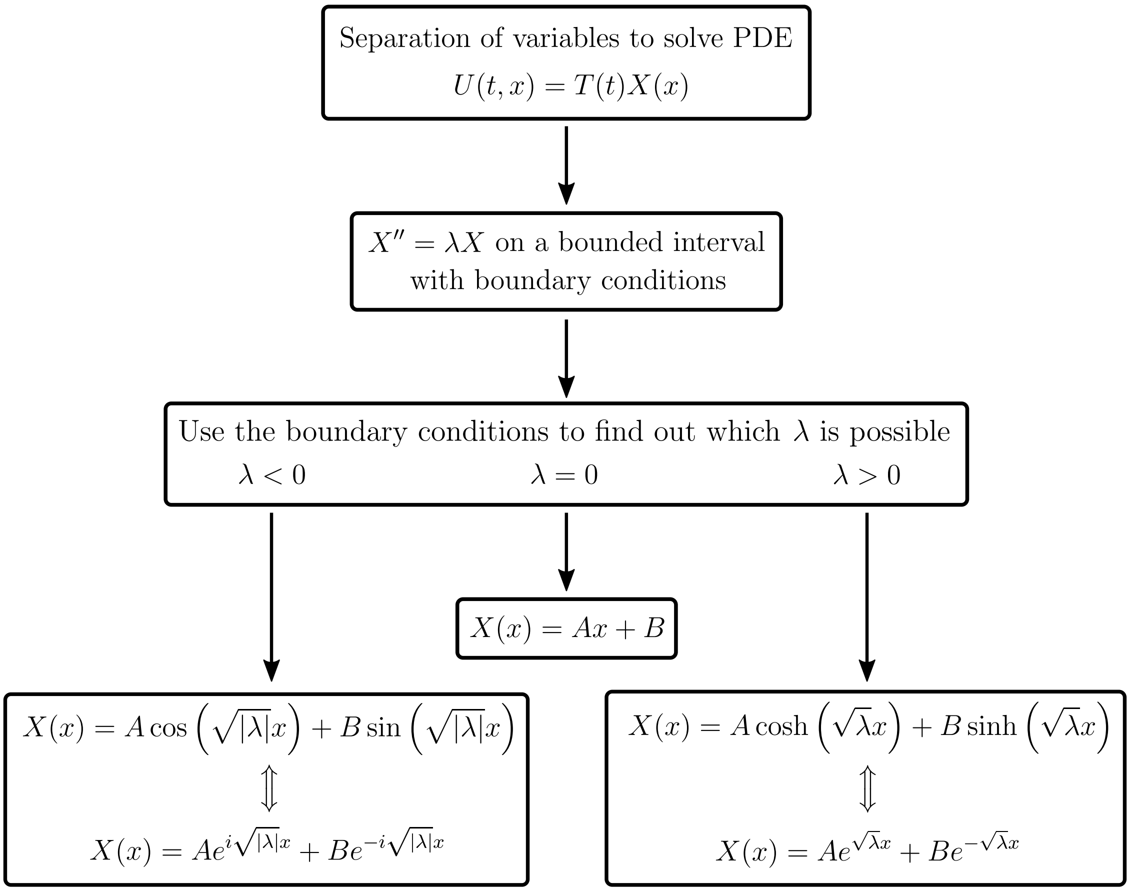
The coefficients will be determined by the initial condition, because we must have

$$u(x,0) = f(x) \implies \sum_{n\in\mathbb{Z}}(a_n\cos(nx) + b_n\sin(nx)) = f(x).$$

So, to complete the last step, we wish to find coefficients a_n and b_n so that, given the initial data $f(x)$ we can express $f(x)$ as such a series. In order to do this, we will represent functions as vectors inside infinite dimensional vector spaces. The idea is that we will view the above sines and cosines as basis vectors, and then we will project the initial data on these basis vectors to express it in terms of them. To be able to do this, we need to develop theory and tools for infinite dimensional vector spaces in which the vectors will be functions. That is just what we will do in the next chapter!

2.3 A schematic summary of using separation of variables to determine what type of functions X will comprise the solution

When we solve partial differential equations using separation of variables, we will follow the same general procedure. The first part of this procedure is to find the space-dependent part of the solution using the boundary conditions. This process is illustrated in Figure 2.4



Don't worry about finding the coefficients A and B.

Figure 2.4: When we use separation of variables to solve partial differential equations, the equation at which we typically arrive for the space-dependent function, $X(x)$, is that the second derivative of X is equal to a constant times X. We follow tradition by naming the constant the Greek letter, lambda, λ, so the equation is $X''(x) = \lambda X(x)$. There are three cases for λ: it could be positive, negative, or zero. According to these three cases, there are three possible types of solutions X as shown above. It is crucial to use the boundary conditions to determine which values λ guarantee that X satisfies the boundary conditions. Simply put: it is the boundary conditions that will narrow down the possible values of λ, usually there will be infinitely many, that you can index with integers. This figure was jointly created with Gottfrid Olsson.

2.3.1 Indications that variable separation and superposition can help solve a PDE.

When will the technique of separating variables and then using superposition of the solutions obtained lead to an actual solution of the partial differential equation at hand? Here are two characteristics that indicate separation of variables and superposition can be helpful:

- linearity in the partial differential equation as well as the methods,
- homogeneity in the partial differential equation and the boundary condition.

Linearity means for example:

$$\frac{d}{dx}(f(x)+g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x).$$

This means that we can add the functions first and then differentiate, or differentiate each function separately and then add them, and the results are the same. This is also true for integration:

$$\int f+g = \int f + \int g.$$

On the other hand

$$\frac{d}{dx}(f(x)g(x)) \neq \frac{d}{dx}f(x)\frac{d}{dx}g(x),$$

because the product rule says that

$$\frac{d}{dx}(f(x)g(x)) = f(x)\frac{d}{dx}g(x) + g(x)\frac{d}{dx}f(x).$$

So, it's not the same to multiply the functions first and then differentiate versus differentiating first and then multiplying. To solve a PDE using separation of variables and superposition, the partial differential equation should be *linear.* This means that if we stuff $(u+v)$ into the PDE, the result will be equal to the result of first stuffing u into the PDE, stuffing v into the PDE, and then adding the result together. We will also need the PDE to be homogeneous, so that when we stuff our super solution into the PDE, each of the individual terms satisfies PDE-of-each-term is zero. So the PDE applied to the whole supersolution-sum is just adding up a bunch of zeros and is again equal to zero. Hence our supersolution satisfies the PDE. We will typically need to use both of these techniques, separation of variables to obtain all possible solutions, and then superposition, adding all of these together with their as-of-now-unknown-coefficients. The reason we must *collect them all,* like Pokemon, is because we will use them *all* to create an orthogonal base for a Hilbert space. We will express the initial data using this orthogonal base and thereby determine the coefficients.

Homogeneity in simple terms means 'things are equal to zero.' For the PDE, when we take all of the terms involving the unknown function to one side of the equation, traditionally the left side, then right side of the equation should equal to zero. Examples of this include

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

that is the homogeneous wave equation with the constant $c = 1$, as well as

$$u_t = u_{xx} \iff u_t - u_{xx} = 0,$$

that is the homogeneous heat equation with the constant $k = 1$. It is also important that we have homogeneous boundary conditions for the unknown function. A homogeneous boundary condition is a set of equations for the unknown function, and possibly some of its derivatives as well, that should vanish. Examples include the Dirichlet boundary condition that demands the unknown function vanishes at the boundary, as well as the Neumann boundary condition that demands the spatial (normal) derivative of the unknown function vanishes at the boundary. Homogeneous boundary conditions are important for the method of superposition, because if we add solutions that satisfy such a boundary condition, the sum will also satisfy the boundary condition because $0+0=0$.

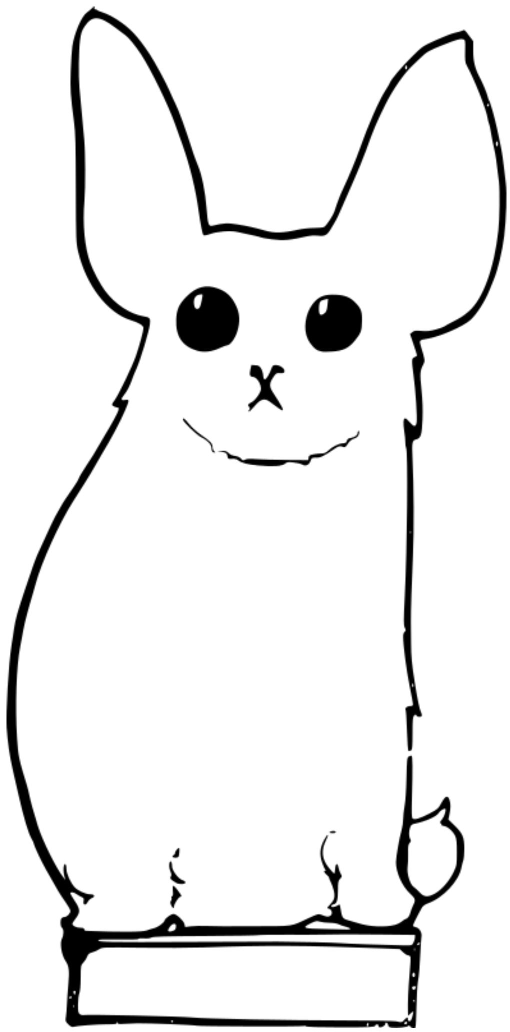


Figure 2.5: When we use separation of variables and superposition to solve a PDE, we will generally need to collect *all* of the solutions, like collecting all Pokemon. This is a public domain image from 1899 that remarkably looks like the Pokemon Pikachu, reminding us to collect all the solutions, because we will use them to create an orthogonal base for a Hilbert space!

2.4 Partial differential equations in mathematical physics

Mathematical physics is the development of mathematical methods for applications to problems in physics. The author of this text is an active researcher in this field, working to develop mathematical methods that apply to problems in physics. One journal that specializes in this topic is the Journal of Mathematical Physics, that defines the field as *the application of mathematics to problems in physics and the development of mathematical methods suitable for such applications and for the formulation of physical theories.* Here we will discuss the basic partial differential equations of classical mathematical physics. One of the most ubiquitous partial differential operators in mathematical physics is the *Laplace operator,* that can be defined in $\mathbb{R}^n$ for all $n = 1, 2, 3, \ldots$ For functions of n real variables, this differential operator is written as ∇^2, or div ∇, or Δ. Different strokes for different folks; that is to say some people prefer a specific notation more than another. Here we will likely use

$$\Delta := \sum_{k=1}^{n} \frac{\partial^2}{\partial x_k^2}.$$

There are three important types of linear PDEs in mathematical: elliptic, parabolic, and hyperbolic. What about *nonlinear* PDEs? These are notoriously difficult to solve, but one method to analyze them is to *linearize* the equation, by studying a linear PDE that approximates the nonlinear one. Consequently, linear PDEs are of fundamental importance to the study of all PDEs in mathematical physics, including nonlinear PDEs. The PDEs which are amenable to solutions via Fourier methods are in general linear and fall into these three classifications. We give a brief collection of the most common equations of these three types.

2.4.1 Elliptic partial differential equations in mathematical physics

$\Delta u = 0$	Laplace's equation
$\Delta u + \lambda^2 u = 0$	Helmholtz's equation, also known as the Laplace eigenvalue equation
$\Delta u = k$	Poisson's equation
$\Delta u + k(E - V)u = 0$	time independent Schrödinger's equation

Table 2.1: The PDEs above are some of the most common elliptic PDEs in mathematical physics.

The poster child elliptic PDE is the *Laplace equation*

$$\Delta u = 0.$$

This equation is satisfied by the electrostatic potential in any region containing no electric charge as well as by the gravitational potential in any region containing no mass. Solutions of the heat and wave equations that are independent of time, known as *steady state solutions,* also satisfy this equation.

2.4.2 Parabolic partial differential equations in mathematical physics

$u_t = k u_{xx}$	one dimensional diffusion equation (heat equation in one dimension)
$u_t = k \Delta u$	heat equation in $\mathbb{R}^n$
$u_t = k u_{xx} - h u_x$	diffusion-convection equation
$u_t = k u_{xx} - k u$	diffusion with lateral heat-concentration loss
$u_t = k u_{xx} + f(x, t)$	diffusion with heat source (or sink)
$i\hbar u_t = -\frac{\hbar^2}{2m} \Delta u + V(\mathbf{x}) u$	time dependent Schrödinger equation

Table 2.2: The PDEs above are some of the most common parabolic PDEs in mathematical physics.

The poster child example of a parabolic PDE is the homogeneous heat equation

$$u_t = k\Delta u.$$

This equation describes the diffusion of heat in a homogeneous material. Here $u(t, \mathbf{x})$ is the temperature at the point $\mathbf{x}$ in the material at the time t. The constant k is the thermal diffusivity of the material. The heat equation is not a fundamental law of physics, as can be seen by the fact that if we add a constant to a solution of the heat equation, the resulting function also satisfies the heat equation. This shows that the heat equation cannot detect the physical fact of absolute zero temperature. Consequently, solutions of the heat equation for extreme temperatures (very high or very low) does not give physically reliable answers.

Another interesting example from quantum mechanics is the time-dependent Schrödinger equation

$$i\hbar u_t = -\frac{\hbar^2}{2m}\Delta u + V(\mathbf{x})u.$$

Here u is the quantum mechanical wave function for a particle of mass m moving in a potential $V(\mathbf{x})$, $\hbar$ is Planck's constant, and $i = \sqrt{-1}$. The physical derivation of the Schrödinger equation is contained in books on quantum mechanics like [15] and [10].

2.4.3 Hyperbolic partial differential equations in mathematical physics

$u_{tt} = \mathfrak{c}^2 u_{xx}$	one dimensional vibrating string (one dimensional wave equation)
$u_{tt} = \mathfrak{c}^2 \Delta u$	wave equation in $\mathbb{R}^n$
$u_{tt} = \mathfrak{c}^2 u_{xx} - hu_t$	vibrating string with friction
$u_{tt} = \mathfrak{c}^2 \Delta u - hu_t$	wave equation with friction in $\mathbb{R}^n$
$u_{tt} = \mathfrak{c}^2 u_{xx} - hu_t - ku$	telegraph equation (also known as transmission line equation)
$u_{tt} = \mathfrak{c}^2 u_{xx} + f(x,t)$	wave equation with forced vibrations

Table 2.3: The PDEs above are some of the most common hyperbolic PDEs in mathematical physics.

The poster child hyperbolic PDE is the homogeneous wave equation

$$u_{tt} = c^2 \Delta u.$$

The function $u(t, \mathbf{x})$ depends on time, t, and $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. It represents a wave traveling through an n-dimensional medium. For $n = 1$, this could be a simplification of a string, where we view the string as a one-dimensional object. The constant c is the speed of propagation of waves in the medium, and $u(t, \mathbf{x})$ is the amplitude of the wave at position $\mathbf{x}$ and time t. In the example of a vibrating string, we can view $u(t, x)$ as the height of the string at the point x at the time t. The wave equation is a mathematical model for several physical phenomena including vibrating strings, drums, air, waves in water, sound waves in air, and electromagnetic waves including light and radio waves. The wave equation can be derived from the laws of physics; see [4, Appendix 1].

2.4.4 Don't let the solution get away; the boundary conditions are here to stay!

When solving these and other partial differential equations of mathematical physics, there is most often a boundary condition and an initial condition. Although it may seem strange to deal with the initial condition *last,* according to our mantra TIDGLAS, it is often best to address the boundary conditions first and save the initial conditions for the end of the solution. The boundary conditions could arise, for example with a vibrating guitar string, due to the physical phenomena we consider, and the geometry of the problem. In this example, the ends of the strings do not move, and that translates into the Dirichlet boundary condition, which requires the solution of the PDE to vanish at the boundary. The boundary conditions are an extremely important consideration and should not be forgotten, lest your guitar strings go wild and crazy like in Figure 2.6. Here is a mnemonic to remind you to always pay attention to the boundary condition.

Don't let the solution run away; the boundary conditions are here to stay!

Another natural boundary condition is the Neumann boundary condition, which requires the normal derivative of the solution of the PDE to vanish at the boundary. This is the boundary condition we would use to solve the heat equation in a bounded region that has an insulated boundary, because this condition means precisely that there is no exchange (gain or loss) of heat across the boundary.

2.5 Exercises

1. [4, 1.1.2] Show that $u(x, y, t) = t^{-1}e^{-(x^2+y^2)/(4kt)}$ satisfies the heat equation $u_t = k(u_{xx} + u_{yy})$ for $t > 0$.

2. [4, 1.2.5(c)] Show that for $n = 1, 2, 3, \ldots$ $u_n(x, y) = \sin(n\pi x)\sinh(n\pi(1 - y))$ satisfies
$$u_{xx} + u_{yy} = 0, \quad u(0, y) = u(1, y) = u(x, 1) = 0.$$

3. [4, 1.3.1] Derive pairs of ordinary differential equations from the following partial differential equations by separation of variables or show that it is not possible:
   (a) $yu_{xx} + u_y = 0$
   (b) $x^2u_{xx} + xu_x + u_{yy} + u = 0$
   (c) $u_{xx} + u_{xy} + u_{yy} = 0$
   (d) $u_{xx} + u_{xy} + u_y = 0$.

4. [4, 1.3.6] Use separation of variables to obtain the family of solutions
$$u^{\pm}_{mn}(x, y, z) = \sin(m\pi x)\cos(n\pi y)e^{\pm\sqrt{m^2+n^2}\pi z},$$
to the problem
$$\nabla^2 u = 0, \quad u(0, y, z) = u(1, y, z) = u_y(x, 0, z) = u_y(x, 1, z) = 0.$$

5. [4, 1.1.1] Show that for $t > 0$, $u(x, t) = t^{-1/2}e^{-x^2/(4kt)}$ satisfies the heat equation
$$u_t = ku_{xx}.$$

6. [4, 1.2.5(a)] Show that for $n = 1, 2, 3, \ldots$ $u_n(x, y) = \sin(n\pi x)\sinh(n\pi y)$ satisfies
$$u_{xx} + u_{yy} = 0, \quad u(0, y) = u(1, y) = u(x, 0) = 0.$$

7. [4, 1.3.5] By separation of variables, derive the solutions $u_n(x, y) = \sin(n\pi x)\sinh(n\pi y)$ of
$$u_{xx} + u_{yy} = 0, \quad u(0, y) = u(1, y) = u(x, 0) = 0.$$

8. [4, 1.3.7] Use separation of variables to find an infinite family of independent solutions to
$$u_t = ku_{xx}, \quad u(0, t) = 0, \quad u_x(\ell, t) = 0,$$
representing heat flow in a rod with one end held at temperature zero and the other end insulated.

9. [4, 1.1.3] Show that $u(x, y) = \log(x^2 + y^2)$ satisfies Laplace's equation $u_{xx} + u_{yy} = 0$ for $(x, y) \neq (0, 0)$.

Figure 2.6: I cannot emphasize enough how important boundary conditions are when solving partial differential equations! Changing the boundary conditions can *completely* change the problem. This figure shows a forget-me-not fluffball, known as a katodanode, playing a guitar with forgotten boundary conditions. The strings are going wild! This little katodanode wants you to remember: *don't let the solution run away; the boundary conditions are here to stay!* Ebba Grönfors & Alva Brycke invented these forget-me-knot-fluffballs to help remember important concepts. That is a a clever mnemonic technique, and we are grateful to Ebba & Alva for contributing!

10. [4, 1.1.4] Show that $u(x,y,z) = (x^2+y^2+z^2)^{-1/2}$ satisfies Laplace's equation $u_{xx}+u_{yy}+u_{zz} = 0$ for $(x,y,z) \neq (0,0,0)$.

11. [4, 1.1.5] Mathematicians often say that we can change the units of time in order to simplify the equations we obtain from physics. Show that if we replace t with $\tau := kt$ then the heat equation $u_t = ku_{xx}$ becomes
$$u_\tau = u_{xx}.$$
Show that if we instead replace t with $\tau = ct$ then the wave equation $u_{tt} = c^2 u_{xx}$ becomes
$$u_{\tau\tau} = u_{xx}.$$

12. [4, 1.1.6] The goal of this exercise is to obtain *d'Alembert's formula* for the solution of the one-dimensional wave equation $u_{tt} = c^2 u_{xx}$. The wave operator is also often known as the *d'Alembertian,* similar to how the Laplace operator (named after Pierre Simon Laplace) is called the Laplacian.
    (a) Show that if $u(y,z) = f(y) + g(z)$ where f and g are C^2 functions of one variable, then $u_{yz} = 0$. Conversely, show that every C^2 solution u of $u_{yz} = 0$ is of this form.
    (b) Let $y = x - ct$ and $z = x + ct$. Show that $u_{tt} - c^2 u_{xx} = -4c^2 u_{yz}$.
    (c) Conclude that the general C^2 solution of the wave equation $u_{tt} = c^2 u_{xx}$ is $u(x,t) = \phi(x-ct) + \psi(x+ct)$ where ϕ and ψ are C^2 functions of one variable. Note that $\phi(x-ct)$ represents a wave traveling to the right with speed c, and $\psi(x+ct)$ represents a wave traveling to the left with speed c.
    (d) Show that the solution of the initial value problem
    $$u_{tt} = c^2 u_{xx}, \quad u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad t > 0, \quad x \in \mathbb{R}$$
    is
    $$u(x,t) = \frac{1}{2}\left[f(x-ct) + f(x+ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(y)dy.$$

13. [4, 1.1.7] The voltage v and current i in an electrical cable along the x-axis satisfy the equations
$$i_x + Cv_t + Gv = 0, \quad v_x + Li_t + Ri = 0,$$
where $C, G, L,$ and R are the capacitance, (leakage) conductance, inductance, and resistance per unit length in the cable. Show that v and i both satisfy the *telegraph equation,*
$$u_{xx} = LCu_{tt} + (RC + LG)u_t + RGu.$$

14. [4, 1.2.2] Consider the nonlinear differential equation $u' = u(1-u)$.
    (a) Show that $u_1(x) = e^x(1+e^x)^{-1}$ and $u_2(x) = 1$ are solutions.
    (b) Show that $u_1 + u_1$ is not a solution.
    (c) For which values of $c \in \mathbb{R}$ is cu_1 a solution? For which values of c is cu_2 a solution?

Chapter 3

Don't get lost in a Hilbert space; find your way with an orthogonal base!

We still do not know how to obtain the coefficients to complete the solutions we have obtained using separation of variables and superposition! Is the suspense killing you? I hope not, because it's going to last for another chapter! To obtain these coefficients and rigorously solve the wave and heat equations, and more generally for solving partial differential equations, *function spaces* play a central role. These are sets of functions that can be mathematicized as *vectors* in the sense that we can view a function as an element of a vector space. These vector spaces will usually be infinite dimensional. This is beautiful because the more dimensions one has, the greater range for creativity, and the more places one can go! For example, if you live in a one dimensional space, you only have two directions: forwards and backwards. Imagine living in such a space, and imagine that there are two other people who live in the same space, but you really don't like them. If they are on either side of you, you are stuck, with no egress! On the other hand, as soon as we increase to a two dimensional universe, there are now infinitely many directions. However, we only need two *basis* directions in order to express all directions. In three dimensions, we need three basis directions to express all directions; two is not enough. For infinite dimensional Hilbert spaces, we will need infinitely many basis vectors, but the fundamental idea of an orthogonal base is the same in all cases, be it finite or infinite dimensional.

In the previous chapter, we used separation of variables and superposition to obtain solutions to the wave and heat equations that were expressed as infinite sums of certain functions, with unknown coefficients for the terms in the sum. To determine those unknown coefficients, we will learn how to mathematicize functions as elements in vector spaces. Then, similar to linear algebra, we can express functions in terms of basis vectors. This is the procedure we will use to finally complete the solutions to the wave and heat equations.

Definition 13 (Banach space)**.** A Banach space is a complete, normed vector space.

Recall that a vector space is a non-empty set that is closed under addition and multiplication by scalars. If we refer to the elements of the vector space as vectors, this means that we can add two vectors, and the result is a vector. Similarly, we can also multiply a vector by a scalar (a real or complex number), and the result is also a vector. Finally, we can talk about the norm of a vector, that is geometrically interpreted as its length. With this interpretation it is natural that the norm satisfies:

$$||v|| \geq 0 \text{ for all vectors, with equality if and only if v=0,}$$

$$||cv|| = |c|||v|| \text{ for all scalars } c,$$

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

The last inequality above is known as the triangle inequality. The n dimensional Euclidean space is a Banach space, and the triangle inequality says that the sum of the lengths of two sides of a triangle is greater than or equal to the length of the third side.

Figure 3.1: This is a quote from Buzz Lightyear, a character in the Pixar animated film *Toy Story*. It is a very good and cute film. Having more dimensions means having more directions in which one can go. With so many directions one might be concerned about the possibility of getting lost. In a Hilbert space, one can specify when directions are perpendicular or not. This helps one to better navigate Hilbert spaces and is one of the reasons that although these spaces can be infinite dimensional, they are very pleasant to work within. Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`.

Definition 14 (Hilbert space)**.** A Hilbert space is a Banach space that is further equipped with a scalar product. If we denote the space by H then the scalar product takes two elements from the Hilbert space and spits out a complex number. The scalar product of u with v is denoted by $\langle u, v\rangle$. Then,

$$u, v \in H \implies \langle u, v\rangle \in \mathbb{C},$$

$$c \in \mathbb{C} \implies \langle cu, v\rangle = c\langle u, v\rangle,$$

$$u, v, w \in H \implies \langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle,$$

$$\langle u, v\rangle = \overline{\langle v, u\rangle},$$

$$\langle u, u\rangle \geq 0, \quad = 0 \iff u = 0.$$

The norm in a Hilbert space is defined by

$$||u|| := \sqrt{\langle u, u\rangle}.$$

Exercise 15. *Verify that $\mathbb{R}^n$ is a Hilbert space by defining the scalar product in the usual way. Show that the same is true for $\mathbb{C}^n$ if we now define*

$$\langle u, v\rangle = \sum_{k=1}^{n} u_k \overline{v_k}, \quad u = (u_1, u_2, \ldots, u_n), \quad v = (v_1, v_2, \ldots, v_n).$$

Definition 16 (Orthonormal basis (ONB) and Orthogonal basis (OB))**.** A set of elements contained in a Hilbert space, H,

$$\{u_\alpha\} \subset H$$

is an *orthonormal basis (ONB)* for H if for any $v \in H$ there exist complex numbers (c_α) such that

$$v = \sum c_\alpha u_\alpha, \quad \langle u_\alpha, u_\beta\rangle = \delta_{\alpha,\beta} = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta. \end{cases}$$

This is the Kronecker δ, that is defined to be

$$\delta_{\alpha,\beta} := \begin{cases} 0 & \alpha \neq \beta \\ 1 & \alpha = \beta. \end{cases}$$

A set of elements contained in a Hilbert space H

$$\{v_\alpha\} \subset H$$

is an orthogonal basis (OB) if for any $v \in H$ there exist complex numbers (b_α) such that

$$v = \sum b_\alpha v_\alpha, \quad \langle v_\alpha, v_\beta \rangle = 0 \quad \forall \alpha \neq \beta.$$

The difference between an orthonormal basis in comparison to an orthogonal basis is that the vectors all have length one. Sometimes this is convenient, sometimes it is a bit of a nuisance to insist on this, hence we are prepared to work with ONBs and OBs. Why haven't we written an index for α? We are not just being lazy, there is a mathematical reason: it is because we do not know how many vectors are in the basis. There could be finitely many or infinitely many, and in fact, there could be uncountably many. For the applications relevant to solving partial differential equations. all Hilbert spaces we shall encounter will have at most countably many basis vectors. This means we will be able to index the elements of a basis by $\mathbb{N}$ or equivalently by $\mathbb{Z}$. The dimension of a Hilbert space is the number of elements in an ONB. Any finite dimensional Hilbert space is in bijection with the standard one

$$\mathbb{C}^n, \quad u, v \in \mathbb{C}^n \implies \langle u, v \rangle = u \cdot \overline{v}.$$

Thus, writing

$$u = (u_1, \ldots, u_n), \quad \text{with each component } u_k \in \mathbb{C},\ k = 1, \ldots, n$$

and similarly for v,

$$\langle u, v \rangle = \sum_{k=1}^{n} u_k \overline{v_k}.$$

The bijection between any finite (n) dimensional Hilbert space and $\mathbb{C}^n$ comes from taking an ONB of the Hilbert space and mapping the elements of the ONB to the standard basis vectors of $\mathbb{C}^n$. Here are some useful basic results for Hilbert spaces.

The elements of an orthonormal (or orthogonal) basis help us to navigate our way through a Hilbert space. We will keep this in mind with the slogan for this chapter:

> Don't get lost in a Hilbert space, find your way with an an orthogonal base!

An orthogonal basis or base is a collection of elements (vectors) in a Hilbert space that are all orthogonal, such that every vector in that Hilbert space can be expressed as a linear combination of the basis vectors. Even though we will often be working in the rather exotic sounding infinite dimensional Hilbert spaces, let's just think about navigating our way around in an everyday sense, like using a compass as shown in Figure 3.2.

3.1 Cauchy-Schwarz Inequality, Triangle Inequality, and Pythagorean Theorem

In this section we will prove several facts concerning the scalar product and the norm of vectors in Hilbert spaces. We begin by proving a useful fact about the scalar product that we will use repeatedly.

Proposition 17. *Let H be a Hilbert space. For any u and v in H,*

$$||u + v||^2 = ||u||^2 + 2\,\mathrm{Re}\langle u, v \rangle + ||v||^2.$$

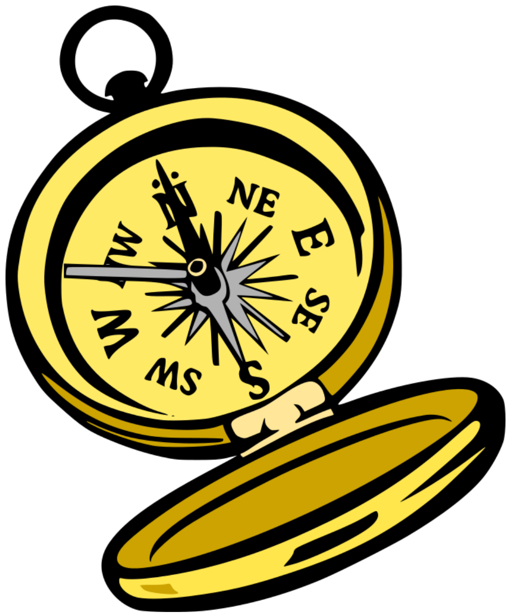


Figure 3.2: Don't get lost in a Hilbert space, find your way with an an orthogonal base! To navigate our way around the Earth, we just need to know which way is north, because then we can deduce east, west, and south! A compass is a light weight magnet, that is magnetized so that the southern pole of the needle is attracted to the Earth's magnetic north pole. Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

Proof: This is a good exercise because it only requires the definitions. First, by definition:

$$\begin{aligned}||u+v||^2 = \langle u+v, u+v\rangle &= \langle u, u+v\rangle + \langle v, u+v\rangle \\ &= \langle u,u\rangle + \langle u,v\rangle + \langle v,v\rangle + \langle v,u\rangle \\ &= ||u||^2 + \langle u,v\rangle + ||v||^2 + \overline{\langle u,v\rangle}.\end{aligned}$$

For any complex number z, adding z to its complex conjugate results in twice the real part of z:

$$z + \overline{z} = 2\,\mathrm{Re}(z).$$

So,

$$\langle u,v\rangle + \overline{\langle u,v\rangle} = 2\,\mathrm{Re}\langle u,v\rangle.$$

We indicate the end of a proof with a paw print, because we could all stand to be reminded of cute animals once in a while. When you see this you can imagine it is an animal friend saying, all right, you completed the proof, that's enough work, now it's time to take a break and pay attention to me! Meow! Woof! If you also have animal friends, then this symbol indicates that you should give your animal friend a pet, you and they deserve it, and petting animal friends has been shown to be a great stress reliever, so it's a great way to follow up a tough proof! For those who don't have animal friends to pet, contributors to this text are sharing cute animal pictures that will be included here as a virtual substitute. The first appearance is my cat, Ada, saying hello in Figure 3.3.

The next fact is so important that it is named after not only one but *two* mathematicians!

Proposition 18 (Cauchy-Schwartz Inequality)**.** *For any Hilbert space, H, for any u and v in H,*

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

Figure 3.3: Ada the cat sends her catly greetings and wishes you success with this text, as does the nice man reflected in the guitar!

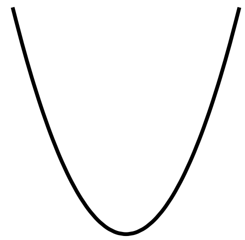


Figure 3.4: An upward quadratic function always has a unique minimum.

Proof: Assume that at least one of the two is non-zero. Let's assume $v \neq 0$, because otherwise we can just swap their names. We begin by considering the length of the vector u plus v scaled by a factor of t. If $t \to 0$, the length tends to $||u||^2$. What happens for other values of t? We compute it:

$$||u + tv||^2 = ||u||^2 + 2t\,\mathrm{Re}\langle u, v\rangle + t^2||v||^2, \quad t \in \mathbb{R}.$$

This is a real valued function of t. It's a quadratic function of t in fact. The derivative is

$$2t||v||^2 + 2\,\mathrm{Re}\langle u, v\rangle.$$

It's an upwards shaped quadratic function, similar to the graph in Figure 3.4. As seen there, this function has a unique minimum. Since the derivative must vanish at this minimum, we compute that the minimum occurs when

$$t = -\frac{\mathrm{Re}\langle u, v\rangle}{||v||^2}.$$

If we then check out what happens at this value of t,

$$||u + tv||^2 = ||u||^2 - 2\frac{\mathrm{Re}\langle u, v\rangle}{||v||^2}\,\mathrm{Re}\langle u, v\rangle + \mathrm{Re}\langle u, v\rangle^2\frac{||v||^2}{||v||^4} = ||u||^2 - \frac{\mathrm{Re}\langle u, v\rangle^2}{||v||^2}.$$

We know that

$$0 \leq ||u + tv||^2$$

so we get

$$0 \leq ||u||^2 - \frac{\mathrm{Re}\langle u, v\rangle^2}{||v||^2} \implies 0 \leq ||u||^2||v||^2 - \mathrm{Re}\langle u, v\rangle^2.$$

This gives us

$$\mathrm{Re}\langle u, v\rangle^2 \leq ||u||^2||v||^2.$$

Well, this is annoying because of that silly Re. I wonder how we could make it turn into $|\langle u, v\rangle|$? Also, we don't want to screw up the $||u||^2||v||^2$ part. Well, we know how the scalar product interacts with complex numbers, for $\lambda \in \mathbb{C}$,

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

So, if for example

$$\langle u, v\rangle = re^{i\theta}, r = |\langle u, v\rangle| \text{ and } \theta \in \mathbb{R}.$$

We can modify u, without changing $||u||$,

$$||e^{-i\theta}u|| = ||u||.$$

Moreover

$$\langle e^{-i\theta}u, v\rangle = e^{-i\theta}\langle u, v\rangle = e^{-i\theta}re^{i\theta} = |\langle u, v\rangle|\,.$$

So, if we repeat everything above replacing u with $e^{-i\theta}u$ we get

$$\mathrm{Re}\langle e^{-i\theta}u, v\rangle^2 \leq ||e^{-i\theta}u||^2||v||^2 = ||u||^2||v||^2,$$

and by the above calculation

$$\langle e^{-i\theta}u, v\rangle = |\langle u, v\rangle| \in \mathbb{R} \implies \mathrm{Re}\langle e^{-i\theta}u, v\rangle^2 = |\langle u, v\rangle|^2 .$$

So, we have

$$|\langle u, v\rangle|^2 \leq ||u||^2||v||^2.$$

Taking the square root of both sides completes the proof of the Cauchy-Schwarz inequality.

Since Hilbert spaces distinguish orthogonality between vectors, there is a triangle inequality, just like we have in the real world. The sum of the lengths of two sides of a triangle is greater than or equal to the third side of the triangle.

Proposition 19 (Triangle Inequality)**.** *For any u and v in a Hilbert space H,*

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

Proof: We just use the previous two results:

$$||u + v||^2 = ||u||^2 + 2\,\mathrm{Re}\langle u, v\rangle + ||v||^2 \leq ||u||^2 + 2||u||||v|| + ||v||^2 = (||u|| + ||v||)^2.$$

Taking the square root we obtain the triangle inequality.

We also have a Pythagorean theorem that says that the sum of the squares of the lengths of the legs of a right triangle is equal to the square of the hypotenuse; one usually learns this as $a^2 + b^2 = c^2$, as shown in Figure 3.5.

Proposition 20 (Pythagorean theorem)**.** *If u and v are orthogonal, then*

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

Moreover, if $\{u_n\}_{n=1}^N$ are orthogonal, then

$$||\sum_{n=1}^N u_n||^2 = \sum_{n=1}^N ||u_n||^2.$$

Proof: The first statement follows from

$$||u + v||^2 = ||u||^2 + 2\,\mathrm{Re}\langle u, v\rangle + ||v||^2 = ||u||^2 + ||v||^2,$$

if u and v are orthogonal, because in that case their scalar product is zero. Moreover, for any collection of orthogonal vectors $\{u_1, \dots, u_n\}$ we proceed by induction. Assume that

$$||u_1 + \ldots + u_{n-1}||^2 = \sum_{k=1}^{n-1} ||u_k||^2.$$

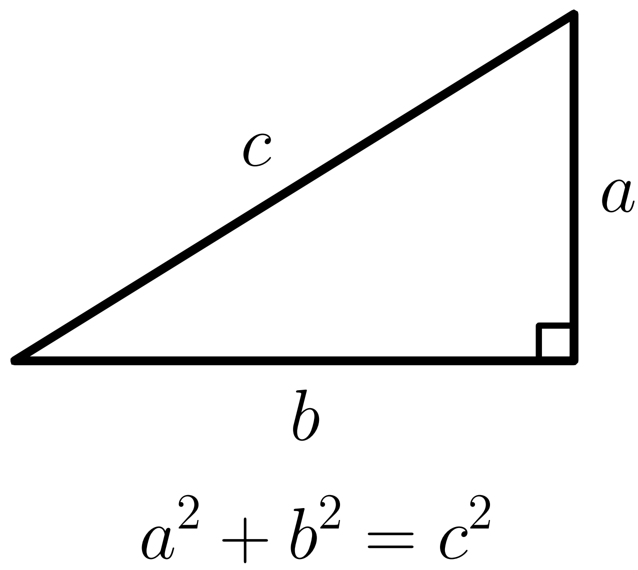


Figure 3.5: The Pythagorean theorem reminds us that the sum of the squares of the lengths of the two perpendicular sides of a right triangle is equal to the the square of the length of the third (longest) side. The two perpendicular sides of a right triangle are called legs, while the third longest side is called the hypotenuse. Thanks to Gottfrid Olsson for contributing this nice illustration!

Then, if u_n is orthogonal to all of $u_1, \ldots, u_{n-1}$ we also have

$$\langle u_n, u_1 + \ldots + u_{n-1} \rangle = \langle u_n, u_1 \rangle + \ldots + \langle u_n, u_{n-1} \rangle = 0 + \ldots + 0.$$

Hence u_n is also orthogonal to the sum,

$$\sum_{k=1}^{n-1} u_k.$$

By the Pythagorean theorem,

$$||u_n + \sum_{k=1}^{n-1} u_k||^2 = ||u_n||^2 + ||\sum_{k=1}^{n-1} u_k||^2.$$

By the induction assumption

$$= ||u_n||^2 + \sum_{k=1}^{n-1} ||u_k||^2 = \sum_{k=1}^{n} ||u_k||^2.$$

3.1.1 Continuity of the scalar product

An important fact that we will often use is that the scalar product is continuous. This means that we can move the scalar product in and out of infinite, convergent sums.

Proposition 21. *Using only the assumptions that the scalar product satisfies:*

$$\langle u, v \rangle = \overline{\langle v, u \rangle}$$

$$\langle au, v \rangle = a \langle u, v \rangle$$

$$\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$$

$$\langle u, u\rangle \geq 0, \quad \langle u, u\rangle = 0 \iff u = 0,$$

then the scalar product is a continuous function from $H \times H \to \mathbb{C}$.

Proof: All statements except the continuity can be demonstrated using the definition of the scalar product, so these are left as an exercise. We focus on the proof of the continuity, because at first glance, it seems like a rather amazing statement. To prove continuity, we would like to show that if u is close to u' and v is close to v', then $\langle u, v\rangle$ is close to $\langle u', v'\rangle$. Consequently, we would like to estimate.

$$|\langle u, v\rangle - \langle u', v'\rangle|.$$

Somehow, we would like to manipulate this so that we have

$$u - u' \text{ and } v - v'.$$

We do this by the typical mathematician's trick of adding zero in a disguised form:

$$\langle u - u', v\rangle = \langle u, v\rangle - \langle u', v\rangle.$$

That shows that

$$\langle u - u', v\rangle + \langle u', v\rangle = \langle u, v\rangle.$$

So, we see that

$$\langle u, v\rangle - \langle u', v'\rangle = \langle u - u', v\rangle + \langle u', v\rangle - \langle u', v'\rangle$$

We can smash the last two terms together because $-1 \in \mathbb{R}$ so

$$-\langle u', v'\rangle = \langle u', -v'\rangle \implies \langle u', v\rangle - \langle u', v'\rangle = \langle u', v - v'\rangle.$$

Hence,

$$|\langle u, v\rangle - \langle u', v'\rangle| = |\langle u - u', v\rangle + \langle u', v - v'\rangle|.$$

By the triangle inequality

$$|\langle u - u', v\rangle + \langle u', v - v'\rangle| \leq |\langle u - u', v\rangle| + |\langle u', v - v'\rangle|.$$

By the Cauchy-Schwarz inequality

$$|\langle u - u', v\rangle| + |\langle u', v - v'\rangle| \leq ||u - u'||||v|| + ||u'||||v - v'||.$$

We therefore see that for any fixed pair $(u, v) \in H \times H$, given $\epsilon > 0$, we can define

$$\delta := \min\left\{\frac{\varepsilon}{2(||v|| + 1)}, \frac{\varepsilon}{2(||u|| + 1)}, 1\right\}.$$

Then we estimate

$$||u - u'|| < \delta \implies ||u'|| < ||u|| + \delta \leq ||u|| + 1,$$

$$||u - u'||||v|| \leq \frac{\varepsilon||v||}{2(||v|| + 1)} < \frac{\varepsilon}{2}.$$

and

$$||u'||||v - v'|| \leq \frac{(||u|| + 1)\varepsilon}{2(||u|| + 1)} \leq \frac{\varepsilon}{2},$$

so we obtain

$$|\langle u, v\rangle - \langle u', v'\rangle| < \varepsilon.$$

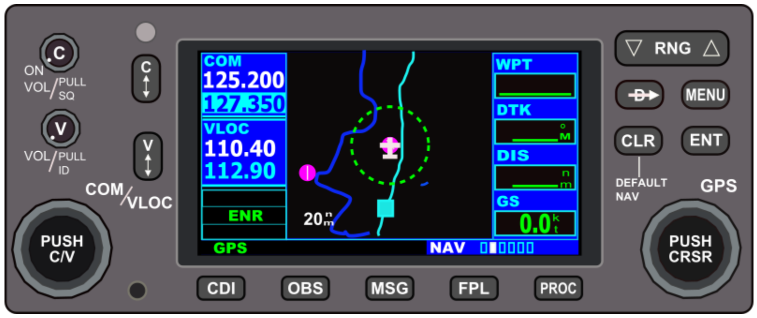


Figure 3.6: Don't get lost in a Hilbert space, find your way with an an orthogonal base! The more dimensions, the more vectors we need in our orthogonal base to navigate. Three dimensional space is an example of a Hilbert space, and one navigates an airplane through this Hilbert space using an orthogonal base - and a GPS system! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

Remark 1. This fact is useful because it allows us to bring limits inside the scalar product. You will see that we do this many times! In particular, if one has two sequences,

$$\{u_n\}_{n\geq 1}, \quad \{v_n\}_{n\geq 1} \text{ in a Hilbert space, } H,$$

and

$$\lim_{n\to\infty} u_n = u \in H, \quad \lim_{n\to\infty} v_n = v \in H,$$

then the continuity of the scalar product implies that

$$\lim_{n\to\infty} \langle u_n, v_n \rangle = \langle u, v \rangle.$$

This fact allows us to prove an infinite dimensional Pythagorean theorem!

We will often be navigating our way around infinite dimensional Hilbert spaces. To navigate around the earth, a simple compass suffices, but to navigate an airplane, one needs more sophisticated navigational aides like in Figure 3.6. Similarly, to navigate our way around an infinite dimensional Hilbert space, it will be very helpful to know that the Pythagorean Theorem is true for infinite-dimensional triangles!

Theorem 22 (Infinite dimensional Pythagorean Theorem)**.** *Assume that* $\{u_k\}_{k\geq 1}$ *are in a Hilbert space, and that*

$$\sum_{k\geq 1} u_k$$

converges to an element u *in that Hilbert space. Further, assume that the* u_k *are pairwise orthogonal. Then we have*

$$||u||^2 = \sum_{k\geq 1} ||u_k||^2.$$

Proof: The meaning of

$$\sum_{k\geq 1} u_k = u$$

is that

$$\lim_{n\to\infty} \sum_{k=1}^{n} u_k = u.$$

This is equivalent to

$$\lim_{n\to\infty} ||\sum_{k=1}^{n} u_k - u|| = 0.$$

The definition of scalar product says that

$$||u||^2 = \langle u, u\rangle.$$

Let us denote

$$U_n := \sum_{k=1}^{n} u_k.$$

Since it is a finite sum of elements of the Hilbert space, this is an element of the Hilbert space, because Hilbert spaces are vector spaces. The continuity of the scalar product shows that

$$\lim_{n\to\infty} \langle U_n, U_n\rangle = \langle U, U\rangle.$$

For each n, we also have

$$\langle U_n, U_n\rangle = \sum_{k=1}^{n} ||u_k||^2,$$

by the usual (finite) Pythagorean Theorem. Hence, we have

$$\lim_{n\to\infty} \sum_{k=1}^{n} ||u_k||^2 = ||U||^2.$$

This shows that the sum on the left converges and is equal to $||U||^2$.

3.2 Let's build an infinite dimensional Hilbert space!

To get a sense of infinite dimensional Hilbert spaces, let's build one. I am not very creative, so I am simply going to take $\mathbb{C}^n$ and let $n \to \infty$. In $\mathbb{C}^n$, the elements of this Hilbert space are vectors, that are represented by a list of n complex numbers. So, if I just let $n \to \infty$, then my elements of this $\mathbb{C}^\infty$ Hilbert space will be a list of infinitely many complex numbers. They are in an order $(c_1, c_2, c_3, c_4, \ldots)$. What else do you call an ordered list of infinitely many complex numbers? That's right, it's a sequence! So, the elements in my Hilbert space are sequences of complex numbers, which we can write more succinctly as $(c_n)_{n\geq 1}$. I am still not very creative, so I will define the scalar product of two sequences in the same way that we define it in $\mathbb{C}^n$, that is

$$\langle (a_n)_{n\geq 1}, (b_n)_{n\geq 1}\rangle := \sum_{n\geq 1} a_n \overline{b_n}.$$

The norm according to this scalar product will then be

$$||(a_n)_{n\geq 1}|| := \sqrt{\langle (a_n)_{n\geq 1}, (a_n)_{n\geq 1}\rangle} = \sqrt{\sum_{n\geq 1} |a_n|^2}.$$

The geometric meaning of the norm of an element of a Hilbert space is that is its length. All elements in a Hilbert space must have finite length. Consequently, we require that all elements of this Hilbert space satisfy:

$$\sqrt{\sum_{n\geq 1} |a_n|^2} < \infty \iff \sum_{n\geq 1} |a_n|^2 < \infty.$$

Then we can check that all of the conditions the scalar product and norm are required to satisfy are met, and consequently, the scalar product satisfies the Cauchy-Schwarz inequality

$$|\langle (a_n)_{n\geq 1}, (b_n)_{n\geq 1}\rangle| \leq ||(a_n)_{n\geq 1}|| ||(b_n)_{n\geq 1}|| < \infty.$$

It turns out that this rather intuitive Hilbert space has a name: it is called 'little ell two' and written ℓ_2. Intuitively, ℓ_2 consists of all infinite dimensional vectors that have a finite length. An orthonormal basis for ℓ_2 looks just like an orthonormal basis for $\mathbb{C}^n$ if we let $n \to \infty$. In particular, the n^{th} element in the standard orthonormal basis for ℓ_2 is the sequence such that the n^{th} component is equal to one, and all others are equal to zero. In other words, it is the sequence in which all elements, save the n^{th} one, are equal to zero. If we denote these basis vectors by $\mathbf{e}_n$, then just as with the standard unit vectors in $\mathbb{C}^n$, we can write any element of ℓ_2 in terms of these as:

$$(a_n)_{n\geq 1} = \sum_{n\geq 1} a_n \mathbf{e}_n.$$

By the definition of the scalar product, these vectors $\mathbf{e}_n$ are orthogonal and normalized, because

$$\langle \mathbf{e}_n, \mathbf{e}_m \rangle = \begin{cases} 1 & n = m \\ 0 & n \neq m. \end{cases}$$

Exercise 23. *Find an example of infinitely many vectors in ℓ_2 that are orthogonal to each other and normalized, but are* not *an orthonormal basis, in the sense that not all vectors in ℓ_2 can be expressed as a linear combination of your example.*

In order to complete the last step of solving the wave and heat equations, we will represent functions as vectors contained in infinite dimensional Hilbert spaces. For example, with the initial value problem for the heat equation on a circular rod, we wished to express the initial data $f(x)$ as a series:

$$u(x,0) = f(x) \implies \sum_{n\in\mathbb{Z}} (a_n \cos(nx) + b_n \sin(nx)) = f(x).$$

The function $f(x)$ is 2π periodic, and so it is enough to prove that we can express f as such a series for $x \in [-\pi, \pi]$. If we can ascertain that the functions $\sin(nx)$ and $\cos(nx)$ create an orthogonal basis for all functions defined on $[-\pi, \pi]$, then we will be able to use Hilbert space theory to determine the required coefficients. Essentially what we will do is to project f onto these basis functions and express f in terms of this orthogonal basis. So, we would like to understand if and when collections of functions, like these sines and cosine functions, are orthogonal vectors in a Hilbert space, and if so, do they span the space, that is can we make an orthonormal basis out of them somehow?

3.3 A Hilbert space of functions

Motivated by the preceding discussion, we would like to consider functions defined on an interval, for now we take $[-\pi, \pi]$, and interpret them as elements of a Hilbert space. Consequently, we need a way of taking in two functions, doing *something* to them, and obtaining a number. For example, if we have f and g, just multiplying them together will not work, because that will result in another function. It will give us a number *at each point,* not just one number for the whole interval. Now, we could try something like

$$\sum f(x)\overline{g(x)},$$

but then we would have to sum over all points between $-\pi$ and π, and there are uncountably many such points. So this is not going to work either. However, a process that is quite similar to a sum, that we can do over this whole interval is... integrate! Consequently, we will define the scalar product as

$$\langle f, g \rangle := \int_{-\pi}^{\pi} f(x)\overline{g(x)} dx.$$

Doing this, the length of f will be equal to

$$\sqrt{\langle f, f\rangle} = \sqrt{\int_{-\pi}^{\pi} |f(x)|^2 dx}.$$

We demand that this is finite, because all elements in a Hilbert space need to have finite length:

$$\sqrt{\langle f, f\rangle} < \infty \iff \int_{-\pi}^{\pi} |f(x)|^2 dx < \infty.$$

Consequently, our next example of a Hilbert space is the set of all functions defined on $[-\pi, \pi]$ that have finite length according to this definition. It turns out that this Hilbert space has a name: $\mathcal{L}^2(-\pi, \pi)$. The part in parentheses indicates the interval on which our functions are defined. We can analogously study $\mathcal{L}^2(a, b)$ for an interval (a, b) with real numbers $a < b$. This Hilbert space consists of all functions f that satisfy

$$\int_a^b |f(x)|^2 dx < \infty,$$

and the scalar product is

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

A word of caution: these definitions of $\mathcal{L}^2$ are what we shall call *working definitions.* They are sufficient to solve problems and for many purposes, but a truly rigorous mathematical definition invokes some measure theory.

Definition 24 (Rigorous $\mathcal{L}^2$)**.** For a bounded real interval (a, b), the Hilbert space $\mathcal{L}^2(a, b)$ consists of the equivalence classes of Lebesgue measurable functions, f such that

$$\int_a^b |f(x)|^2 dx < \infty,$$

where f and g are in the same equivalence class if and only if

$$\int_a^b |f(x) - g(x)|^2 dx = 0.$$

Please do not be disconcerted by this definition. Every function you have ever encountered in your mathematical education up to this point - unless you have taken a full course on measure theory - is measurable. So one can happily ignore the requirement of 'Lebesgue measurable.' One can quite simply assume that every function is measurable, so the key question is whether the result of squaring the function and integrating from a to b is finite? If so, your function is in (an equivalence class) of $\mathcal{L}^2(a, b)$. If not, then it's not. What about this 'equivalence class' terminology? To be considered equivalent, if we subtract the two functions, take the absolute value of the difference, square it and integrate, the result should be zero. Now let's just think about that. $|f(x) - g(x)|^2 \geq 0$ is always true. Consequently, if the result of integrating this is zero, then it pretty much means that $|f(x) - g(x)|^2$ is equal to zero on pretty much the entire interval. There is a mathematically precise expression for this: it is zero *almost everywhere.* Note further that $|f(x) - g(x)|^2 = 0$ if and only if $f(x) = g(x)$. Consequently, in the Hilbert space $\mathcal{L}^2(a, b)$ we consider two functions to be equivalent if they are equal *almost everywhere.* Mathematically, this means that the set at which f and g are not equal has *no length,* or in precise terms, Lebesgue measure zero. A point has no length. Two points have no length. But any open non-empty interval *does* have a length. For working purposes, it suffices to just think about $\mathcal{L}^2(a,b)$ as functions that are 'square integrable,' meaning that the integral of the function squared over the interval on which we are working is finite. The scalar product of two such functions is

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

Since we will need to know that the $\int_a^b |f(x)|^2 dx < \infty$ to conclude that f is in $\mathcal{L}^2(a,b)$, the following estimate will be useful.

Proposition 25 (The Standard Estimate)**.** *Assume f is defined on some interval $[a,b]$. Assume that f satisfies a bound of the form $|f(x)| \le M$ for $x \in [a,b]$.*[1] *Then,*

$$\left| \int_a^b f(x)dx \right| \le (b-a)M.$$

Proof: Standard estimate!

$$\left| \int_a^b f(x)dx \right| \le \int_a^b |f(x)|dx \le \int_a^b M dx = M(b-a).$$

Exercise 26. *An example of a function that is not in $\mathcal{L}^2(-\pi,\pi)$ is the function*

$$f(x) := \begin{cases} \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

The reason is because if we square $f(x)$ and try to integrate it, the integral tends to infinity. Use this function for inspiration to find other functions that are not in $\mathcal{L}^2(-\pi,\pi)$. Then, prove that any function that is bounded on $[-\pi,\pi]$ is in $\mathcal{L}^2(-\pi,\pi)$. If a function is continuous on $[-\pi,\pi]$ can you conclude that it is in $\mathcal{L}^2(-\pi,\pi)$? Prove or give a counterexample.

We would like to understand orthonormal bases for these Hilbert spaces. Since it works in the same way for all Hilbert spaces, we develop the theory for Hilbert spaces in general, and then we will apply the theory to study the specific Hilbert spaces that are useful for solving problems.

3.4 Bessel's inequality and orthonormal sets

Bessel's inequality is very useful. Heuristically, what it says is that first of all, if you have an orthonormal set in a Hilbert space, and you project a vector onto the span of this set, then the projected vector's length is not any longer than the original vector. Geometrically, this should jive with your intuition. Let's do a finite dimensional example. Consider the vector $\mathbf{v} = (1,2,3)$ in the Hilbert space $\mathbb{C}^3$. It has length equal to $\sqrt{1+4+9} = \sqrt{14}$. Let's consider the projection onto the span of $\mathbf{e}_1$ and $\mathbf{e}_2$. The projection is defined to be

$$\langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 = (1,2,0).$$

[1] We actually only need this for "almost every" x, where this 'almost every' has a precise measure theoretic meaning. Please be patient - when I find time - I will write an appendix containing all this measure theory. For now, please just accept this, thanks!

The length of this vector is $\sqrt{5}$. It is shorter than the length of $\mathbf{v}$ because the two vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are not an orthogonal basis for the Hilbert space. If we instead project $\mathbf{v}$ onto the span of $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$ the result is

$$\langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 + \langle \mathbf{v}, \mathbf{e}_3 \rangle \mathbf{e}_3 = \mathbf{v}.$$

Now let's do an infinite dimensional example with the Hilbert space ℓ_2. Consider the element

$$\mathbf{x} = \left(\frac{1}{n}\right)_{n \geq 1}.$$

We first check to make sure this is indeed in ℓ_2:

$$||\mathbf{x}||^2 = \sum_{n \geq 1} \frac{1}{n^2} < \infty.$$

Interestingly, using techniques from this course and indeed this chapter you will learn how to compute this series! We have a similar set of basis vectors for ℓ_2, just infinitely many of them, $\{\mathbf{e}_n\}_{n \geq 1}$. Now, let's see what happens if we project $\mathbf{x}$ onto all of these *except* the first one. The result is

$$\sum_{n \geq 2} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n = (0, 1/2, 1/3, 1/4, \ldots).$$

The length of this new vector is

$$\sqrt{\sum_{n \geq 2} \frac{1}{n^2}} < \sqrt{\sum_{n \geq 1} \frac{1}{n^2}}.$$

So, again we see that if we project onto the span of all basis vectors except one, the resulting vector is shorter than the original vector. Bessel's inequality summarizes this observation: the length of the projection of a vector onto the span of an orthonormal set is less than or equal to the length of the original vector.

Theorem 27 (Bessel's Inequality for Hilbert spaces)**.** *Let $\{\phi_n\}_{n \in \mathbb{N}}$ be an orthonormal set in a Hilbert space H. Then if $f \in H$,*

$$g := \sum_{n \in \mathbb{N}} \langle f, \phi_n \rangle \phi_n \in H,$$

and we have the inequality

$$||g||^2 = \sum_{n \in \mathbb{N}} |\langle f, \phi_n \rangle|^2 \leq ||f||^2.$$

Proof: It is tempting, and indeed I nearly fell for it, to use the infinite dimensional Pythagorean theorem. We are not allowed to do that, however, because although each term $\langle f, \phi_n \rangle \phi_n \in H$, we do not know from the start that this entire infinite sum converges to an element of H. We must prove that. So, to prove that, we will *first* prove the inequality

$$\sum_{n \in \mathbb{N}} |\langle f, \phi_n \rangle|^2 \leq ||f||^2.$$

Now, for each finite sum, it's just a finite sum so it's an element of the Hilbert space, and the Pythagorean theorem says that

$$F_N := \sum_{n=1}^N \langle f, \phi_n \rangle \phi_n \in H \implies \| \sum_{n=1}^N \hat{f}_n \phi_n \|^2 = \sum_{n=1}^N |\hat{f}_n|^2.$$

Above, we have used the convenient notation

$$\hat{f}_n := \langle f, \phi_n \rangle.$$

We will say more about this later. To proceed we look at the distance between F_N and f. This distance squared is

$$||F_N - f||^2 = \langle F_N - f, F_N - f \rangle = ||F_N||^2 - \langle f, F_N \rangle - \langle F_N, f \rangle + ||f||^2.$$

We just need to figure out those middle terms. The first one is

$$-\langle f, F_N\rangle = -\langle f, \sum_{n=1}^N \hat{f}_n \phi_n\rangle = -\sum_{n=1}^N \overline{\hat{f}_n}\langle f, \phi_n\rangle = -\sum_{n=1}^N |\hat{f}_n|^2.$$

We have used the properties of the scalar product together with the definition of $\hat{f}_n$. If you don't see how this works, please review the properties of the scalar product until you understand these manipulations. Next we compute

$$-\langle F_N, f\rangle = -\langle \sum_{n=1}^N \hat{f}_n \phi_n, f\rangle = -\sum_{n=1}^N \hat{f}_n \langle \phi_n, f\rangle = -\sum_{n=1}^N \hat{f}_n \overline{\hat{f}_n} = -\sum_{n=1}^N |\hat{f}_n|^2.$$

We have again used the properties of the scalar product together with the definition of $\hat{f}_n$. If you don't see how this works, please review the properties of the scalar product until you understand these manipulations. Consequently, we have

$$||F_N - f||^2 = \sum_{n=1}^N |\hat{f}_n|^2 - 2\sum_{n=1}^N |\hat{f}_n|^2 + ||f||^2 = -\sum_{n=1}^N |\hat{f}_n|^2 + ||f||^2.$$

The two terms have opposite signs. Let us think about the meaning of the left side. Or just look at it naively and say 'it is stuff squared.' Yes, more precisely, it is real stuff squared. Real stuff squared is always greater than or equal to zero. If we think geometrically, it is the distance (squared) between two vectors, and distances are always greater than or equal to zero. So we have

$$0 \le ||F_N - f||^2 = -\sum_{n=1}^N |\hat{f}_n|^2 + ||f||^2 \implies \sum_{n=1}^N |\hat{f}_n|^2 \le ||f||^2.$$

The right side is fixed and finite. So, we can let $N \to \infty$ on the left side, and that right side just stays put, so we arrive at

$$\sum_{n\ge 1} |\hat{f}_n|^2 \le ||f||^2.$$

We can now use this to prove that the sequence $\{F_N\}_{N\ge 1}$ is a Cauchy sequence. Then the fact that Hilbert spaces are complete implies that the sequence has a limit that is an element of the Hilbert space, so we will have that

$$\lim_{N\to\infty} F_N = \sum_{n\ge 1} \hat{f}_n \phi_n = g \in H.$$

To prove that the sequence is Cauchy, we start with $\varepsilon > 0$, an arbitrary small number that is handed to us. Since we have proven that

$$\sum_1^\infty |\hat{f}_n|^2 < \infty,$$

there exists $N \in \mathbb{N}$ such that

$$\sum_N^\infty |\hat{f}_n|^2 < \varepsilon^2.$$

This is because the tail of any convergent series can be made as small as we like. So, now if we have $N_1 \ge N_2 \ge N$, we estimate

$$||F_{N_1} - F_{N_2}||^2 = ||\sum_{N_2+1}^{N_1} \hat{f}_n \phi_n||^2 = \sum_{N_2+1}^{N_1} |\hat{f}_n|^2$$

$$\leq \sum_{N_2+1}^{\infty} |\hat{f}_n|^2 \leq \sum_{N}^{\infty} |\hat{f}_n|^2 < \varepsilon^2.$$

Consequently we have that for all $N_1 \geq N_2 \geq N$,

$$||F_{N_1} - F_{N_2}|| < \varepsilon.$$

This is the definition of being a Cauchy sequence. Every Cauchy sequence in a complete space has a limit in that space (this is the definition of being complete!). Consequently, we obtain that

$$\lim_{N \to \infty} \sum_{n=1}^{N} \hat{f}_n \phi_n = g \in H.$$

It is only *now* that we can invoke the infinite dimensional Pythagorean Theorem! Knowing that this sum is an element of the Hilbert space and that ϕ_n are orthonormal, we also have that $\hat{f}_n \phi_n \in H$ are orhthogonal to each other. We therefore have

$$||g||^2 = \sum_{n \geq 1} ||\hat{f}_n \phi_n||^2 = \sum_{n \geq 1} |\hat{f}_n|^2 ||\phi_n||^2 = \sum_{n \geq 1} |\hat{f}_n|^2 \leq ||f||^2.$$

3.5 The 3 equivalent conditions to determine if an orthonormal set in a Hilbert space is in fact an orthonormal basis

Perhaps what makes the following theorem so nice is the pleasant setting of a Hilbert space, or translated directly from German, a Hilbert room. Hilbert rooms are cozy. The reason is because there is a notion of orthogonality, so it is very easy to find one's way around, much like the grid-like streets in the USA. We use the elements of an orthogonal base to navigate our way through a Hilbert space. We can think of them as road signs that guide our way, with our slogan:

> Don't get lost in a Hilbert space, find your way with an an orthogonal base!

An orthogonal basis or base is a collection of elements (vectors) in a Hilbert space that are all orthogonal, like the streets in the USA as in Figure 3.7, such that every vector in that Hilbert space can be expressed as a linear combination of the basis vectors. This is useful for many purposes including solving partial differential equations, computing seemingly impossible sums, and determining best approximations. For finite dimensional Hilbert spaces, like $\mathbb{C}^n$, if we have n vectors that are all orthogonal then we immediately know that they are an orthogonal basis. It is not so simple for infinite dimensional Hilbert spaces, but we will manage nonetheless![2]

For infinite dimensional Hilbert spaces, we cannot just say 'infinitely many' orthogonal vectors is enough to be a basis. Recall the example of ℓ^2. The vectors $\{\mathbf{e}_n\}_{n \geq 2}$ are not a basis, because there is no way to write the vector $\mathbf{x} = (1/n)_{n \geq 1}$ in terms of only these. We need $\mathbf{e}_1$. Moreover, we saw that when we project $\mathbf{x}$ onto $\{\mathbf{e}_n\}_{n \geq 2}$, the projected vector is *shorter* than $\mathbf{x}$. It turns out that this is one way to ascertain whether or not an orthonormal set is a basis. If Bessel's inequality is in fact an *equality,* that is if the length of the projected vector is equal to that of the original vector, then the set is a basis. There are in total three equivalent 'checks' one can perform to determine whether or not an orthonormal set in a Hilbert space is a bases.

[2] Anecdote alert! I traveled alone to France when I was 15. I spoke pretty good French. Landed in Paris, figured out the Metro, made it out to the street. Where the F were the street signs??? It took me a while to figure out that in France, they put street signs *on buildings.* It was a bit of a funny logic puzzle, standing there thinking, the French people must have the names of the street somewhere, because they have to navigate. Where the heck are they? Looking looking until aha! On the buildings! Rather space-saving and clever.

Theorem 28 (The 3 equivalent conditions for an ONS to be an ONB)**.** *Let $\{\phi_n\}_{n\in\mathbb{N}}$ be orthonormal in a Hilbert space, H. The following are equivalent:*

$$(1) \quad f \in H \text{ and } \langle f, \phi_n\rangle = 0 \forall n \in \mathbb{N} \implies f = 0.$$

$$(2) \quad f \in H \implies f = \sum_{n\in\mathbb{N}} \langle f, \phi_n\rangle \phi_n.$$

$$(3) \quad ||f||^2 = \sum_{n\in\mathbb{N}} |\langle f, \phi_n\rangle|^2 .$$

The last of these is known as Parseval's equation. *If any of these three equivalent conditions hold, then we say that $\{\phi_n\}$ is an orthonormal basis of H.*

Proof: We shall proceed in order prove (1) $\implies$ (2), then (2) $\implies$ (3), and finally (3) $\implies$ (1). Stay calm and carry on.

First we assume statement (1) holds, and then we shall show that (2) must hold as well. Bessel's Inequality Theorem says that

$$g := \sum_{n\in\mathbb{N}} \langle f, \phi_n\rangle \phi_n \in H.$$

So, we would like to prove that in fact $g = f$, somehow using the fact that statement (1) holds true.

Idea!

let's try to show that $f - g = 0$. This will imply that $f = g$. To use (1) we should compute then

$$\langle f - g, \phi_n\rangle.$$

Let's do this.

$$\langle f - g, \phi_n\rangle = \langle f, \phi_n\rangle - \langle g, \phi_n\rangle.$$

We insert the definition of g as the series,

$$\langle g, \phi_n\rangle = \langle \sum_{m\geq 1} \langle f, \phi_m\rangle \phi_m, \phi_n\rangle = \sum_{m\geq 1} \langle f, \phi_m\rangle \langle \phi_m, \phi_n\rangle = \langle f, \phi_n\rangle.$$

Above, we have used in the second equality the linearity of the inner product and the continuity of the inner product. In the third equality, we have used that $\langle \phi_m, \phi_n\rangle$ is 0 if $m \neq n$, and is 1 if $m = n$. Hence, only the term with $m = n$ survives in the sum. Thus,

$$\langle f - g, \phi_n\rangle = \langle f, \phi_n\rangle - \langle g, \phi_n\rangle = \langle f, \phi_n\rangle - \langle f, \phi_n\rangle = 0, \quad \forall n \in \mathbb{N}.$$

By (1), this shows that $f - g = 0 \implies f = g$.

Next, we shall assume that (2) holds, and we shall use this to demonstrate (3). By (2),

$$f = \sum_{n\in\mathbb{N}} \hat{f}_n \phi_n, \quad \hat{f}_n := \langle f, \phi_n\rangle.$$

To obtain (3), we can simply apply our infinite dimensional Pythagorean theorem, which says that

$$||f||^2 = \sum_{n\in\mathbb{N}} ||\hat{f}_n \phi_n||^2 = \sum_{n\in\mathbb{N}} |\hat{f}_n|^2 ||\phi_n||^2 = \sum_{n\in\mathbb{N}} |\hat{f}_n|^2.$$

Finally, we assume (3) holds and use it to show that (1) must also hold. This is pleasantly straightforward. We assume that for some f in our Hilbert space, $\langle f, \phi_n\rangle = 0$ for all n. Using (3), we compute

$$||f||^2 = \sum_{n\in\mathbb{N}} |\langle f, \phi_n\rangle|^2 = \sum_{n\in N} 0 = 0.$$

The only element in a Hilbert space with norm equal to zero is the 0 element. Thus $f = 0$.

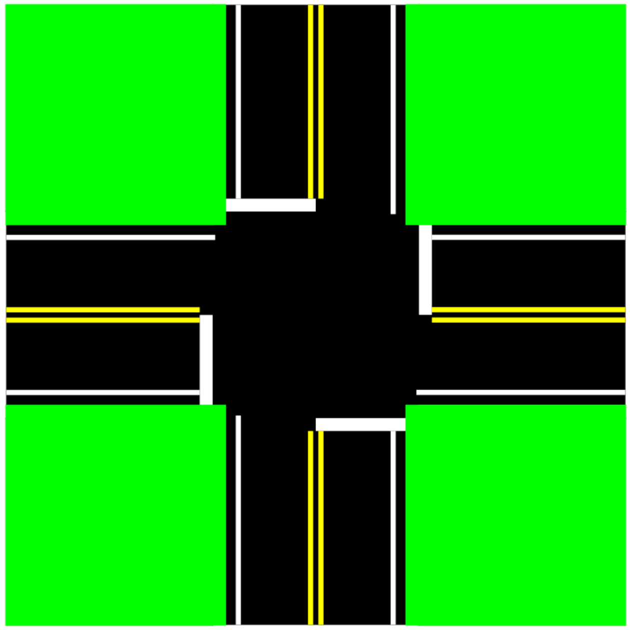


Figure 3.7: Don't get lost in a Hilbert space, find your way with an orthogonal base! The elements of an orthonormal (or just orthogonal but not normalized) basis are like street signs that help us navigate the Hilbert space. They help us to see the Hilbert space like an infinite dimensional version of the grid-like structure of streets in the US as shown here. Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`.

3.6 Fourier series on Hilbert spaces are the best approximations!

We have been discussing the projection of vectors in a Hilbert space onto the span of sets of orthonormal vectors. This in fact has a name.

Definition 29. The *Fourier series* of an element f of a Hilbert space, H, with respect to an orthonormal set $\{\phi_n\}$ is defined to be

$$\sum_n \langle f, \phi_n \rangle \phi_n.$$

The coefficients are known as *Fourier coefficients,* and are often denoted

$$\hat{f}_n = \langle f, \phi_n \rangle.$$

If the set is orthogonal, but not normalized, then the Fourier coefficients are defined to be

$$\hat{f}_n = \frac{\langle f, \phi_n \rangle}{||\phi_n||^2},$$

and the Fourier series is also defined as above,

$$\sum_n \hat{f}_n \phi_n.$$

Exercise 30. *Use the theorem on the 3 equivalent conditions for an ONS to be an ONB to show that the Fourier series of f is actually* equal *to f if and only if the orthonormal set satisfies one of the three conditions of the theorem.*

3.6.1 The Best Approximation!

Although the Fourier series of f might not be equal to f, it is the *best* approximation to f in the following sense.

Theorem 31 (Best Approximation)**.** *Let $\{\phi_n\}_{n\in\mathbb{N}}$ be an orthonormal set in a Hilbert space, H. If $f \in H$, and*

$$\sum_{n\in\mathbb{N}} c_n\phi_n \in H,$$

then

$$||f - \sum_{n\in\mathbb{N}} \langle f, \phi_n\rangle \phi_n|| \le ||f - \sum_{n\in\mathbb{N}} c_n\phi_n||,$$

and equality holds $\iff$ $c_n = \langle f, \phi_n\rangle$ is true $\forall n \in \mathbb{N}$.

Proof: We make a few definitions: let

$$g := \sum \widehat{f_n}\phi_n, \quad \widehat{f_n} = \langle f, \phi_n\rangle,$$

and

$$\varphi := \sum c_n\phi_n.$$

Idea!

$$||f-\varphi||^2 = ||f-g+g-\varphi||^2 = ||f-g||^2 + ||g-\varphi||^2 + 2\,\mathrm{Re}\langle f-g, g-\varphi\rangle.$$

Idea!

$$\langle f-g, g-\varphi\rangle = 0.$$

Just write it out (stay calm and carry on):

$$\langle f, g\rangle - \langle f, \varphi\rangle - \langle g, g\rangle + \langle g, \varphi\rangle$$

$$= \sum \overline{\widehat{f_n}}\langle f, \phi_n\rangle - \sum \overline{c_n}\langle f, \phi_n\rangle - \sum \widehat{f_n}\langle \phi_n, \sum \widehat{f_m}\phi_m\rangle + \sum \widehat{f_n}\langle \phi_n, \sum c_m\phi_m\rangle$$
$$= \sum |\widehat{f_n}|^2 - \sum \overline{c_n}\widehat{f_n} - \sum |\widehat{f_n}|^2 + \sum \widehat{f_n}\overline{c_n} = 0,$$

where above we have used the fact that ϕ_n are an orthonormal set. Then, we have

$$||f-\varphi||^2 = ||f-g||^2 + ||g-\varphi||^2 \ge ||f-g||^2,$$

with equality iff

$$||g-\varphi||^2 = 0.$$

Let us now write out what this norm is, using the definitions of g and φ. By their definitions,

$$g - \varphi = \sum(\widehat{f_n} - c_n)\phi_n.$$

By the Pythagorean theorem, due to the fact that the ϕ_n are an orthonormal set, and hence multiplying them by the scalars, $\widehat{f_n} - c_n$, they remain orthogonal, we have

$$||g-\varphi||^2 = \sum \left|\widehat{f_n} - c_n\right|^2.$$

This is a sum of non-negative terms. Hence, the sum is only zero if all of the terms in the sum are zero. The terms in the sum are all zero iff

$$\left|\widehat{f_n} - c_n\right| = 0 \forall n \iff c_n = \widehat{f_n} \forall n \in \mathbb{N}.$$

Corollary 32. *Assume that $\{\phi_n\}$ is an OS in a Hilbert space H. Then the best approximation to $f \in H$ of the form*

$$\sum_{n=1}^{N} c_n \phi_n$$

is given by taking

$$c_n = \frac{\langle f, \phi_n \rangle}{||\phi_n||^2}.$$

Exercise 33. *Prove this corollary using the best approximation theorem.*

3.7 Orthogonal bases for the Hilbert space $\mathcal{L}^2(-\pi, \pi)$

We will compute that the functions e^{inx} are orthogonal, and if we divide by their norm, then we obtain an orthonormal set.

Proposition 34. *On the interval $[-\pi, \pi]$, the functions*

$$\phi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}$$

are an orthonormal set with respect to the scalar product,

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)} dx.$$

Proof: By definition, we consider

$$\int_{-\pi}^{\pi} \frac{e^{inx}}{\sqrt{2\pi}} \frac{\overline{e^{imx}}}{\sqrt{2\pi}} dx.$$

We bring the constant factor out in front of the integral the constant factor, and we recall that $\overline{e^{imx}} = e^{-imx}$, so we are computing

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-imx} dx.$$

Exercise 35. *Why is*

$$\overline{e^{imx}} = e^{-imx}?$$

Explain in your own words or prove it algebraically.

So, we compute,

$$\int_{-\pi}^{\pi} e^{ix(n-m)} dx = \begin{cases} 2\pi & m = n \\ \left.\frac{e^{ix(n-m)}}{n-m}\right|_{x=-\pi}^{\pi} & n \neq m \end{cases}.$$

Now, we know that

$$e^{i\pi(n-m)} = \begin{cases} 1 & n-m \text{ is even} \\ -1 & n-m \text{ is odd.} \end{cases}.$$

To see this, I just imagine where we are on the Liseberghjul... Or you can write this out as

$$e^{i\pi(n-m)} = \cos(\pi(n-m)) + i\sin(\pi(n-m)).$$

The sine term is always zero since n and m are integers, and the cosine is either 1 or -1. Similarly,

$$e^{-i\pi(n-m)} = \begin{cases} 1 & n-m \text{ is even} \\ -1 & n-m \text{ is odd.} \end{cases}.$$

So in all cases, when $n \neq m$,

$$e^{i\pi(n-m)} - e^{-i\pi(n-m)} = 0.$$

Hence,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-imx}dx = \begin{cases} \frac{2\pi}{2\pi} = 1 & n = m \\ 0 & n \neq m \end{cases}.$$

This is precisely what it means to be orthonormal!

Exercise 36. *Show that the set of functions*

$$\left\{ \frac{1}{\sqrt{2\pi}}, \quad \frac{\sin(nx)}{\sqrt{\pi}}, \quad \frac{\cos(nx)}{\sqrt{\pi}} \right\}_{n\geq 1}$$

is also an orthonormal set in the Hilbert space $\mathcal{L}^2(-\pi,\pi)$.

We will frequently use, but not prove the following theorem.

Theorem 37 (Orthogonal bases of complex exponentials and trigonometric functions)**.** *The set of functions*

$$\left\{ \frac{e^{inx}}{\sqrt{2\pi}} \right\}_{n\in\mathbb{Z}}$$

is an orthonormal basis for the Hilbert space $\mathcal{L}^2(-\pi,\pi)$. *The set of functions*

$$\left\{ \frac{1}{\sqrt{2\pi}}, \quad \frac{\sin(nx)}{\sqrt{\pi}}, \quad \frac{\cos(nx)}{\sqrt{\pi}} \right\}_{n\geq 1}$$

is an orthonormal basis for the Hilbert space $\mathcal{L}^2(-\pi,\pi)$.

We can effectively use this theorem without proof, but we explain roughly why the theorem is analogous to the spectral theorem for hermitian matrices. The functions in Theorem 37 all satisfy a certain differential equation:

$$f_n''(x) = -n^2 f_n(x), \quad f_n(x) \in \left\{e^{inx}, \sin(nx), \cos(nx)\right\}, \quad \forall n \in \mathbb{Z}.$$

For this reason, we consider the differential operator $\frac{d^2}{dx^2}$. It is linear because for any functions f and g and constants a and b,

$$\frac{d^2}{dx^2}(af(x) + bg(x)) = af''(x) + bg''(x).$$

This linear differential operator acts on the Hilbert space $\mathcal{L}^2(-\pi,\pi)$. The functions f_n above are *eigenfunctions* with *eigenvalues* $-n^2$. These functions all satisfy periodic boundary conditions:

$$f_n(-\pi) = f_n(\pi), \quad f_n'(-\pi) = f_n'(\pi).$$

As a consequence, the differential operator (with these boundary conditions) satisfies a condition that is like an analogue to the assumption that the matrix in the spectral theorem is hermitian. Consequently, there is a spectral theorem for this operator that says that for the Hilbert space, there is an orthogonal basis that consists of eigenfunctions. Since

$$\sin(nx) = \frac{e^{inx} - e^{-inx}}{2i}, \quad \cos(nx) = \frac{e^{inx} + e^{-inx}}{2}, \tag{3.7.1}$$

two orthogonal bases are

$$\left\{e^{inx}\right\}_{n\in\mathbb{Z}} \text{ and } \{1, \sin(nx), \cos(nx)\}_{n\geq 1}.$$

3.8 Trigonometric Fourier series

Since two orthogonal bases for the Hilbert space $\mathcal{L}^2(-\pi,\pi)$ are

$$\left\{e^{inx}\right\}_{n\in\mathbb{Z}} \text{ and } \{1, \sin(nx), \cos(nx)\}_{n\geq 1},$$

we can express all functions that satisfy

$$\int_{-\pi}^{\pi} |f(x)|^2 < \infty$$

as Fourier series on this interval!

Definition 38 (Trigonometric Fourier series)**.** Assume f is defined $[-\pi,\pi]$. The Fourier coefficients of f with respect to the basis functions e^{inx} are defined to be

$$c_n := \frac{\langle f, e^{inx}\rangle}{||e^{inx}||^2} = \frac{\int_{-\pi}^{\pi} f(x)\overline{e^{inx}}dx}{\int_{-\pi}^{\pi} |e^{inx}|^2 dx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

The Fourier series of f with respect to this basis is

$$\sum_{n\in\mathbb{Z}} c_n e^{inx}.$$

The Fourier coefficients of f with respect to the basis functions $\{1, \cos(nx), \sin(nx)\}_{n\geq 1}$ are defined to be

$$a_n := \frac{\langle f, \cos(nx)\rangle}{||\cos(nx)||^2} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\overline{\cos(nx)}dx, \quad n \geq 1,$$

$$b_n := \frac{\langle f, \sin(nx)\rangle}{||\sin(nx)||^2} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\overline{\sin(nx)}dx, \quad n \geq 1.$$

The Fourier series of f with respect to the basis functions $\{1, \cos(nx), \sin(nx)\}_{n\geq 1}$ is

$$c_0 + \sum_{n\geq 1} a_n \cos(nx) + b_n \sin(nx).$$

Figure 3.8: Here is a little katodanode fluffball finding its way around the Hilbert space, $\mathcal{L}^2(-\pi, \pi)$ with help of the orthogonal base $\{1, \cos(nx), \sin(nx)\}_{n\geq 1}$. Thanks to Ebba Grönfors for this cute mnemonic illustration!

3.8.1 Computing trigonometric Fourier series: an example

Let's start with the function $f(x) = |x|$. By definition, the Fourier coefficients with respect to the orthogonal basis functions e^{inx} are:

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} |x|e^{-inx}dx, \quad c_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|x|dx = \frac{2\pi^2}{2(2\pi)} = \frac{\pi}{2}.$$

Since

$$|x| = \begin{cases} -x & x < 0 \\ x & x \geq 0 \end{cases}$$

we compute:

$$\int_{-\pi}^{0} -xe^{-inx}dx, \quad \int_{0}^{\pi} xe^{-inx}dx.$$

We do substitution in the first integral to change it:

$$\int_{-\pi}^{0} -xe^{-inx}dx = \int_{0}^{\pi} xe^{inx}dx = \frac{xe^{inx}}{in}\bigg|_0^{\pi} - \int_0^{\pi} \frac{e^{inx}}{in}dx$$

$$= \frac{\pi e^{in\pi}}{in} - \frac{e^{in\pi}}{(in)^2} + \frac{1}{(in)^2}.$$

Similarly we also use integration by parts to compute

$$\int_0^{\pi} xe^{-inx}dx = \frac{xe^{-inx}}{-in}\bigg|_0^{\pi} - \int_0^{\pi}\frac{e^{-inx}}{(-in)}dx$$

$$= \frac{\pi e^{-in\pi}}{-in} - \frac{e^{-in\pi}}{(-in)^2} + \frac{1}{(-in)^2}.$$

Adding them up and using the 2π periodicity, we get

$$\frac{2e^{in\pi}}{n^2} - \frac{2}{n^2} = \frac{2(-1)^n - 2}{n^2}.$$

OBS! We need to divide by 2π to get

$$c_n = \frac{(-1)^n - 1}{\pi n^2}, \quad n \in \mathbb{Z} \setminus \{0\}.$$

The Fourier series is therefore

$$\frac{\pi}{2} + \sum_{n \in \mathbb{Z},\ \text{odd}} e^{inx}\left(-\frac{2}{\pi n^2}\right).$$

Exercise 39. *Use these calculations to compute the series*

$$c_0 + \sum_{n \geq 1} a_n \cos(nx) + b_n \sin(nx)$$

and to show that all of the b_n are equal to zero.

3.9 Trigonometric Fourier series to compute sums and π falling out of the sky

In mathematics and its applications, there are many infinite series that nobody in the world knows how to compute analytically. For example, if I just make up a sum that I know converges, like

$$\sum_{n\geq 1} \frac{1}{2n^2+3n^3+4n^4}. \tag{3.9.1}$$

I honestly have no idea how to compute that sum analytically. We could approximate it by computing up to the millionth term with a computer, and the result would be fairly close to the value of the sum. For example, we can estimate the *tail* of the series

$$\sum_{n\geq N} \frac{1}{2n^2+3n^3+4n^4} \leq \sum_{n\geq N} \frac{1}{2n^2}.$$

The series

$$\sum_{n\geq 1} \frac{1}{2n^2} = \frac{\pi^2}{12}.$$

Why on earth does π^2 appear on the right side? We will see exactly why and how this happens by using a Fourier series to compute the sum! So, in fact, Fourier series can be useful because we can use them to compute series that we can in turn use to estimate the tails of other series that we cannot compute analytically! This is awesome, because we can then obtain a rigorous, analytical estimate for the tail of the unknown series. Consequently, we can *rigorously* compute the value of the series (3.9.1) up to any accuracy we desire! It is a clever mathematical trick to do the seemingly impossible!

3.9.1 The Basel problem

To compute the sum

$$\sum_{n\geq 1} \frac{1}{n^2}$$

the first step is to search for a function whose Fourier series bears some resemblance to the sum. This is a famous series, so famous that its computation has a name and a history. It is called the Basel problem and was posed by Pietro Mengoli in 1650 and solved by Leonhard Euler in 1734. Euler generalized the problem considerably, and his ideas helped to inspire Bernhard Riemann's 1859 paper *On the Number of Primes Less Than a Given Magnitude,* [20] in which he defined the zeta function that is now known as the Riemann zeta function. The original hand-written script is found here http://www.claymath.org/sites/default/files/riemann1859.pdf. Riemann's zeta function is defined and briefly explored in exercise 17 of this chapter. One of the most famous unsolved math problems is *Riemann's hypothesis,* and concerns the locations of the zeros of the Riemann zeta function. It can be equivalently formulated to be a remainder estimated in the Prime Number Theorem. This exemplifies how seemingly different areas of mathematics are connected. Moreover, Riemann's zeta function is a particular example of a spectral zeta function, and these are quite important in physics. This will be further explored in §4.8. For now, let us return to computing the sum $\frac{1}{n^2}$.

We will do this using a Fourier series, somehow. Let's start with the simplest function ever, 1. Well, its Fourier series is just that, 1, so that won't help us. Next, perhaps, we could try the function x. We compute using integration by parts that

$$n \neq 0 \implies \int_{-\pi}^{\pi} xe^{-inx}dx = \frac{xe^{-inx}}{-in}\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} \frac{e^{-inx}}{-in}dx = \frac{2\pi(-1)^n}{-in} - \frac{1}{-in}\,\frac{e^{-inx}}{-in}\Big|_{-\pi}^{\pi} = \frac{2\pi(-1)^n}{-in}.$$

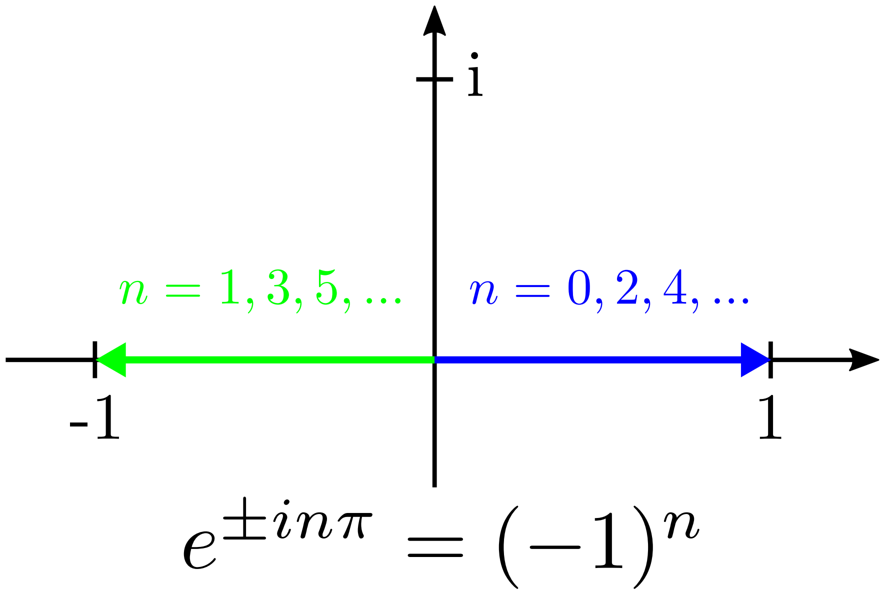


Figure 3.9: To help you remember this useful fact, here is a picture created by Gottfrid Olsson!

We have used the fact that $e^{\pm in\pi} = (-1)^n$. I remember this using the unit circle, or Lisebergs ferris wheel; see the rather barebones illustration in Figure 3.9. When n is even, the point $e^{\pm in\pi}$ is sitting at 1, whereas when n is odd, the point $e^{\pm in\pi}$ is sitting at -1. The value of the $\pm$ does not matter!

Since it is an odd function, the zeroth Fourier coefficient is zero:

$$\int_{-\pi}^{\pi} x dx = 0.$$

So, we therefore have computed the Fourier coefficients for $f(x) = x$ on the interval $(-\pi, \pi)$ are

$$c_n = \begin{cases} \frac{1}{2\pi}\frac{2\pi(-1)^n}{-in} & n \neq 0 \\ 0 & n = 0. \end{cases}.$$

The Fourier series is therefore

$$\sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{(-1)^n e^{inx}}{-in}.$$

We see that if we take the coefficients $|c_n|^2$ the result is... $\frac{1}{n^2}$. So, we can use this Fourier series to solve the famous Basel problem! Parseval's equality tells us that if we use the orthonormal basis

$$\left\{\frac{e^{inx}}{\sqrt{2\pi}}\right\}_{n \in \mathbb{Z}} \implies \hat{f}_n := \int_{-\pi}^{\pi} f(x)\overline{\frac{e^{inx}}{\sqrt{2\pi}}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{-inx} dx$$

$$\implies ||f||^2 = \int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{n \in \mathbb{Z}} |\hat{f}_n|^2.$$

Since

$$c_n = \frac{\hat{f}_n}{\sqrt{2\pi}} \implies \sqrt{2\pi} c_n = \hat{f}_n$$

Figure 3.10: Having solved Basel's problem, we have earned ourselves some pie, as we have also seen the reason that π would rather surprisingly appear in the calculation of the series $\sum n^{-2}$. Intuitively, the reason is that using a Fourier series to compute the sum of the Fourier coefficients of a polynomial function on $(-\pi, \pi)$, by Parseval's equality, the sum is related to the $\mathcal{L}^2$ norm of the polynomial. That is equal to the integral of the polynomial squared on $(-\pi, \pi)$. If we think about integrating polynomials, the result is another polynomial, which we evaluate at the endpoints. Stuffing $\pm\pi$ into a polynomial, we'll get powers of π. For this reason, even though it is not obvious, now we see why we can expect that evaluating sums like $\sum_{n\geq 1} \frac{1}{p(n)}$ for polynomials $p(n)$, we can actually expect πs to come flying out! Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`.

this tells us that

$$\int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{n\in\mathbb{Z}} |\hat{f}_n|^2 = \sum_{n\in\mathbb{Z}} |\sqrt{2\pi} c_n|^2.$$

On the left side we compute

$$\int_{-\pi}^{\pi} x^2 dx = \left.\frac{x^3}{3}\right|_{-\pi}^{\pi} = \frac{2\pi^3}{3}.$$

On the right side we compute

$$\sum_{n\in\mathbb{Z}} |\sqrt{2\pi} c_n|^2 = \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{2\pi}{n^2} = 4\pi \sum_{n\geq 1} \frac{1}{n^2}.$$

We therefore have

$$\frac{2\pi^3}{3} = 4\pi \sum_{n\geq 1} \frac{1}{n^2} \implies \frac{\pi^2}{6} = \sum_{n\geq 1} \frac{1}{n^2}.$$

This is a beautiful fact! If I were to look at the sum on the right of $\frac{1}{n^2}$ for natural numbers n, I would not really guess that the result would involve π. However, from the perspective of trigonometric Fourier series, it is quite natural, because the $\mathcal{L}^2$ norm of a polynomial on the interval $(-\pi, \pi)$ will have a lot of π in it. Speaking of π, see Figure 3.10.

3.9.2 Parseval's equality to compute another series

We have computed the Fourier series for the function $|x|$ on the interval $(-\pi, \pi)$ with respect to the orthogonal basis, and it is

$$\frac{\pi}{2} + \sum_{n\in\mathbb{Z},\ \text{odd}} e^{inx} \left(-\frac{2}{\pi n^2}\right).$$

To apply the Parseval equality, we would like to express $f(x) = |x|$ as

$$\sum_{n\in\mathbb{Z}} \hat{f}_n \phi_n(x), \quad \hat{f}_n = \langle f, \phi_n \rangle.$$

So, let's see how to obtain these $\hat{f}_n$ from the coefficients we computed:

$$\frac{\pi}{2} = \frac{\pi\sqrt{2\pi}}{2}\frac{1}{\sqrt{2\pi}} = \frac{\pi\sqrt{2\pi}}{2}\phi_0 \implies \hat{f}_0 = \frac{\pi\sqrt{2\pi}}{2}.$$

Next,

$$-\frac{2}{\pi n^2}e^{inx} = -\frac{2\sqrt{2\pi}}{\pi n^2}\frac{e^{inx}}{\sqrt{2\pi}} \implies \hat{f}_n = \begin{cases} 0 & n \text{ even} \\ -\frac{2\sqrt{2\pi}}{\pi n^2} & n \text{ odd.} \end{cases}$$

So, the Parseval equation states that

$$||f||^2 = \int_{-\pi}^{\pi} |x|^2 dx = \sum_{n\in\mathbb{Z}} |\hat{f}_n|^2 = \frac{\pi^2(2\pi)}{4} + \sum_{n\geq 1 \text{ odd}} \frac{4(2\pi)}{\pi^2 n^4} = \frac{\pi^3}{2} + \sum_{n\geq 1 \text{ odd}} \frac{8}{\pi n^4}.$$

On the other hand,

$$\int_{-\pi}^{\pi} |x|^2 dx = 2\int_0^{\pi} x^2 dx = 2\frac{\pi^3}{3}.$$

So we have obtained the equality

$$2\frac{\pi^3}{3} = \frac{\pi^3}{2} + \sum_{n\geq 1 \text{ odd}} \frac{8}{\pi n^4}.$$

We can re-arrange this equality to compute the sum

$$\sum_{n\geq 1 \text{ odd}} \frac{1}{n^4} = \frac{\pi}{8}\pi^3\left(\frac{2}{3} - \frac{1}{2}\right) = \frac{\pi^4}{48}.$$

It is again not entirely obvious from looking at the left side that one should expect the sum to involve π, but indeed, we see that π again falls out of the sky!

3.10 Computing best approximations

Can we find the four numbers $\{c_n\}_{n=0}^3$ that minimize the integral

$$\int_{-\pi}^{\pi} |f - \sum_{j=0}^{3} c_n e^{inx}|^2 dx,$$

for

$$f(x) = \begin{cases} 0 & -\pi < x < 0 \\ 1 & 0 \leq x \leq \pi \end{cases}?$$

Yes, we can, and the method is to apply the best approximation theorem! The functions e^{inx} are orthogonal on $\mathcal{L}^2(-\pi, \pi)$, so we can apply the best approximation theorem! It says that the best approximation is to set

$$c_n = \frac{\hat{f}_n}{||e^{inx}||^2} = \frac{\langle f, \phi_j n \rangle}{||\phi_n||^2}, \quad \phi_n(x) = e^{inx}.$$

We therefore compute

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx = \begin{cases} \frac{1}{2} & n = 0 \\ \frac{(-1)^n - 1}{-2\pi i n} & j = 1, 2, 3 \end{cases}$$

3.11 Hilbert spaces in mathematical physics

The concept of a Hilbert space offers one of the best formulations of quantum mechanics that was developed by John von Neumann [22]. The pure states of a quantum mechanical system are represented by unit vectors, known as state vectors, that sit inside a separable Hilbert space, known as the state space. A Hilbert space is called *separable* if and only if it has a finite or countable orthonormal basis; this is investigated further in exercise 19 of this chapter. For example, the position and momentum states for a single non-relativistic spin zero particle is an element of a certain $\mathcal{L}^2$ space. The states for the spin of a single proton are unit elements of the two-dimensional complex Hilbert space $\mathbb{C}^2$. The elements of this space are called spinors. Each observable is represented by a self-adjoint linear operator acting on the state space. If we denote such an operator by L then it satisfies

$$L : H \to H, \quad L(af + bg) = aL(f) + bL(g), \quad a, b \in \mathbb{C}, \quad f, g \in H$$

for the Hilbert space H, and

$$\langle Lf, g\rangle = \langle f, Lg\rangle.$$

The eigenstates of an observable correspond to the eigenvectors of the operator, that are the elements of H such that there is $\lambda \in \mathbb{C}$ with $L(f) = \lambda f$. The associated eigenvalue λ corresponds to the value of the observable in that eigenspace. An example is shown in Figure 3.11. Note that by the definition of the scalar product and its properties, this immediately implies that the eigenvalues are in fact in $\mathbb{R}$.

The scalar product between two state vectors is a probability amplitude. During an ideal measurement of a quantum mechanical system, the probability that a system collapses from a given initial state to a particular eigenstate is equal to the square of the absolute value of the probability amplitudes between the initial and final states. For a general system, states are typically represented as statistical mixtures of pure states. Mathematically, this means that the pure states are our orthonormal base, and general states are expressed in terms of this orthonormal base. Hilbert spaces are not only useful for quantum mechanics, but they can also be used to describe classical mechanics and dynamical systems. Learning about Hilbert spaces from a general mathematical perspective, as we take here, prepares you to be able to use them to describe a myriad of physical systems!

3.12 Exercises

1. Show that for $\mathbf{v}, \mathbf{w}$ in $\mathbb{C}^n$, then $|\langle \mathbf{v}, \mathbf{w}\rangle| = ||\mathbf{v}|| ||\mathbf{w}||$ if and only if $\mathbf{v}$ and $\mathbf{w}$ are scalar multiples of each other. Show that $||\mathbf{v} + \mathbf{w}|| = ||\mathbf{v}|| + ||\mathbf{w}||$ if and only if $\mathbf{v}$ and $\mathbf{w}$ are positive scalar multiples of each other.

2. [4, 3.3.2] Show that for all f, g in a Hilbert space one has

$$|||f|| - ||g||| \leq ||f - g||.$$

3. [4, 3.3.9] Suppose $\{\phi_n\}$ is an orthonormal basis for a Hilbert space. Show that for any f, g in the Hilbert space

$$\langle f, g\rangle = \sum \langle f, \phi_n\rangle \overline{\langle g, \phi_n\rangle}.$$

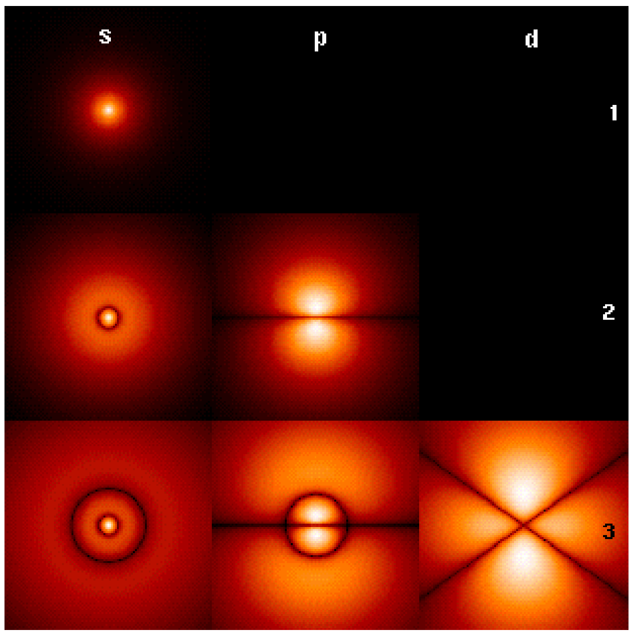


Figure 3.11: Here are the first few hydrogen atom orbitals. These are the cross-sections of the probability density that are color-coded with dark corresponding to zero density and light corresponding to highest density. The angular momentum quantum number, ℓ, is denoted in each column using the spectroscopic letter code (s means $\ell = 0$; p means $\ell = 1$; d means $\ell = 2$). The main quantum number $n = 1, 2, 3, \ldots$ is marked on the right side of each row. For all pictures the magnetic quantum number $m = 0$, and the cross-sectional plane is in the $x - z$ plane. Image source: `https://en.wikipedia.org/wiki/File:HAtomOrbitals.png` and license `https://creativecommons.org/licenses/by-sa/3.0/deed.en`.

4. Here, the elements of our Hilbert space are functions that are defined on $(0, 1)$. We would like to define the scalar product as follows

$$\langle f, g\rangle = \int_0^1 f^2(x)g^2(x)dx.$$

Why will this definition fail to satisfy the requisite properties of the scalar product?

5. [4, 3.4.3] Let D be the unit disk $\{x^2 + y^2 \leq 1\}$ and let $f_n(x, y) = (x + iy)^n$. Show that $\{f_n\}_{n\geq 0}$ is an orthogonal set in $\mathcal{L}^2(D)$, and compute $||f_n||$ for all n.

6. [4, 3.3.10.c] Evaluate the following series by applying Parseval's equation to one of the Fourier expansions in Table 3.1

$$\sum_{n\geq 1} \frac{n^2}{(n^2+1)^2}.$$

7. [4, 3.3.10.b] Evaluate the following series by applying Parseval's equation to one of the Fourier expansions in Table 3.1

$$\sum_{n\geq 1} \frac{1}{(2n-1)^6}.$$

8. Determine coefficients c_0, c_1, c_{-1} that minimize

$$\int_{-\pi}^{\pi} |e^x - c_0 - c_1 e^{ix} - c_{-1}e^{-ix}|^2 dx.$$

9. Show that for all $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ we have

$$||\mathbf{v} + \mathbf{w}||^2 + ||\mathbf{v} - \mathbf{w}||^2 = 2(||\mathbf{v}||^2 + ||\mathbf{w}||^2).$$

10. [4, 3.3.1] Show that if $\{f_n\}_{n\geq 1}$ are elements of a Hilbert space, H, and we have for some $f \in H$ that

$$\lim_{n\to\infty} f_n = f,$$

then for all $g \in H$ we have

$$\lim_{n\to\infty} \langle f_n, g\rangle = \langle f, g\rangle.$$

11. Let $f(x) = 1 + ix$ and $g(x) = 2 + ix^2$. Calculate the scalar products $\langle f, g\rangle$ and $\langle g, f\rangle$ on $\mathcal{L}^2(0, 1)$.

12. [4, 3.4.1] Show that $\{e^{2\pi i(mx+ny)}\}_{n,m\in\mathbb{Z}}$ is an orthogonal set in $\mathcal{L}^2(R)$ where R is any square whose sides have length one and are parallel to the coordinate axes.

13. Evaluate the following series by applying Parseval's equation to one of the Fourier expansions in Table 3.1

$$\sum_{n\geq 1} \frac{1}{n^4}.$$

14. [4, 3.3.10.d] Evaluate the following series by applying Parseval's equation to one of the Fourier expansions in Table 3.1

$$\sum_{n\geq 1} \frac{\sin^2(na)}{n^4}.$$

15. Determine coefficients c_0, c_1, c_{-1} that minimize

$$\int_{-\pi}^{\pi} |\cosh(x) - c_0 - c_1 e^{ix} - c_{-1} e^{-ix}|^2 dx.$$

16. [4, 3.3.9] Assume that $\{\phi_n\}_{n\geq 1}$ is an orthonormal basis for $\mathcal{L}^2(a,b)$. Show that for any $f, g \in \mathcal{L}^2(a,b)$

$$\langle f, g \rangle = \sum_{n\geq 1} \langle f, \phi_n \rangle \overline{\langle g, \phi_n \rangle}.$$

17. Assume that f is continuously differentiable, 2π periodic, and real-valued. Is $f \in \mathcal{L}^2(-\pi,\pi)$? Is f' in $\mathcal{L}^2(-\pi,\pi)$? Can you use the assumptions on f to deduce anything about $\langle f, f' \rangle$?

18. This is something I never figured out, but admittedly did not spend a ton of time trying to solve. Can you find a Fourier series that can be used to compute

$$\sum_{n\geq 1} \frac{1}{n^3}?$$

19. In this exercise we will prove that a Hilbert space H is separable if and only if it has an orthonormal basis that is finite or countable. Separable means that there is a countable subset $\{q_n\}_{n\in\mathbb{N}}$ such that for any $x \in H$ and any $\epsilon > 0$, there is some $m \in \mathbb{N}$ such that $||q_m - x|| < \epsilon$. This property is called *dense.* One of the ways the real numbers are often defined is that it is the smallest well-ordered set that contains the rational numbers and has the property that every non-empty subset of the real numbers that is bounded above has a least upper bound contained in the real numbers. For more about this, see [21, Chapter 3].

   (a) The rational numbers $\mathbb{Q}$ are dense in the set of real numbers; that is given any $\epsilon > 0$ and real number x there is a rational number q such that $|x - q| < \epsilon$. Use this to prove that the set

   $$\mathbb{Q} + i\mathbb{Q} = \{z \in \mathbb{C} : z = q + ir, \quad i = \sqrt{-1}, \quad q, r \in \mathbb{Q}\}$$

   is dense in the Hilbert space $\mathbb{C}$.

   (b) Use the preceding exercise to prove that every finite dimensional Hilbert space is separable. Note here that you may use the fact that a finite union of countable sets is countable.

   (c) Assume that a Hilbert space is separable. Denote the countable dense subset as $\{q_n\}_{n\in\mathbb{N}}$ above. Discard any elements of this set that are equal to zero. Here, we will use this set to build a countable orthonormal basis. Define

   $$v_1 := \frac{q_1}{||q_1||}.$$

   If q_2 is equal to a scalar multiple of v_1, then we throw it away and replace it by q_3. Define inductively in this way

   $$v_{n+1} = \frac{q_{n+1} - \sum_{j=1}^{n} \langle q_{n+1}, v_j \rangle v_j}{||q_{n+1} - \sum_{j=1}^{n} \langle q_{n+1}, v_j \rangle v_j||}.$$

   Prove by induction that the set $\{v_n\}_{n\geq 1}$ is orthonormal. Why is this set either finite or countable?

   (d) To prove that the set $\{v_n\}_{n\geq 1}$ is a basis, assume that some $x \in H$ satisfies $\langle x, v_n \rangle = 0$ for all n. Show first that this implies that $\langle x, q_n \rangle = 0$ for all n. Next, show that for any $\epsilon > 0$, $||x|| < \epsilon$. Use this together with the three equivalent conditions to be an orthonormal basis to conclude that $\{v_n\}_{n\geq 1}$ is an orthonormal basis for H.

(e) Assume now that a Hilbert space has a countable orthonormal basis. Denote this basis by $\{e_n\}_{n\in\mathbb{N}}$. Consider sets of the form

$$Q_N := \{\sum_{n=1}^{N} q_n e_n : q_n \in \mathbb{Q} + i\mathbb{Q}\}.$$

Why are these sets countable? Using the fact that a countable union of countable sets is countable,

$$\bigcup_{N\in\mathbb{N}} Q_N = Q \subset H \text{ is countable.}$$

Prove that Q is also dense in H and thereby conclude that H is separable.

20. We can define the Hilbert space $\mathcal{L}^2$ for unbounded intervals as well, like[3]

$$\mathcal{L}^2(0,\infty) = \{f : (0,\infty) \to \mathbb{C}, \quad \int_0^\infty |f(x)|^2 dx < \infty\}$$

Can you find a sequence of functions $\{f_n\}$ in $\mathcal{L}^2(0,\infty)$ such that for all $x > 0$ we have $\lim_{n\to\infty} f_n(x) = 0$, *but* the $\mathcal{L}^2$ norms, $||f_n|| = \sqrt{\int_0^\infty |f_n(x)|^2 dx} \not\to 0$?

[3]Of course, this is a 'working definition,' whereas the concise definition is almost-everywhere equivalence classes of measurable functions.

1.	$f(x)=x$	$2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin(nx)$.
2.	$f(x)=\|x\|$	$\frac{\pi}{2}-\frac{4}{\pi}\sum_{n\geq 1}\frac{\cos((2n-1)x)}{(2n-1)^2}$
3.	$f(x)=\begin{cases}0 & -\pi<x<0\\ x & 0<x<\pi\end{cases}$	$\frac{\pi}{4}-\frac{2}{\pi}\sum_{n\geq 1}\frac{\cos((2n-1)x)}{(2n-1)^2}+\sum_{n\geq 1}\frac{(-1)^{n+1}}{n}\sin(nx)$
4.	$f(x)=\sin^2(x)$	$\frac{1}{2}-\frac{1}{2}\cos(2x)$
5.	$f(x)=\begin{cases}-1 & -\pi<x<0\\ 1 & 0<x<\pi\end{cases}$	$\frac{4}{\pi}\sum_{n\geq 1}\frac{\sin((2n-1)x)}{2n-1}$
6.	$f(x)=\begin{cases}0 & -\pi<x<0\\ 1 & 0<x<\pi\end{cases}$	$\frac{1}{2}+\frac{2}{\pi}\sum_{n\geq 1}\frac{\sin((2n-1)x)}{2n-1}$
7.	$f(x)=\|\sin(x)\|$	$\frac{2}{\pi}-\frac{4}{\pi}\sum_{n\geq 1}\frac{\cos(2nx)}{4n^2-1}$
8.	$f(x)=\|\cos(x)\|$	$\frac{2}{\pi}-\frac{4}{\pi}\sum_{n\geq 1}\frac{(-1)^n\cos(2nx)}{4n^2-1}$
9.	$f(x)=\begin{cases}0 & -\pi<x<0\\ \sin(x) & 0<x<\pi\end{cases}$	$\frac{1}{\pi}-\frac{2}{\pi}\sum_{n\geq 1}\frac{\cos(2nx)}{4n^2-1}+\frac{1}{2}\sin(x)$
10.	$f(x)=x^2$	$\frac{\pi^2}{3}+4\sum_{n\geq 1}\frac{(-1)^n\cos(nx)}{n^2}$
11.	$f(x)=x(\pi-\|x\|)$	$\frac{8}{\pi}\sum_{n\geq 1}\frac{\sin((2n-1)x)}{(2n-1)^3}$
12.	$f(x)=e^{bx}$	$\frac{\sinh(b\pi)}{\pi}\sum_{n\in\mathbb{Z}}\frac{(-1)^n}{b-in}e^{inx}$
13.	$f(x)=\sinh x$	$\frac{2\sinh(\pi)}{\pi}\sum_{n\geq 1}\frac{(-1)^{n+1}n}{n^2+1}\sin(nx)$

Table 3.1: Here is a small collection of trigonometric Fourier expansions for the functions defined on $(-\pi,\pi)$. The series on the right are all 2π periodic functions, so the graph of these functions looks like the graph of $f(x)$ on $(-\pi,\pi)$, and then copy-pasted repeatedly over the rest of the real line. Here we collect expansions of functions in $\mathcal{L}^2(-\pi,\pi)$ in terms of the orthogonal basis $\{1,\cos(nx),\sin(nx)\}_{n\geq 1}$.

Chapter 4

Trigonometric Fourier series and their applications: rock and roll, hot rods, and π falling out of the sky!

Let's go back to the beginning and recall our very first example: a vibrating string of length ℓ whose ends are held fixed. Now we can really rock and roll!

4.1 Rock and roll for real: completing the solution to the vibrating string problem

The string length is ℓ. We defined

$$u(x,t) := \text{ the height of the string at the point } x \in [0,\ell] \text{ at time } t \in [0,\infty[.$$

The sitting-still height to be height 0, so if the string is not moving at all, then we would just have $u(x,t) = 0$ for all points x and all times t. To distinguish between up and down, we use positive and negative numbers. So a height $u(x,t) > 0$ is above the sitting-still position, whereas a height $u(x,t) < 0$ is below the sitting-still position.

The fact that ends are sitting still means that

$$u(0,t) = u(\ell,t) = 0 \quad \forall t.$$

The wave equation according to the laws of physics says that:

$$u_{tt} = \mathfrak{c}^2 u_{xx}.$$

We used separation of variables to obtain a solution of the form

$$u(x,t) = \sum_{k \in \mathbb{Z}} u_k(x,t) = \sum_{k \in \mathbb{Z}} A_k \sin\left(\frac{k\pi x}{\ell}\right)\left(B_k \cos\left(\frac{k\pi t\mathfrak{c}}{\ell}\right) + C_k \sin\left(\frac{k\pi t\mathfrak{c}}{\ell}\right)\right).$$

We said that these coefficients will be determined by the two initial conditions, the position of the string at time zero and its velocity

$$u(x,0) = f(x), \quad u_t(x,0) = g(x).$$

Now we can finally do the last step: TIDGLAS (the initial data goes last). We begin with a simplification by observing that since sine is odd, and cosine is even, we can combine the terms with $\pm k$ together:

$$A_k \sin\left(\frac{k\pi x}{\ell}\right)\left(B_k \cos\left(\frac{k\pi t\mathfrak{c}}{\ell}\right) + C_k \sin\left(\frac{k\pi t\mathfrak{c}}{\ell}\right)\right) + A_{-k} \sin\left(\frac{-k\pi x}{\ell}\right)\left(B_{-k} \cos\left(\frac{-k\pi t\mathfrak{c}}{\ell}\right) + C_{-k} \sin\left(\frac{-k\pi t\mathfrak{c}}{\ell}\right)\right)$$

$$= \sin\left(\frac{k\pi x}{\ell}\right)\left[(A_k B_k - A_{-k}B_{-k})\cos\left(\frac{k\pi t\mathfrak{c}}{\ell}\right) + (A_k C_k + A_{-k}C_{-k})\sin\left(\frac{k\pi t\mathfrak{c}}{\ell}\right)\right].$$

So, if we just use new coefficients

$$a_k := (A_k B_k - A_{-k}B_{-k}), \quad b_k = (A_k C_k + A_{-k}C_{-k}),$$

then our solution is equal to

$$u(x,t) = \sum_{k\geq 1}\sin\left(\frac{k\pi x}{\ell}\right)\left(a_k\cos\left(\frac{k\pi t\mathfrak{c}}{\ell}\right) + b_k\sin\left(\frac{k\pi t\mathfrak{c}}{\ell}\right)\right).$$

The first initial condition is that we wish

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

We set $t = 0$ in the series, so we wish that

$$\sum_{k\geq 1}\sin\left(\frac{k\pi x}{\ell}\right)a_k = f(x).$$

The functions $\sin\left(\frac{k\pi x}{\ell}\right)$ are an orthogonal basis! The reason for this is because they are all of the solutions to the eigenvalue problem, that is to find all λ and all X that satisfy

$$X''(x) = \lambda X(x), \quad X(0) = X(\ell) = 0.$$

This problem was the result of separating variables. This is an example of a *regular Sturm-Liouville problem,* that we will study in more detail and generality in chapter 6 For now, it suffices to know that the solutions we found and enumerated as $\sin(n\pi x/\ell)$ are an orthogonal basis for the $\mathcal{L}^2(0,\ell)$. Consequently, we can express every other member of this Hilbert space in terms of this basis!

If the function $f(x)$ represents the initial positions of the string at time $t = 0$, then it must be continuous, because the string is connected, not broken apart presumably. So, since f is continuous, it is also bounded on the interval $[0,\ell]$ and is a proud member of the Hilbert space $\mathcal{L}^2(0,\ell)$. So f can be expressed in terms of this orthogonal basis! The coefficients are therefore:

$$a_k = \frac{\langle f, \sin(k\pi x/\ell)\rangle}{||\sin(k\pi x/\ell)||^2} = \frac{\int_0^\ell f(x)\overline{\sin(k\pi x/\ell)}dx}{\int_0^\ell \sin^2(k\pi x/\ell)dx}.$$

When you have a specific f, you can solve for these coefficients by integrating. For general f, it suffices to specify these coefficients in this way. We use a similar procedure to find the coefficients b_k. Taking the time derivative and setting $t = 0$, we wish that

$$u_t(x,0) = g(x) \iff \sum_{k\geq 1}\sin\left(\frac{k\pi x}{\ell}\right)b_k\frac{k\pi\mathfrak{c}}{\ell} = g(x).$$

If $g(x)$ is the velocity along the string at time $t = 0$, then by the same physical reasoning (no crazy ass string action!) this function should also be continuous, hence bounded, and a proud member of the Hilbert space $\mathcal{L}^2(0,\ell)$. We can therefore expand g in terms of this orthogonal basis as

$$\sum_{k\geq 1}\hat{g}_k\sin\left(\frac{k\pi x}{\ell}\right), \quad \hat{g}_k = \frac{\langle g, \sin(k\pi x/\ell)\rangle}{||\sin(k\pi x/\ell)||^2} = \frac{\int_0^\ell g(x)\overline{\sin(k\pi x/\ell)}dx}{\int_0^\ell \sin^2(k\pi x/\ell)dx}.$$

Consequently, we obtain the coefficients b_k via:

$$b_k\frac{k\pi\mathfrak{c}}{\ell} = \hat{g}_k \implies b_k = \frac{\hat{g}_k\ell}{k\mathfrak{c}\pi}.$$

We have therefore solved the problem and can enjoy the rock and roll like in Figure 4.1.

Figure 4.1: Here is my friend playing with the band Sodom. There is a whole lot of sound due to vibrating strings! Photograph copyright Moritz 'Mumpi' Künster.

4.2 It's getting hot in here: completing the solution for heat flow on a circular rod

It is finally time for the initial data that we have been saving until last: The Initial Data Goes LASt. We consider the flow of heat, corresponding to an initial temperature distribution on a circular rod. In mathematical terms, we are solving the *initial value problem for the homogeneous heat equation.* The temperature of a circular rod is initially given by

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

and for times $t > 0$ the temperature function satisfies the heat equation, that is the partial differential equation

$$u_t = ku_{xx},$$

for a constant $k > 0$. The fact that the rod is circular means that we can define the function $f(x)$ for all real x by simply demanding that

$$f(x + 2\pi) = f(x).$$

This is because the angles x and $x + 2\pi$ correspond to the exact same point on the circular rod. Similarly, the same is also true for the function u, that is

$$u(x + 2\pi, t) = u(x, t),$$

for all x and all $t > 0$. We used the two methods: separation of variables and superposition to obtain

$$u(x, t) = \sum_{n\in\mathbb{Z}} u_n(x, t) = \sum_{n\in\mathbb{Z}} e^{-n^2tk}(a_n \cos(nx) + b_n \sin(nx)).$$

Our mantra was TIDGLAS: the initial data goes last.

Now it is time to determine these coefficients! First we make a simplification using the facts that cosine is even and sine is odd, as well as the fact that $(-n)^2 = n^2$:

$$e^{-n^2tk}(a_n \cos(nx) + b_n \sin(nx)) + e^{-(-n)^2tk}(a_{-n} \cos(-nx) + b_{-n} \sin(-nx))$$
$$= e^{-n^2tk} \left[(a_n + a_{-n}) \cos(nx) + (b_n - b_{-n}) \sin(nx)\right].$$

If we then define new coefficients $\alpha_n = a_n + a_{-n}$, $\beta_n = b_n - b_{-n}$, then our solution takes the form

$$u(x, t) = \sum_{n\geq 0} e^{-n^2tk}(\alpha_n \cos(nx) + \beta_n \sin(nx)).$$

Figure 4.2: There is a bit of magic surrounding Fourier series and their applications. These series were introduced by Joseph Fourier as a method for solving the heat equation [5]. The solution is expressed as a series of trigonometric functions. Fourier did not have the mathematical tools to prove that these trigonometric series could represent any arbitrary function; that is what Hilbert provided with the development of Hilbert spaces in general and $\mathcal{L}^2$ spaces in particular. It does not seem obvious that any arbitrary function on a bounded interval can be expressed as a series of trigonometric facts, and many scientists were sceptical that Fourier's method would be able to solve our 'hot rod' problem for arbitrary initial data. So, it's a bit like the cow jumping over the moon, the miraculous fact that we can express all square integrable functions on bounded interval as trigonometric Fourier series and thereby solve the 'hot rod' problem for arbitrary initial data! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

.

We would like this to equal $f(x)$ when $t = 0$, so we would like to choose α_n and β_n so that

$$u(x,0) = \sum_{n \geq 0} (\alpha_n \cos(nx) + \beta_n \sin(nx)) = f(x).$$

The function f is the initial temperature of the rod. If there is no freaky weird quantum stuff happening, then this should be a continuous function. Consequently, it is bounded and therefore a proud member of the Hilbert space $\mathcal{L}^2(-\pi, \pi)$, so we can express it in terms of the orthogonal basis $\{1, \cos(nx), \sin(nx)\}_{n \geq 1}$. The coefficients are:

$$\alpha_0 = \frac{\int_{-\pi}^{\pi} f(x)}{2\pi}, \quad \alpha_n = \frac{\langle f, \cos(nx) \rangle}{|| \cos(nx)||^2} = \frac{\int_{-\pi}^{\pi} f(x)\overline{\cos(nx)}dx}{\int_{-\pi}^{\pi} |\cos(nx)|^2 dx},$$

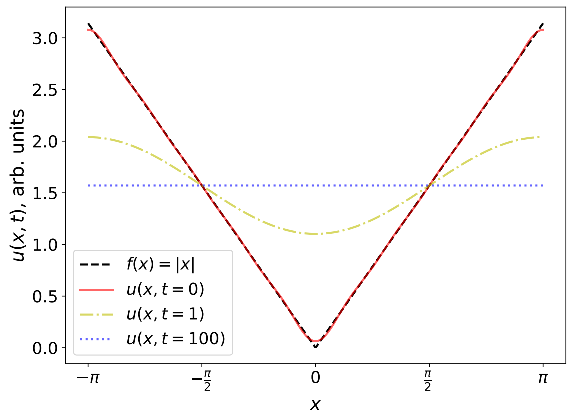
$$\beta_n = \frac{\langle f, \sin(nx) \rangle}{|| \sin(nx)||^2} = \frac{\int_{-\pi}^{\pi} f(x)\overline{\sin(nx)}dx}{\int_{-\pi}^{\pi} |\sin(nx)|^2 dx}.$$

In this way, we can solve the hot rod problem, that is the initial value problem for the homogeneous heat equation on a circular rod, for any continuous (or even just $\mathcal{L}^2$) initial data! That is rather miraculous, like the cow jumping over the moon with a hot rod in Figure 4.2.

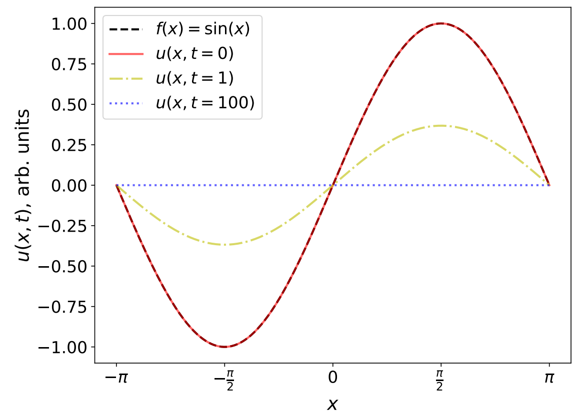
Figure 4.3: Berlioz the cat friend of Eric Lindgren is also a contributor to this text.

4.2.1 Visualization of the solution to the heat equation on a rod

For practical applications, one may compute the solution as a series, and if the series converges quickly, approximate the solution using the first several terms in the series. Here we show the graphs of the truncated series corresponding to solving the heat equation on a circular rod with different choices of initial data. These figures and the python script used to create them were contributed by Eric Lindgren, with help from his distinguished cat Berlioz. The python script used to create these simulations and figures is contained in §B.2.



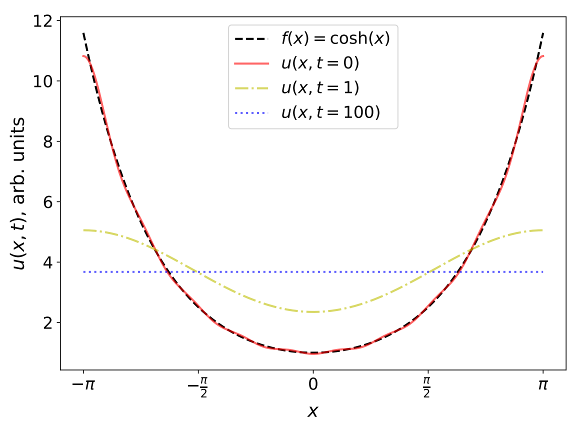
(a) $f(x) = |x|$.



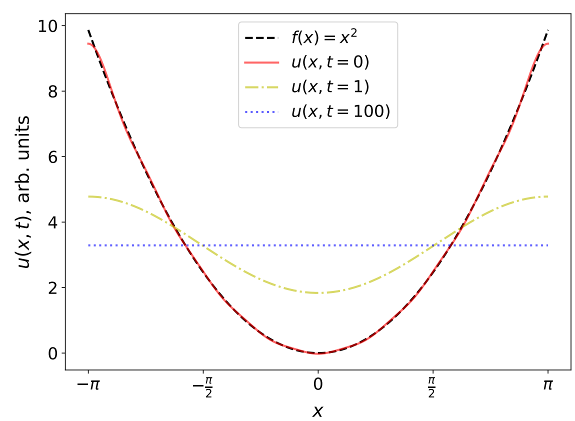
(b) $f(x) = \sin(x)$

Figure 4.4: The heat equation on a 'hot rod' with initial data $f(x)$.

The figures here show the numerical approximation of the solution of the heat equation on a circular rod with initial data $f(x)$. The approximation is obtained by taking the first ten terms in the series expansion of the solution. As time goes on, just as physics predicts, the heat within the rod disperses evenly throughout the rod.



(a) $f(x) = \cosh(x)$.



(b) $f(x) = x^2$

Figure 4.5: The heat equation on a 'hot rod' with initial data $f(x)$.

4.3 What happens to trigonometric Fourier series outside $(-\pi, \pi)$?

The sets of functions $\{e^{inx}\}_{n\in\mathbb{Z}}$ and $\{1, \cos(nx), \sin(nx)\}_{n\geq 1}$ are two orthogonal bases for the Hilbert space $\mathcal{L}^2(-\pi, \pi)$. We have used them to express elements of this Hilbert space as trigonometric series, either of the form

$$\sum_{n\in\mathbb{Z}} c_n e^{inx}$$

or of the form

$$c_0 + \sum_{n\geq 1} a_n \cos(nx) + b_n \sin(nx),$$

depending on the context. For an $f \in \mathcal{L}^2(-\pi, \pi)$, the coefficients are

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{e^{inx}} dx,$$

and

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\overline{\cos(nx)} dx, \quad n \geq 1,$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\overline{\sin(nx)} dx, \quad n \geq 1.$$

Then the series is equal to f as an element of the Hilbert space $\mathcal{L}^2(-\pi, \pi)$. This is only looking within the interval $(-\pi, \pi)$. What happens if we plug in points x to the series

$$\sum_{n\in\mathbb{Z}} c_n e^{inx}, \quad c_0 + \sum_{n\geq 1} a_n \cos(nx) + b_n \sin(nx),$$

for x outside this interval? Well, these series are 2π periodic because

$$e^{in(x+2\pi)} = e^{inx}e^{in2\pi} = e^{inx}, \quad \cos(n(x+2\pi)) = \cos(nx), \quad \sin(n(x+2\pi)) = \sin(nx).$$

So, outside the interval, these series converge to the function that is equal to f in $(-\pi, \pi)$ and is extended to the rest of the real line to be 2π periodic. To visualize this function, take the graph of f in $(-\pi, \pi)$ and copy-paste it left and right, on $(-3\pi, -\pi)$ and $(\pi, 3\pi)$. Keep copy-pasting. What happens if $f(-\pi) \neq f(\pi)$? There will be a jump in the graph. It's not super obvious to what the Fourier series will converge, and the theorem that reveals the answer takes quite a lot of work to prove. One ingredient in the proof is a general property of periodic functions given in Lemma 41.

Definition 40. A function $f : \mathbb{R} \to \mathbb{R}$ is periodic with (minimal) period p if and only if for all $x \in \mathbb{R}$, $f(x+p) = f(x)$, and moreover, $p > 0$ is the smallest real number for which this is true. Then, such a function is also np for any positive integer n, but for $n > 1$, np is no longer the minimal period.

For example, $\sin(x)$ is periodic with minimal period 2π. Our heat equation examples, $f_n(x) = a_n \cos(nx) + b_n \sin(nx)$ are periodic with minimal period $2\pi/n$. They are also 2π periodic, it's just that this is not the minimal period. We shall prove a super useful little lemma about periodic functions and their integrals.

Lemma 41 (Integration of periodic functions lemma)**.** *If f is p periodic then for any $a \in \mathbb{R}$*

$$\int_a^{a+p} f(x)dx$$

is the same.

Exercise 42. *Give an example for how this fails to be true if the function f is* not *periodic. That is, take some non-periodic function and show that integrating it from say a to $a+p$ is not the same as integrating it from c to $c+p$.*

Proof: If we think about it, we want to show that the function

$$g(a) := \int_a^{a+p} f(x)dx$$

is a constant function. This looks awfully similar to the fundamental theorem of calculus. Now, this statement above is not true for non-periodic functions. So, we're going to need to use the assumption that f is periodic with period p. This tells us that f has the same value at both endpoints of the integral, so

$$f(a) = f(a+p) \implies f(a+p) - f(a) = 0.$$

Now, since we want to consider a as a variable, we don't want it at both the top and the bottom of the integral defining g. Instead, we can use linearity of integration to write

$$g(a) = \int_0^{a+p} f(x)dx - \int_0^{a} f(x)dx.$$

Then, using the fundamental theorem of calculus on each of the two terms on the right,

$$g'(a) = f(a+p) - f(a) = 0.$$

Above, we use the fact that f is periodic with period p. Hence, $g'(a) \equiv 0$ for all $a \in \mathbb{R}$. This tells us that g is a constant function, so its value is the same for all $a \in \mathbb{R}$.

4.3.1 Pointwise convergence of Fourier Series: don't let your work go to waste, with Fourier series just copy-paste!

A trigonometric Fourier series will always converge to a 2π periodic function, because the terms in the series are all 2π periodic. We can expand any arbitrary function in $\mathcal{L}^2(-\pi, \pi)$ in a Fourier series on that interval. Super. It is crucial to remember that if we start looking at the series *outside* that interval it will be a 2π periodic function; that is it will look like the graph of f on $(-\pi, \pi)$ has been copied and pasted all across the real line; see Figure 4.6!

We will prove the pointwise convergence of Fourier series for functions that are:

1. continuous on $(-\pi, \pi)$ with the exception of at most finitely many points at which there is a jump discontinuity (the left and right limits both exist and are finite, just not equal);
2. differentiable on $(-\pi, \pi)$ with the exception of at most finitely many points at which there is a jump discontinuity (the left and right limits both exist and are finite, just not equal).

Functions which satisfy these two conditions are called *piecewise* $\mathcal{C}^1$. More generally, we can define *piecewise* $\mathcal{C}^k$ *functions* in an analogous way.

Definition 43. A function is piecewise $\mathcal{C}^k$ on a bounded interval, I, if there is a set $S \subset I$ such that f is $\mathcal{C}^k$ on $I \setminus S$, and the set S is either empty or contains finitely many points. Moreover, at each of the points in S the left and right limits of $f^{(j)}$ exist and are finite for all $j = 0, 1, \ldots, k$.

The theorem on the pointwise convergence of Fourier series states that for a function f that is piecewise $\mathcal{C}^1$ on $[-\pi, \pi]$, the series converges to a function that is 2π periodic on $\mathbb{R}$. Within the interval $[-\pi, \pi]$, the series converges to $f(x)$ if f is continuous at x, or to the average of the left and right sided limits at x, if f has a jump discontinuity there.

Theorem 44 (Pointwise convergence of Fourier series)**.** *Assume that f is piecewise $\mathcal{C}^1$ on $[-\pi, \pi]$. Then, copy-paste to the rest of $\mathbb{R}$ so that f satisfies $f(x + 2\pi) = f(x)$ for all x; more precisely extend f to a 2π periodic function on $\mathbb{R}$. Denote the left limit at x by $f(x_-)$ and the right limit by $f(x_+)$, so that for each $x \in \mathbb{R}$,*

$$f(x_-) := \lim_{t \to x, t<x} f(t), \quad f(x_+) := \lim_{t \to x, t>x} f(t).$$

Let

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$$

Then

$$\lim_{N\to\infty} \sum_{-N}^{N} c_n e^{inx} = \frac{1}{2} \left(f(x_-) + f(x_+) \right), \quad \forall x \in \mathbb{R}.$$

A visualization of the theorem

One says that a picture is worth a thousand words. Perhaps a simulation accompanied by a picture is equally valuable! When we create a Fourier series of a function on the interval, like for example $|x|$ on the interval $(-\pi, \pi)$, the series is of the form

$$\sum_{n\in\mathbb{Z}} c_n e^{inx} \text{ or equivalently } \sum_{n\geq 0} a_n \cos(nx) + b_n \sin(nx).$$

Note that the terms in the sum all satisfy:

$$c_n e^{in(x+2\pi)} = c_n e^{inx}, \quad a_n \cos(n(x + 2\pi)) = a_n \cos(nx), \quad b_n \sin(n(x + 2\pi)) = b_n \sin(nx).$$

So, adding 2π to x does *not* change the series. This means that the series, that converges to $|x|$ for $|x| < \pi$, will *not* converge to $|x|$ for $|x| > \pi$. Instead, it will converge to a zig-zag function as illustrated in Figure 4.6. *Don't let your work go to waste; with Fourier series just copy-paste!*

Aided by the visualization, we will now proceed with the proof of the theorem of pointwise convergence of Fourier series.

Proof of the theorem on pointwise convergence of Fourier series

Proof: This is a big theorem, because it requires several clever ideas in the proof. Smaller theorems can be proven by just "following your nose." So, to try to help with the proof, we're going to highlight the big ideas. To learn the proof, you can start by learning all the big ideas in the order in which they're used. Once you've got these down, then try to fill in the math steps starting at one idea, working to get to the next idea. The big ideas are like light posts guiding your way through the dark and spooky math.

Idea!

Fix a point $x \in \mathbb{R}$. This first step is more getting into a frame of mind. Think of x as fixed. Then the numbers $f(x_-)$ and $f(x_+)$ are just the left and right limits of f at x, so these are also fixed. Our goal is to prove that:

$$\lim_{N\to\infty} \sum_{-N}^{N} c_n e^{inx} = \frac{1}{2} \left(f(x_-) + f(x_+) \right).$$

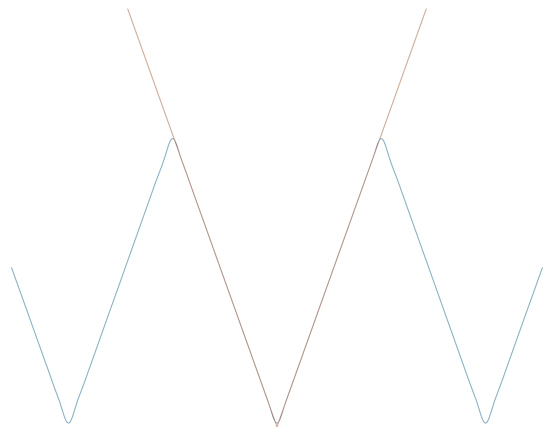


Figure 4.6: The Fourier series of the function $|x|$ on the interval $(-\pi, \pi)$ is shown here, together with the function $|x|$. Note that $|x|$ keeps climbing up to the left and the right outside the interval, whereas the Fourier series does a zig-zag pattern. The reason is because trigonometric Fourier series is 2π periodic. Don't let your work go to waste; with Fourier series just copy-paste! Thanks to Anton Rosén for contributing this figure as well as the matlab code! to create it in §B.1

This is completely equivalent to proving

$$\lim_{N\to\infty} \sum_{-N}^{N} c_n e^{inx} - \frac{1}{2}\left(f(x_-) + f(x_+)\right) = 0.$$

Let us call

$$\star = \sum_{-N}^{N} c_n e^{inx} - \frac{1}{2}\left(f(x_-) + f(x_+)\right).$$

Idea!

The main idea is to try to make the two things look like each other, that is we want to make $\sum c_n e^{inx}$ look like the average of the left and right limits of f. To get $\sum c_n e^{inx}$ looking more like f, we write out what each term in this sum really is:

$$c_n e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny} e^{inx} dy.$$

We must use a different variable for integration because x is a fixed point.

Idea!

We want to get some x inside of f to be able to relate to $\frac{f(x_+)+f(x_-)}{2}$. Make a change of variables to do this:

$$y = t + x, \quad dy = dt \implies c_n e^{inx} = \frac{1}{2\pi} \int\limits_{-\pi-x}^{\pi+x} f(t+x) e^{-int} dt.$$

Idea!

Slide the integral back to being from $-\pi$ to π because the integrand is 2π periodic, so the integral over any interval of length 2π is the same. Then we have

$$c_n e^{inx} = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(t+x) e^{-int} dt.$$

Idea!

Investigate the sum, because

$$\star = \sum_{-N}^{N} c_n e^{inx} - \frac{1}{2}\left(f(x_-) + f(x_+)\right) = \sum_{-N}^{N} \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(t+x) e^{-int} dt - \frac{1}{2}\left(f(x_-) + f(x_+)\right)$$

$$= \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(t+x) \sum_{-N}^{N} e^{int} dt - \frac{1}{2}\left(f(x_-) + f(x_+)\right).$$

So, let us see what we can say about

$$\sum_{-N}^{N} e^{int}.$$

Idea!

Use the fact that

$$e^{int} + e^{-int} = 2\cos(nt) \implies \sum_{-N}^{N} e^{int} = 1 + 2\sum_{1}^{N} \cos(nt)$$

to compute

$$\frac{1}{2} = \frac{1}{2\pi} \int\limits_{-\pi}^{0} \sum_{-N}^{N} e^{int} dt, \quad \frac{1}{2} = \frac{1}{2\pi} \int\limits_{0}^{\pi} \sum_{-N}^{N} e^{int} dt.$$

Idea!

Use this to express

$$\frac{1}{2}f(x_-) = \frac{1}{2\pi}\int_{-\pi}^{0} f(x_-)\sum_{-N}^{N} e^{int}dt, \quad \frac{1}{2}f(x_+) = \frac{1}{2\pi}\int_{0}^{\pi} f(x_+)\sum_{-N}^{N} e^{int}dt,$$

so that we can equivalently write

$$\begin{aligned}\star &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t+x)\sum_{-N}^{N} e^{int}dt - \frac{1}{2\pi}\int_{-\pi}^{0} f(x_-)\sum_{-N}^{N} e^{int}dt - \frac{1}{2\pi}\int_{0}^{\pi} f(x_+)\sum_{-N}^{N} e^{int}dt \\ &= \frac{1}{2\pi}\left(\int_{-\pi}^{0} (f(t+x) - f(x_-))\sum_{-N}^{N} e^{int}dt + \int_{0}^{\pi} (f(t+x) - f(x_+))\sum_{-N}^{N} e^{int}dt\right).\end{aligned}$$

Idea!

Return to that sum to see if we can simplify it somehow: $\sum_{-N}^{N} e^{int}$. This is like a geometric series, that can be explicitly computed

$$\sum_{-N}^{N} e^{int} = e^{-iNt}\sum_{0}^{2N} e^{int} = e^{-iNt}\frac{e^{i(2N+1)t} - 1}{e^{it} - 1} = \frac{e^{i(N+1)t} - e^{-iNt}}{e^{it} - 1},$$

having used the fact about geometric series that if $z \neq 1$ then

$$\sum_{0}^{M} z^m = \frac{z^{M+1} - 1}{z - 1}.$$

So,

$$\star = \frac{1}{2\pi}\left(\int_{-\pi}^{0} (f(t+x) - f(x_-))\frac{e^{i(N+1)t} - e^{-iNt}}{e^{it} - 1}dt + \int_{0}^{\pi} (f(t+x) - f(x_+))\frac{e^{i(N+1)t} - e^{-iNt}}{e^{it} - 1}dt\right).$$

Idea!

Collect everything except the $e^{i(N+1)t}$ and e^{-iNt} and use it to define a new function

$$g(t) := \begin{cases} \frac{f(t+x) - f(x_-)}{e^{it} - 1} & -\pi < t < 0 \\ \frac{f(t+x) - f(x_+)}{e^{it} - 1} & 0 < t < \pi \end{cases}$$

Check to see that this does not do anything terrible at $t = 0$ by evaluating its left and right limits. Note that at all other points in $[-\pi, \pi]$ g inherits the properties of f because the denominator is non-zero. We use l'hopital's rule to compute

$$\lim_{t \to 0^-} g(t) = \frac{f'(x_-)}{i}, \quad \lim_{t \to 0^+} g(t) = \frac{f'(x_+)}{i}.$$

These are just the left and right limits of f' at x which both exist since f is piecewise $\mathcal{C}^1$. So g is also piecewise $\mathcal{C}^1$ and therefore bounded and therefore also in $\mathcal{L}^2(-\pi, \pi)$.

Idea!

Recognize

$$\star = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)(e^{i(N+1)t} - e^{-iNt})dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{i(N+1)t}dt - \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-iNt}dt$$

are Fourier coefficients of g, specifically

$$\star = \hat{g}_{-N-1} - \hat{g}_N.$$

Idea!

Since g is in $\mathcal{L}^2$, we have the equality for the ONB $\{e^{inx}/\sqrt{2\pi}\}$

$$||g||^2 = \sum_{\mathbb{Z}} \left| \langle g, \frac{e^{inx}}{\sqrt{2\pi}} \rangle \right|^2 = \frac{1}{2\pi} \sum_{\mathbb{Z}} |\langle g, e^{inx} \rangle|^2.$$

The norm on the left is finite, so this series converges, and so the individual terms tend to zero as $|n| \to \infty$. Consequently

$$\hat{g}_{-N-1} = \frac{\langle g, e^{-i(N+1)x} \rangle}{2\pi} \to 0, \quad \hat{g}_N = \frac{\langle g, e^{iNx} \rangle}{2\pi} \to 0$$

as $N \to \infty$, thereby guaranteeing that $\star \to 0$ and completing the proof!

As a corollary to the theorem on the pointwise convergence of Fourier series, we obtain that piecewise $\mathcal{C}^1$ functions that have the same Fourier series are equal.

Corollary 45. *Assume that f and g are piecewise $\mathcal{C}^1$ on $[-\pi, \pi]$. Assume that at any point at which f is discontinuous, it satisfies*

$$f(x) = \frac{f(x_+) + f(x_-)}{2},$$

and the same is true for g. If f and g have the same Fourier coefficients, then $f(x) = g(x)$ for all $x \in (-\pi, \pi)$. Moreover, the extensions of f and g that are equal to f and g on $(-\pi, \pi)$ and extended to be 2π periodic on $\mathbb{R}$ are equal on $\mathbb{R}$.

Proof: By assumption, f and g have the same Fourier series. Let us write the partial series

$$S_N(x) = \sum_{-N}^{N} c_n e^{inx}.$$

By the theorem on the pointwise convergence of Fourier series,

$$\lim_{N\to\infty} S_N(x) = \frac{f(x_+) + f(x_-)}{2} = \frac{g(x_+) + g(x_-)}{2}, \quad \forall x \in \mathbb{R}. \tag{4.3.1}$$

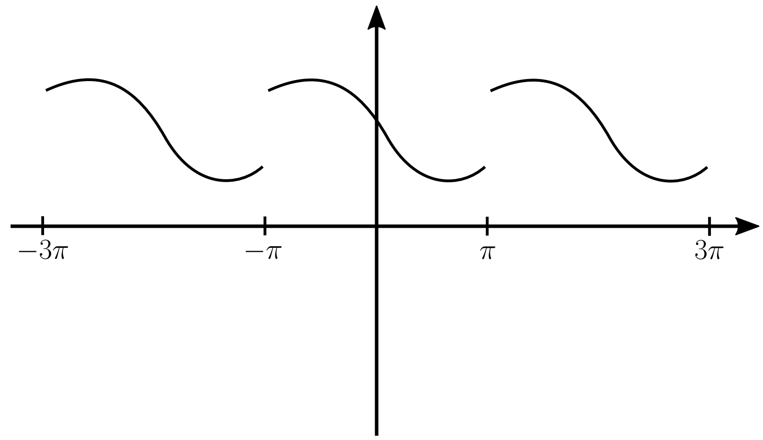


Figure 4.7: To illustrate what happens when we expand a function defined on an interval in a trigonometric Fourier series, I used Microsoft Word. I drew a little squiggly curve and set it on coordinate axes. Then, I selected my squiggly curve and copy+pasted it two times. This is exactly what happens mathematically when we expand an arbitrary function into a trigonometric Fourier series. The resulting series converges to our squiggle inside the interval. To obtain what happens outside the interval, we must *copy and paste* the picture inside the interval. This is an improved version of my original figure, thanks to Gottfrid Olsson.

Now, at a point where f is continuous,

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

Similarly, at a point where g is continuous

$$\frac{g(x_+) + g(x_-)}{2} = g(x).$$

So, by the assumptions on f and g, we have for all $x \in (-\pi, \pi)$

$$f(x) = \frac{f(x_+) + f(x_-)}{2}, \quad g(x) = \frac{g(x_+) + g(x_-)}{2}.$$

Thus, by (4.3.1),

$$f(x) = g(x) \quad \forall x \in (-\pi, \pi).$$

Consequently, their 2π periodic extensions are also equal on all of $\mathbb{R}$.

Don't let your work go to waste; with Fourier series just copy-paste!

To what does the Fourier series of e^x converge when x is not in the interval $(-\pi, \pi)$? We build a Fourier series for a function defined on the interval $(-\pi, \pi)$ of the form:

$$\sum_{n \in \mathbb{Z}} c_n e^{inx}.$$

Each of the terms e^{inx} is a 2π periodic function. Hence the Fourier series is also a 2π periodic function. So, for $x = 2\pi$, the series does *not* converge to $e^{2\pi}$. Rather, it converges to e^0 because, writing

$$S(x) = \sum_{n \in \mathbb{Z}} c_n e^{inx}, \quad S(x + 2k\pi) = S(x) \quad \forall k \in \mathbb{Z}.$$

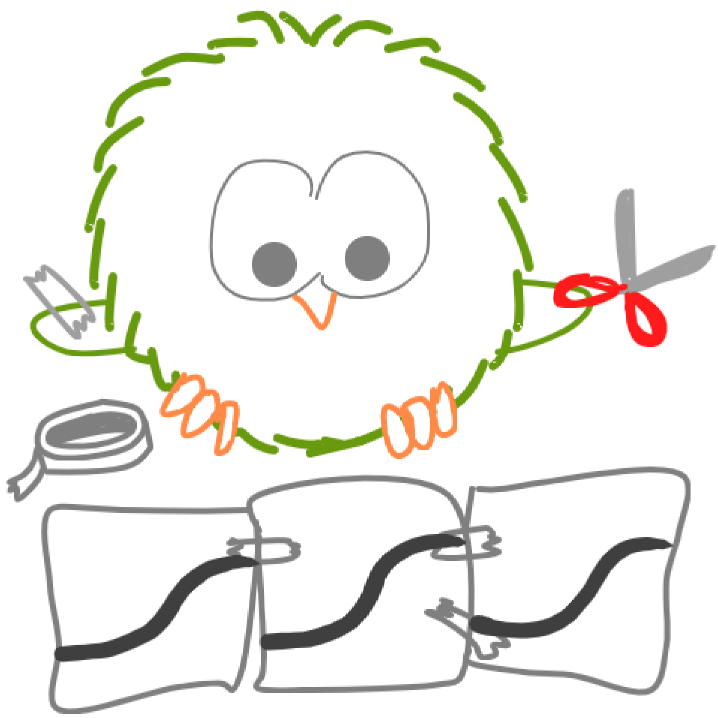


Figure 4.8: Here is a katodanode fluffball reminding us not to let our work go to waste; with Fourier series just copy-paste! Thanks to Ebba Grönfors for this contribution!

For $x \in (-\pi, \pi)$, by the Theorem we proved, we have that $S(x) = e^x$. However, for x outside this interval, the series converges to the function which is equal to e^x on $(-\pi, \pi)$ and is extended to be 2π periodic. Hence the series converges to the value at 0 since $2\pi = 0 + 2\pi$, and the series is 2π periodic. This is a really important subtlety. This phenomenon is depicted in Figure 4.7, that I created using Microsoft Word and *quite literally copy and pasting!* The original figure has been improved by Gottfrid Olsson, but it conveys the same message:

Don't let your work go to waste; with Fourier series just copy-paste!

4.4 Applying the pointwise convergence of Fourier series to compute sums and catch more π falling out of the sky!

We wish to compute the sum

$$\sum_{n=0}^{\infty} \frac{1}{1+n^2}.$$

It is not immediately obvious how to compute the sum above using a Fourier series, because there is the $1 + n^2$ downstairs, so somehow we need to get a one added downstairs. Moreover, this does not look like $|c_n|^2$, because that would require $c_n = \frac{1}{\sqrt{1+n^2}}$, and it is not very obvious how we could obtain such a square root in a Fourier coefficient. So, this indicates that to compute this series, we may seek to apply the theorem on the pointwise convergence. To find a function whose Fourier series will help us to compute the series, we should ponder the expression

$$\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

Let us again start with the most simple function $f(x)$ as possible, and then increase complexity as needed. For starters, if $f(x)$ were equal to just one, then we would be integrating e^{-inx}. For $n \neq 0$, a function whose derivative is e^{-inx} is

$$\frac{e^{-inx}}{-in}.$$

When we compute the integral, we evaluate at the two end points, so what we would get is $\frac{1}{-in}$ times some stuff involving n and π. That is close to what we want. However there is the $+1$ in the denominator. So, how can we get something more like $n+1$ downstairs? Well, we can put a $+1$ upstairs into our exponential function. That is, we integrate

$$\int_{-\pi}^{\pi} e^x e^{-inx}dx = \left.\frac{e^{(1-in)x}}{1-in}\right|_{x=-\pi}^{\pi}.$$

See, we end up with $1-in$ downstairs! For this reason, we choose the function $f(x) = e^x$ to expand in a Fourier series. We compute:

$$\int_{-\pi}^{\pi} e^x e^{-inx}dx = \left.\frac{e^{x(1-in)}}{1-in}\right|_{x=-\pi}^{x=\pi} = \frac{e^{\pi}e^{-in\pi}}{1-in} - \frac{e^{-\pi}e^{in\pi}}{1-in} = (-1)^n \frac{2\sinh(\pi)}{1-in}.$$

Hence, the Fourier coefficients are

$$\frac{1}{2\pi}(-1)^n \frac{2\sinh(\pi)}{1-in},$$

and the Fourier series for e^x on this interval is

$$e^x = \sum_{-\infty}^{\infty} \frac{(-1)^n \sinh(\pi)}{\pi(1-in)} e^{inx}, \quad x \in (-\pi, \pi).$$

We can pull out some constants,

$$e^x = \frac{\sinh(\pi)}{\pi} \sum_{-\infty}^{\infty} \frac{(-1)^n e^{inx}}{1-in}, \quad x \in (-\pi, \pi).$$

In the series we wish to evaluate, there is no $(-1)^n$. We would like this to go away. We can achieve that by selecting a value of x so that

$$(-1)^n e^{inx} = 1 \text{ for all } n, \text{ with a wisely chosen } x.$$

Recall Figure 3.9. If we choose $x = \pi$ or $x = -\pi$, then $e^{inx} = (-1)^n$. So, we make the choice say, $x = \pi$. Then the series is

$$\sum_{n\in\mathbb{Z}} \frac{1}{1-in}.$$

To what does this converge? The series is the trigonometric Fourier series expansion of the function e^x on the interval $(-\pi, \pi)$. We are looking at the value of the series at the point π. What happens here? Remember the copy-past procedure like in Figure 4.7. At the point π, we must copy-paste the graph of e^x on the interval $(-\pi, \pi)$. So, from the left as we approach π, we are approaching e^{π}. However, when we go past π, we jump down to the left of the graph on $(-\pi, \pi)$, and hence the graph jumps down to $e^{-\pi}$. The

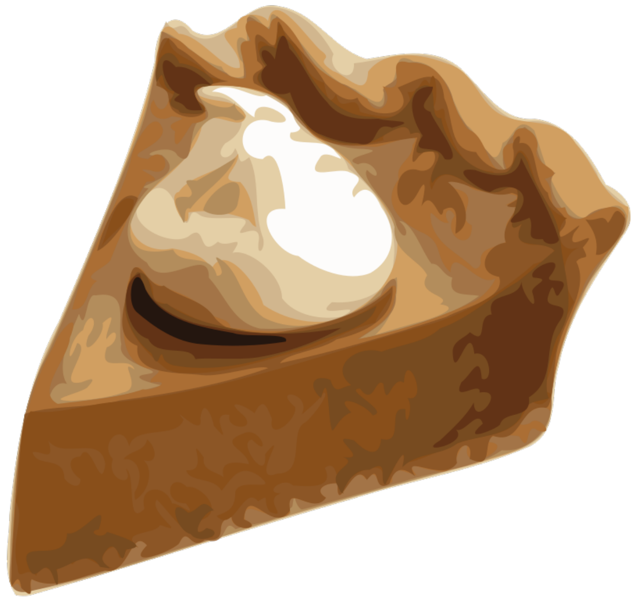


Figure 4.9: We have computed another series of the form $\sum_n \frac{1}{p(n)}$ where $p(n)$ is a polynomial function of the integer n. Again, it does not seem obvious that π would appear out of such a series, but indeed it does. This times it comes with an added topping of hyperbolic trigonometric functions, which we visualize as whipped cream on top of a piece of pumpkin pie. Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

Theorem on Pointwise Convergence of Fourier Series therefore tells us that the Fourier series converges to the average of these two limits, that is

$$\sum_{n\in\mathbb{Z}} \frac{1}{1-in} = \frac{e^{\pi}+e^{-\pi}}{2}.$$

We now consider the sum, and we pair together $\pm n$ for $n \in \mathbb{N}$, writing

$$\sum_{-\infty}^{\infty} \frac{1}{1-in} = 1 + \sum_{n\geq 1} \frac{1}{1-in} + \frac{1}{1+in} = 1 + \sum_{n\geq 1} \frac{2}{1+n^2}.$$

Hence we have found that

$$\frac{e^{\pi}+e^{-\pi}}{2} = \frac{\sinh(\pi)}{\pi} \sum_{-\infty}^{\infty} \frac{(-1)^n e^{in\pi}}{1-in} = \frac{\sinh(\pi)}{\pi} \left(1 + \sum_{n\geq 1} \frac{2}{1+n^2} \right).$$

Now, we just need to re-arrange. On the left we have the definition of $\cosh(\pi)$. So, moving over the $\sinh(\pi)$ we have

$$\frac{\pi\cosh(\pi)}{\sinh(\pi)} = 1 + 2\sum_{n\geq 1} \frac{1}{1+n^2} \implies \left(\frac{\pi\cosh(\pi)}{\sinh(\pi)} - 1 \right) \frac{1}{2} = \sum_{n\geq 1} \frac{1}{1+n^2},$$

and thus

$$\sum_{n\geq 0} \frac{1}{1+n^2} = \left(\frac{\pi\cosh(\pi)}{\sinh(\pi)} - 1 \right) \frac{1}{2} + 1.$$

Interestingly, in this sum we are not only treated with another appearance of π, but also with hyperbolic trigonometric functions, check out Figure 4.9!

4.4.1 Practice makes perfect: applying trigonometric series to compute another sum and catch some more π!

Let us do another example. We wish to use a Fourier series to compute

$$\sum_{n\geq 1} \frac{(-1)^n}{n^2+b^2}.$$

Since we had a similar series with b replaced by 1 for the function e^x, a natural function to try here is e^{bx}. We compute the Fourier coefficients with respect to the orthogonal basis $\{e^{inx}\}_{n\in\mathbb{Z}}$ on $(-\pi,\pi)$.

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{bx}e^{-inx}dx = \frac{1}{2\pi(b-in)}e^{(b-in)\pi} - \frac{1}{2\pi(b-in)}e^{(b-in)(-\pi)}.$$

As in Figure 3.9 that shows the values of $e^{\pm in\pi}$ for n even and odd, we can simplify,

$$c_n = \frac{1}{2\pi(b-in)}(-1)^n e^{b\pi} - \frac{1}{2\pi(b-in)}(-1)^n e^{-b\pi} = \frac{(-1)^n}{2\pi(b-in)}\left(e^{b\pi}-e^{-b\pi}\right) = \frac{(-1)^n}{\pi(b-in)}\sinh(b\pi).$$

The Fourier series is therefore

$$\frac{1}{\pi}\sinh(b\pi)\sum_{n\in\mathbb{Z}}\frac{(-1)^n}{b-in}e^{inx}.$$

In this case, we want to *keep* the $(-1)^n$, that is we do not want e^{inx} to screw it up. So, what value of x can we stuff into the series so that we preserve $(-1)^n$ for all n? We want $x=0$. The series is at this point

$$\frac{1}{\pi}\sinh(b\pi)\sum_{n\in\mathbb{Z}}\frac{(-1)^n}{b-in}.$$

Let us re-arrange things a wee bit:

$$\frac{1}{\pi}\sinh(b\pi)\sum_{n\in\mathbb{Z}}\frac{(-1)^n}{b-in} = \frac{\sinh(b\pi)}{\pi b} + \frac{1}{\pi}\sinh(b\pi)\sum_{n\geq 1}\frac{(-1)^n}{b-in} + \frac{1}{\pi}\sinh(b\pi)\sum_{n\leq 1}\frac{(-1)^n}{b-in}.$$

Let us re-write

$$\frac{1}{\pi}\sinh(b\pi)\sum_{n\leq 1}\frac{(-1)^n}{b-in} = \frac{1}{\pi}\sinh(b\pi)\sum_{n\geq 1}\frac{(-1)^n}{b+in},$$

with the observation that

$$(-1)^n = (-1)^{-n}.$$

Consequently the series is:

$$\frac{\sinh(b\pi)}{\pi b} + \frac{1}{\pi}\sinh(b\pi)\sum_{n\geq 1}\left(\frac{(-1)^n}{b-in}+\frac{(-1)^n}{b+in}\right)$$

$$= \frac{\sinh(b\pi)}{\pi b} + \frac{1}{\pi}\sinh(b\pi)\sum_{n\geq 1}(-1)^n\frac{b+in+b-in}{(b-in)(b+in)}$$

$$= \frac{\sinh(b\pi)}{\pi b} + \frac{1}{\pi}\sinh(b\pi)\sum_{n\geq 1}(-1)^n\frac{2b}{b^2+n^2}.$$

The Theorem on Pointwise Convergence of Fourier Series to say that at the point $x=0$ the Fourier series of this function converges to

$$\frac{f(0_+)+f(0_-)}{2}, \quad f(x)=e^{bx}$$

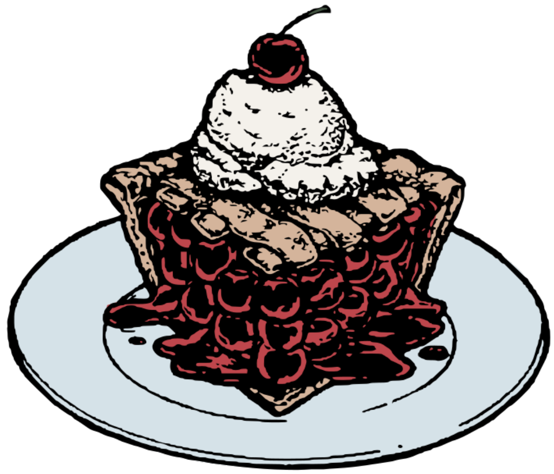


Figure 4.10: Having computed yet another series of the form $\sum \frac{1}{p(n}$ for a polynomial function of the integer n, we have yet again seen the result containing... π. We further have hyperbolic trigonometric functions, and a great generality because the value of b is arbitrary. So, we've earned a piece of cherry pie, representing the π that seemingly miraculously appears, together with the topping of hyperbolic trig function whipping cream, and a b-shaped cherry on top! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

because $0 \in (-\pi, \pi)$. The function is continuous here, with left and right limits both equal to its value at $x = 0$, and that is one. The series therefore converges to 1, and so

$$1 = \frac{\sinh(b\pi)}{\pi b} + \frac{1}{\pi}\sinh(b\pi)\sum_{n\geq 1}(-1)^n \frac{2b}{b^2+n^2}.$$

Re-arranging, we get

$$1 - \frac{\sinh(b\pi)}{\pi b} = \frac{2b\sinh(b\pi)}{\pi}\sum_{n\geq 1}\frac{(-1)^n}{b^2+n^2} \implies \frac{\pi}{2b\sinh(b\pi)} - \frac{1}{2b^2} = \sum_{n\geq 1}\frac{(-1)^n}{b^2+n^2}.$$

Having computed these seemingly impossible series, all of which involved some π, let's treat ourselves to a wacky educational video about π here: https://www.youtube.com/watch?v=XanjZw5hPvE. Warning: that video is an artefact of its time and may be considered offensive or inappropriate in current society. For a more wholesome reward, check out Figure 4.10.

4.5 Differentiating and integrating trigonometric Fourier series

If we take a function and obtain its Fourier series in terms of the bases $\{e^{inx}\}_{n\in\mathbb{Z}}$ or equivalently $\{1, \cos(nx), \sin(nx)\}_{n\in\mathbb{N}}$ on $\mathcal{L}^2(-\pi,\pi)$, the series we obtain is a 2π periodic function that is defined on all of $\mathbb{R}$. For this reason, when we study and demonstrate results about Fourier series, we may often assume that the function is 2π periodic to begin with. If we were only interested in the interval $(-\pi,\pi)$, then this extension does not change anything about our original function defined on $(-\pi,\pi)$. The reason for the 2π periodicity is simply because the Fourier series, while it may be equal to our function defined on $(-\pi,\pi)$ in that interval, will be equal to its 2π periodic extension outside that interval; keep in mind Figure 4.7!

Under certain assumptions there is a relationship between the Fourier coefficients of a function and the Fourier coefficients of its derivative. If the function fails to satisfies all of the assumptions of Theorem 46, then the conclusion need not be true.

Theorem 46. *Assume that f is 2π periodic, continuous, and piecewise $\mathcal{C}^1$. Let a_n, b_n, and c_n be the Fourier coefficients as we have defined them previously, and let a_n', b_n', c_n' be the Fourier coefficients of f' according to the same definition. Then we have*

$$a_n' = nb_n, \quad b_n' = -na_n, \quad c_n' = inc_n.$$

Proof: DO NOT DIFFERENTIATE THE FOURIER SERIES TERMWISE. To do this, you would need to prove that the series can be differentiated termwise, which at this point we do not have the techniques to demonstrate. So, it will be an incomplete and incorrect proof. Not a good thing.

Instead, we will use the definition of Fourier coefficients and integration by parts:

$$c_n' = \frac{1}{2\pi}\int_{-\pi}^{\pi} f'(x)e^{-inx}dx = \frac{1}{2\pi} f(x)e^{-inx}\Big|_{x=-\pi}^{x=\pi} - \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(-ine^{-inx})dx$$

$$= \frac{in}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx = inc_n.$$

Above, we have used the fact that f is 2π periodic, and e^{-inx} is also 2π periodic so

$$\frac{1}{2\pi} f(x)e^{-inx}\Big|_{x=-\pi}^{x=\pi} = 0.$$

In the last step we use the definition of c_n. Recall that

$$a_n = c_n + c_{-n}, \quad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx, \quad \forall n \in \mathbb{N}_{\geq 1},$$

and

$$b_n = i(c_n - c_{-n}), \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx, \quad \forall n \in \mathbb{N}_{\geq 1},$$

with

$$a_0 = c_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx,$$

and the same relationship holds true for a_n', b_n', c_n'. We therefore compute

$$a_n' = c_n' + c_{-n}' = inc_n - inc_{-n} = in(c_n - c_{-n}) = nb_n,$$
$$b_n' = i(c_n' - c_{-n}') = i(inc_n + inc_{-n}) = -n(c_n + c_{-n}) = -na_n.$$

Now, using the theorem we have just proven, we obtain

Corollary 47. *Assume that f is 2π periodic, continuous, piecewise $\mathcal{C}^1$, and assume that f' is also piecewise $\mathcal{C}^1$. Then, if*

$$\sum_{-\infty}^{\infty} c_n e^{inx}$$

is the Fourier series for f, we have that

$$\sum_{n\in\mathbb{Z}} inc_n e^{inx}$$

is the Fourier series for f'.

Theorem 48. *Assume that f is 2π periodic, continuous, and piecewise $\mathcal{C}^1$. Then the Fourier series of f converges absolutely uniformly to f on all of $\mathbb{R}$.*

Proof: By assumption, f' is piecewise continuous. Bessel's inequality tells us that

$$\sum_{\mathbb{Z}} |c_n'|^2 < \infty.$$

We use the preceding theorem to say that for all $n \neq 0$,

$$|c_n| = \left| c_n' \frac{1}{n} \right|.$$

Hence we can estimate

$$\sum_{n\in\mathbb{Z}} |c_n e^{inx}| = \sum_{n\in\mathbb{Z}} |c_n| = |c_0| + \sum_{n\in\mathbb{Z}\setminus 0} \frac{|c_n'|}{|n|}.$$

By Bessel's inequality

$$\sum_{n\in\mathbb{Z}} |c_n'|^2 < \infty,$$

and we know very well that

$$\sum_{n\in\mathbb{Z}\setminus 0} |n|^{-2} < \infty.$$

So, using the Cauchy-Schwarz inequality on ℓ^2, we have

$$\sum_{n\in\mathbb{Z}} |c_n| = |c_0| + \sum_{n\in\mathbb{Z}\setminus 0} \frac{|c_n'|}{|n|} \leq |c_0| + \sqrt{\sum_{n\in\mathbb{Z}\setminus 0} |c_n'|^2} \sqrt{\sum_{n\in\mathbb{Z}\setminus 0} |n|^{-2}} < \infty.$$

Therefore the Fourier series converges absolutely, and uniformly for all $x \in \mathbb{R}$, because we see that the convergence estimates are independent of the point x. Since the function is continuous, the limit of the series is, by the Theorem on the pointwise convergence of Fourier series

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

We can repeat this idea to show that the more differentiable a function is, the faster its Fourier series converges.

Theorem 49. *Let f be 2π periodic, and assume that f is $\mathcal{C}^{k-1}$, and $f^{(k-1)}$ is piecewise $\mathcal{C}^1$, and f is piecewise $\mathcal{C}^k$. Then the Fourier coefficients of f satisfy*

$$\sum |n^k a_n|^2 < \infty, \quad \sum |n^k b_n|^2 < \infty, \quad \sum |n^k c_n|^2 < \infty.$$

If $|c_n| \leq c|n|^{-k-\alpha}$ for some $c > 0$ and $\alpha > 1$, for all $n \neq 0$, then $f \in \mathcal{C}^k$.

Proof: We apply the theorem relating the Fourier coefficients of f to those of the derivatives of f. We do it k times. We get

$$c_n^{(k)} = (in)^k c_n.$$

Next, we apply Bessel's inequality to conclude that since f is piecewise $\mathcal{C}^k$, $f^{(k)}$ is bounded on the interval hence it is in L^2 on the interval, and so

$$\sum |c_n^{(k)}|^2 < \infty.$$

Since

$$|c_n^{(k)}| = |n|^k |c_n|$$

this shows that

$$\sum |n^k c_n|^2 < \infty.$$

We have similar estimates for a_n and b_n using the same theorem, specifically

$$|a_n^{(k)}| = |n^k a_n|, \quad |b_n^{(k)}| = |n^k b_n|.$$

Hence,

$$\sum |n^k a_n| < \infty, \quad \sum |n^k b_n| < \infty.$$

Now we demonstrate the result which says that if the Fourier coefficients are sufficiently rapidly decaying, then the function f is actually in $\mathcal{C}^k$. Let

$$g(x) := f^{(k-1)}(x).$$

Then g is continuous and by assumption it is piecewise $\mathcal{C}^1$. Therefore, by the theorem on the pointwise convergence of Fourier series, the Fourier series of g converges to $g(x)$ for all x in $\mathbb{R}$. Next, we use the assumption and the fact that the Fourier coefficients of g are

$$c_n^{(k-1)} = (in)^{k-1} c_n.$$

Therefore

$$\sum_{n \in \mathbb{Z}} |c_n^{(k-1)} e^{inx}| = \left| c_0^{(k-1)} \right| + \sum_{n \neq 0} |n^{k-1}||c_n| \le \left| c_0^{(k-1)} \right| + c \sum_{n \neq 0} |n|^{k-1-k-\alpha} < \infty.$$

Hence, the series converges absolutely and uniformly in $\mathbb{R}$. Moreover, differentiating the series termwise is legitimate, because the result

$$\sum_{n \in \mathbb{Z}} inc_n^{(k-1)} e^{inx}$$

also converges absolutely and uniformly in $\mathbb{R}$:

$$\sum_{n \in \mathbb{Z}} |inc_n^{(k-1)}| \le \sum_{n \neq 0} |n||c_n^{(k-1)}| \le c \sum_{n \neq 0} |n||n|^{k-1-k-\alpha} < \infty$$

because $\alpha > 1$. Since the series is equal to $g(x) = f^{(k-1)}(x)$ for all $x \in \mathbb{R}$, and the series is a differentiable function for all $x \in \mathbb{R}$, this shows that g is differentiable for all $x \in \mathbb{R}$. Moreover, g' is continuous on $\mathbb{R}$, because the series defines a continuous function.[1] This is the case because the series defining g' converges absolutely and uniformly for all of $\mathbb{R}$. Hence, $f^{(k-1)}$ is in $\mathcal{C}^1$ on all of $\mathbb{R}$, and therefore f is in $\mathcal{C}^k$ on all of $\mathbb{R}$.

We will prove a theorem about integrating Fourier series. To get warmed up, here is an exercise.

Exercise 50. *Show that if you compute the indefinite integrate*

$$\int e^{inx} dx, \quad n \in \mathbb{Z} \setminus \{0\},$$

the result is also a 2π periodic function. What happens in the case $n = 0$?

[1] This is true because the series should really be viewed as the limit of the partial series, and each partial series defines a smooth, thus also continuous, function. The uniform limit of continuous functions is itself a continuous function.

Theorem 51. *Let f be a 2π periodic function which is piecewise continuous. Define*

$$F(x) := \int_0^x f(t)dt.$$

If $c_0 = 0$, then

$$F(x) = C_0 + \sum_{n \neq 0} \frac{c_n}{in} e^{inx}, \quad C_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x)dx.$$

Similarly,

$$F(x) = \frac{1}{2} A_0 + \sum_{n \geq 1} \frac{a_n}{n} \sin(nx) - \frac{b_n}{n} \cos(nx).$$

Proof: We first note that F is continuous and piecewise $\mathcal{C}^1$, because it is the integral of a piecewise continuous function. Moreover, assuming $c_0 = 0$, we see that

$$F(x+2\pi) - F(x) = \int_0^{x+2\pi} f(t)dt - \int_0^x f(t)dt = \int_x^{x+2\pi} f(t)dt = \int_{-\pi}^{\pi} f(t)dt = 2\pi c_0 = 0.$$

Above we have used the nifty lemma that allows us to slide around integrals of periodic functions. So, F satisfies the assumptions of the theorem on pointwise convergence of Fourier series. We therefore have pointwise convergence of the Fourier series of F. Moreover, applying the theorem relating the Fourier coefficients of $F' = f$ to those of F, we have

$$C_n = \frac{c_n}{in} \quad n \neq 0.$$

The reason is because $c_n = C_n'$, and the theorem says $C_n' = inC_n$ which shows $c_n = inC_n$, which we can re-arrange as above. The formula for C_0 is just the usual formula for it, because we can't say anything more specific without knowing more information on f. The re-statement in terms of a and b follows from the relationship between these and the c_n Fourier coeffiients.

Remark 2. If $c_0 \neq 0$, then define a new function

$$g(t) := f(t) - c_0.$$

Since f is 2π periodic, so is g. Then, apply the theorem above to g. Note that

$$G(x) = \int_0^x g(t)dt = F(x) - c_0 x.$$

Moreover, the Fourier coefficients of g,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x) - c_0) e^{-inx} dx = c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad \forall n \neq 0.$$

So, the series for $G(x)$ from the theorem is

$$\widetilde{C_0} + \sum_{n \neq 0} \frac{c_n}{in} e^{inx},$$

with

$$\widetilde{C_0} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (F(x) - c_0 x)\, dx = C_0.$$

So, in fact, it is the same C_0, where we have used the oddness of the function x above. Then, we get something of a corollary which says that in general, the series in the theorem,

$$C_0 + \sum_{n\neq 0} \frac{c_n}{in} e^{inx}, \quad C_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)dx$$

converges to $F(x) - c_0 x$.

4.5.1 Computing a sum using the Integration Theorem for Fourier Series

In general, this is a more complicated way to compute series, so if there is an alternative method, like Parseval's equality, that would most likely be easier. However, this is a useful technique, and so we proceed with an example, by computing

$$\sum_{n\geq 1} \frac{1}{n^4}.$$

Since we used $f(x) = x$ to compute the series $\sum n^{-2}$ it is a reasonable guess that $f(x) = x^2$ could be used to compute this series. To practice a different, but equivalent technique, we shall use the orthogonal basis $\{1, \cos(nx), \sin(nx)\}_{n\in\mathbb{N}}$ to expand this function on $(-\pi, \pi)$. Since x^2 is even, and $\sin(nx)$ is odd for all $n \geq 1$, the integrals

$$\int_{-\pi}^{\pi} x^2 \sin(nx)dx = 0 \quad \forall n \geq 1.$$

We compute the cosine terms,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2 \cos(nx)dx = \frac{2}{\pi}\int_0^{\pi} x^2 \cos(nx)dx.$$

We do this integral:

$$\int_0^{\pi} x^2 \cos(nx)dx = \int x^2 \left(\frac{\sin(nx)}{n}\right)' dx = x^2 \frac{\sin(nx)}{n}\Big|_0^{\pi} - \int_0^{\pi} 2x\frac{\sin(nx)}{n} dx.$$

Above we did integration by parts. The first part vanishes. The second term we handle with integration by parts again,

$$\int_0^{\pi} x\sin(nx)dx = \int_0^{\pi} x\left(-\cos(nx)/n\right)' dx = -\frac{x\cos(nx)}{n}\Big|_0^{\pi} + \int_0^{\pi} \cos(nx)/n dx.$$

Now this time the second term vanishes because integrating gives us a sine which is 0 at 0 and at π. So, recalling the constant factors, we get

$$\int_0^{\pi} x^2 \cos(nx)dx = \frac{2\pi\cos(\pi n)}{n^2} = \frac{2\pi(-1)^n}{n^2}.$$

Hence our coefficients,

$$a_n = \frac{2(2)(-1)^n}{n^2}.$$

Moreover, we also compute the term

$$a_0 = \frac{1}{\pi}\int\limits_{-\pi}^{\pi} x^2 dx = \frac{2\pi^3}{3\pi} = \frac{2\pi^2}{3}.$$

Hence, the Fourier series expansion of x^2 is

$$\frac{\pi^2}{3} + 4\sum_{n\geq 1}\frac{(-1)^n\cos(nx)}{n^2}.$$

Let $x=\pi$. Since our periodically extended function, x^2 is continuous on all of $\mathbb{R}$, the Fourier series converges to its value at $x=\pi$ which means

$$\pi^2 = \frac{\pi^2}{3} + 4\sum_{n\geq 1}\frac{(-1)^n(-1)^n}{n^2} \implies \frac{\pi^2}{6} = \sum_{n\geq 1}\frac{1}{n^2}.$$

To get up to summing n^{-4} we will use Theorem 51. We see that

$$c_0 = \frac{\pi^2}{3}.$$

We also see that for $f(t)=t^2$,

$$F(x) := \int\limits_0^x f(t)dt = \frac{x^3}{3}.$$

The series from the theorem is

$$C_0 + 4\sum_{n\geq 1}\frac{(-1)^n\sin(nx)}{n^3}.$$

The term

$$C_0 = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} F(x)dx = 0,$$

because $F(x)$ above is odd. Hence, the theorem together with the remark after it says that

$$4\sum_{n\geq 1}\frac{(-1)^n\sin(nx)}{n^3} = \frac{x^3}{3} - \frac{\pi^2 x}{3}, \quad x\in[-\pi,\pi].$$

To proceed, we're going to need to use the theorem once more to get n^4 in the denominator. Before we do this, let's multiply everything by 3 to make it nicer. Then

$$x^3 - \pi^2 x = 12\sum_{n\geq 1}\frac{(-1)^n\sin(nx)}{n^3}, \quad x\in[-\pi,\pi].$$

So, here we have

$$f(t) = t^3 - \pi^2 t \implies F(x) = \int\limits_0^x f(t)dt = \frac{x^4}{4} - \frac{\pi^2 x^2}{2}.$$

We see also that

$$c_0 = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} f(t)dt = 0.$$

Hence, we apply the theorem directly to F. The theorem says

$$F(x) = C_0 + 12 \sum_{n \geq 1} -\frac{(-1)^n \cos(nx)}{n^4}.$$

We compute

$$C_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x)dx = \frac{1}{\pi} \int_0^{\pi} \frac{x^4}{4} - \frac{\pi^2 x^2}{2} dx = \frac{\pi^4}{20} - \frac{\pi^4}{6}.$$

Therefore

$$F(x) = \frac{x^4}{4} - \frac{\pi^2 x^2}{2} = \frac{\pi^4}{20} - \frac{\pi^4}{6} - 12 \sum_{n \geq 1} \frac{(-1)^n \cos(nx)}{n^4}, \quad x \in [-\pi, \pi].$$

We do the same trick now of choosing

$$x = \pi \implies \cos(nx) = \cos(n\pi) = (-1)^n, \quad (-1)^n(-1)^n = 1 \forall n.$$

Hence,

$$F(\pi) = \frac{\pi^4}{4} - \frac{\pi^4}{2} = \frac{\pi^4}{20} - \frac{\pi^4}{6} - 12 \sum_{n \geq 1} \frac{1}{n^4}.$$

Re-arranging things

$$\sum_{n \geq 1} \frac{1}{n^4} = \frac{1}{12} \left(\frac{\pi^4}{20} - \frac{\pi^4}{6} + \frac{\pi^4}{2} - \frac{\pi^4}{4} \right).$$

Just for fun, we determine what this is...

$$\frac{\pi^4}{20} - \frac{\pi^4}{6} + \frac{\pi^4}{2} - \frac{\pi^4}{4} = \frac{\pi^4}{2} \left(\frac{1}{10} - \frac{1}{3} + \frac{1}{2} \right) = \frac{\pi^4}{2} \left(\frac{3 - 10 + 15}{30} \right)$$

$$= \frac{\pi^4}{2} \left(\frac{8}{30} \right) = \frac{2\pi^4}{15}.$$

So, recalling the factor of $\frac{1}{12}$, we see that

$$\sum_{n \geq 1} \frac{1}{n^4} = \frac{2\pi^4}{(12)(15)} = \frac{\pi^4}{6(15)} = \frac{\pi^4}{90}.$$

We have again earned ourselves some pie. To honor the esteemed mathematical physicist, Stephen Hawking, this time we shall a mincemeat pie as in Figure 4.11.

4.6 Fourier sine and cosine series: even and odd extensions

In the preceding example, we were able to express x^2 as a Fourier series that only required cosines. We may wish to be able to do that in general, for functions that from the start might *not* be even or odd. One way to do this is to start with our function on an interval, and then extend it to twice that interval evenly or oddly. To show how this is done, we consider the interval $(0, \pi)$. Assume that we have some arbitrary $\mathcal{L}^2$ function $f(x)$ defined on this interval. First, we define its odd extension:

$$f_o(x) := \begin{cases} f(x) & x \in [0, \pi] \\ -f(-x) & x \in (-\pi, 0). \end{cases}$$

Figure 4.11: We have managed to compute yet another seemingly impossible sum, and again we have seemingly magically appearing πs. To celebrate, here we have a piece of mincemeat pie. According to my English friends, this is not a pie involving meat, but rather a sweet pie that is filled with a mixture of dried fruits and spices, traditionally served during the Christmas season. Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

In the exercises of this chapter, we will prove that for the odd extension f_o, the a_n Fourier coefficients coefficients are all zero, and the b_n coefficients are

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx = \frac{2}{\pi}\int_0^{\pi} f(x)\sin(nx)dx.$$

Here we used the fact that sine is also an odd function, and therefore the product of f_o and $\sin(nx)$ is an *even* function for all n. In this way we can obtain a Fourier series for f that contains only sine terms! This is pretty awesome, because we do not need f to be an odd function.

On the other hand, if we wish to obtain a Fourier cosine series for f, we extend it evenly, defining

$$f_e(x) := \begin{cases} f(x) & x \in [0,\pi] \\ f(-x) & x \in (-\pi, 0). \end{cases}$$

In the exercises of this chapter, we will prove that for the even extension f_e, the b_n Fourier coefficients all vanish, and the a_n coefficients are

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx = \frac{2}{\pi}\int_0^{\pi} f(x)\cos(nx)dx, \quad n \geq 0.$$

Above we used the fact that cosine is even, and so its product with f_e is also even. In this way, we may define Fourier sine and cosine series for functions defined on $[0,\pi]$. The Fourier sine series is defined to be

$$\sum_{n\geq 1} b_n \sin(nx), \quad b_n = \frac{2}{\pi}\int_0^{\pi} f(x)\sin(nx)dx$$

whereas the Fourier cosine series is

$$c_0 + \sum_{n\geq 1} a_n\cos(nx), \quad c_0 = \frac{1}{\pi}\int_0^{\pi} f(x)dx, \quad a_n = \frac{2}{\pi}\int_0^{\pi} f(x)\cos(nx)dx, \quad n \geq 1.$$

More generally, we can do the same procedure on any interval of the form $[0, \ell]$ to obtain sine and cosine series.

Theorem 52. *Let f be a function which is piecewise $\mathcal{C}^1$ on $[0, \pi]$. Then the Fourier sine and cosine series converge to $f(x)$ for all $x \in (0, \pi)$ at which f is continuous. For other points, they converge to*

$$\frac{1}{2}\left(f(x_-) + f(x_+)\right).$$

More generally, for a function that is piecewise $\mathcal{C}^1$ on $[0, \ell]$, the Fourier sine and cosine series are defined to be

$$\sum_{n\geq 1} b_n \sin(n\pi x/\ell), \quad b_n = \frac{2}{\ell}\int_0^\ell f(x)\sin(n\pi x/\ell)dx,$$

and

$$c_0 + \sum_{n\geq 1} a_n \cos(n\pi x/\ell), \quad c_0 = \frac{1}{\ell}\int_0^\ell f(x)dx, \quad a_n = \frac{2}{\ell}\int_0^\ell f(x)\cos(n\pi x/\ell)dx, \quad n \geq 1.$$

These series converge to $f(x)$ at all points $x \in (0, \ell)$ at which f is continuous, and for other points in the same interval they converge to the average of the left and right sided limits.

Proof: First, we extend the function either evenly or oddly. Next, we extend it to all of $\mathbb{R}$ to be 2π periodic. Like Riker in Star Trek the Next Generation, we just *make it so.* We're only proving a statement about points in $(0, \pi)$. So, what happens outside of this set of points, well it don't matter. We apply the theorem on pointwise convergence of Fourier series now. In the case where the interval is $(0, \ell)$, we note that the function

$$g(t) := f(x\pi/\ell), \quad t = \frac{x\pi}{\ell} \implies x \in (0, \ell) \iff t \in (0, \pi)$$

is defined on $(0, \pi)$ for f defined on $(0, \ell)$. We therefore therefore apply the theorem from the interval $(0, \pi)$ obtain its series for g, and then use a change of variables to obtain that for f.

4.7 Fourier trigonometric series on an arbitrary interval

We can obtain a Fourier series for a function f defined on any interval. It is convenient to write the interval as $[a - \ell, a + \ell]$ for some $a \in \mathbb{R}$, and some $\ell > 0$. To do this we define

$$g(t) := f\left(\frac{t\ell}{\pi} + a\right) = f(x).$$

Then when $t \in [-\pi, \pi]$, the corresponding $x \in [a - \ell, a + \ell]$, that is

$$\frac{t\ell}{\pi} + a = x, \quad t = \frac{(x-a)\pi}{\ell}.$$

So, we use the standard procedure to compute expand g into a Fourier series of the form

$$\sum_{n\in\mathbb{Z}} c_n e^{int},$$

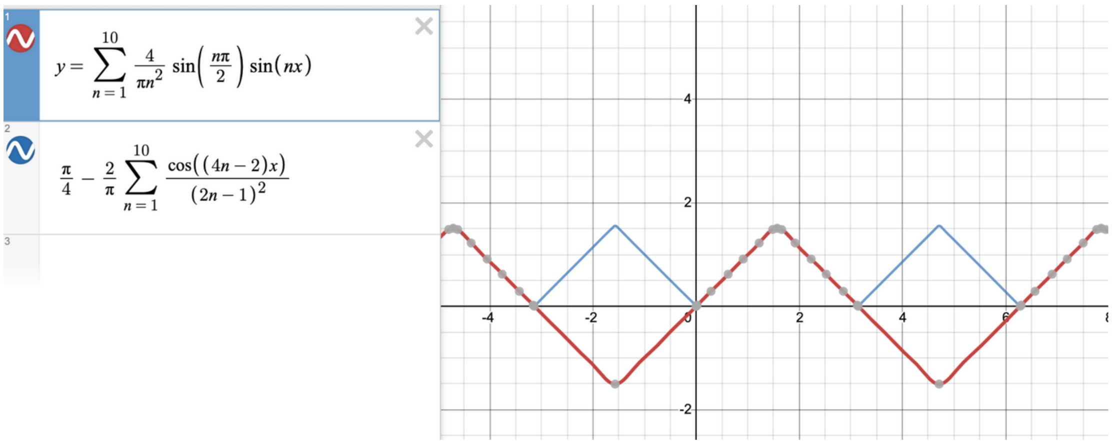


Figure 4.12: The same function is expanded in a sine series and a cosine series on the interval $(0, \pi)$. The function is equal to x from 0 to $\frac{\pi}{2}$ and $\pi - x$ from $\frac{\pi}{2}$ to π. Both the sine and cosine series converge to the function on this interval, but then outside the interval, the sine series is odd whereas the cosine series is even. This figure was created by Edvin Martinson; thank you!

with coefficients

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) e^{-int} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(\frac{t\ell}{\pi} + a\right) e^{-int} dt.$$

Substituting in the integral,

$$x = \frac{t\ell}{\pi} + a, \quad dx = \frac{\ell dt}{\pi}$$

the coefficients become:

$$c_n = \frac{1}{2\pi} \frac{\pi}{\ell} \int_{a-\ell}^{a+\ell} f(x) e^{-in(x-a)\pi/\ell} dx = \frac{1}{2\ell} \int_{a-\ell}^{a+\ell} f(x) e^{-in(x-a)\pi/\ell} dx.$$

Then, we get by substituting for t in terms of x the Fourier series for f,

$$\sum_{n \in \mathbb{Z}} c_n e^{in\left(\frac{(x-a)\pi}{\ell}\right)}.$$

A similar relationship holds for the Fourier cosine and sine coefficients:

$$a_n = \frac{1}{\ell} \int_{a-\ell}^{a+\ell} f(x) \cos(n(x-a)\pi/\ell) dx, \quad b_n = \frac{1}{\ell} \int_{a-\ell}^{a+\ell} f(x) \sin(n(x-a)\pi/\ell) dx,$$

and the Fourier series has the form

$$c_0 + \sum_{n \geq 1} a_n \cos(n(x-a)\pi/\ell) + b_n \sin(n(x-a)\pi/\ell).$$

This series will converge to f on the interval $[a-\ell, a+\ell]$, and to the 2ℓ periodic extension of f on the rest of $\mathbb{R}$.

Theorem 53. *Assume that f is defined on an interval $[a-\ell, a+\ell]$ for some $a \in \mathbb{R}$, and some $\ell > 0$, such that f is piecewise $\mathcal{C}^1$ on this interval. Then the Fourier series for f, defined by*

$$\sum_{n\in\mathbb{Z}} c_n e^{in\left(\frac{(x-a)\pi}{\ell}\right)}, \quad c_n = \frac{1}{2\ell}\int_{a-\ell}^{a+\ell} f(x) e^{-in(x-a)\pi/\ell} dx,$$

or equivalently the series

$$c_0 + \sum_{n\geq 1} a_n \cos(n(x-a)\pi/\ell) + b_n \sin(n(x-a)\pi/\ell)$$

converges to $f(x)$ for all $x \in (a-\ell, a+\ell)$ at which f is continuous. At a point $x \in (a-\ell, a+\ell)$ where f is not continuous, the series converges to

$$\frac{f(x_+) + f(x_-)}{2}. \tag{4.7.1}$$

Exercise 54. *Prove the theorem. As a hint: apply the Theorem on Pointwise Convergence of Fourier Series to the function g above.*

4.8 Fourier series in mathematical physics

The historical origin of Fourier series is their application to solving the heat equation [5]. We will continue to use Fourier series to solve partial differential equations in the following chapter. Another application of Fourier series is to compute seemingly incomputable series. For example, we can compute a series of the form

$$\sum_{n\geq 1} a(n)$$

as long as we can obtain an estimate for the terms of the form

$$|a(n)| \leq C\frac{1}{n^2}, \quad \forall n \geq N, \quad \text{for some constant } N.$$

The way this works is that we can exactly compute

$$\sum_{n=N}^{\infty} \frac{C}{n^2} = \frac{C\pi^2}{6} - C\sum_{n=1}^{N-1}\frac{1}{n^2}.$$

Then, taking N sufficiently large, we can make the right side above as small as we wish. Consequently, since the tail of the series satisfies

$$\left|\sum_{n\geq N} a(n)\right| \leq \sum_{n=N}^{\infty}\frac{C}{n^2} = \frac{C\pi^2}{6} - C\sum_{n=1}^{N-1}\frac{1}{n^2},$$

we can calculate the value of the series by summing $a(n)$ from 1 up to N, and then estimating the tail as above. We can make this tail as small as we want, and therefore compute the series $\sum_{n\geq 1} a(n)$ to any level of accuracy we wish!

4.8.1 Zeta functions in mathematical physics and number theory

The series that we use above has further connections to physics and number theory via a mathematical function known as a zeta function. Riemann's zeta function is

$$\zeta(s) = \sum_{n\geq 1}\frac{1}{n^s}.$$

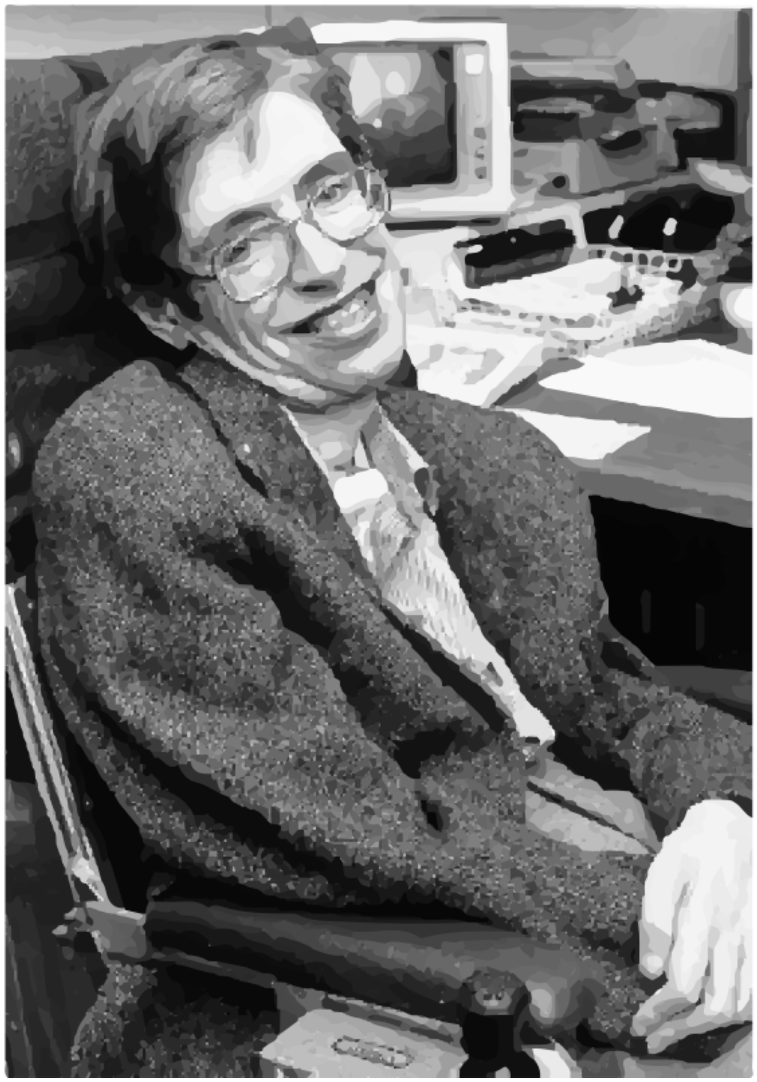


Figure 4.13: Stephen Hawking lived from 1942 until 2018. He was a theoretical physicist and worked as both a professor of physics as well as a professor of mathematics. In spite of significant physical challenges, including the loss of speech, he made tremendous contributions to mathematical physics in particular, science in general, and society at large. I really enjoyed his book ‘A brief history of time.’ Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

This will be investigated in the exercises. This is one particular example of a more general *spectral zeta function* that appears in mathematical physics. On the mathematical side, we study spectral zeta functions that are defined by the eigenvalues of a differential operator. In case zero is an eigenvalue, we exclude it from the sum. As an example, let's consider the vibrating string with fixed ends. Part of our solution involved finding all solutions X to:

$$X''(x) = \lambda X(x), \quad X(0) = X(\ell) = 0, \quad \text{for some constant } \lambda \in \mathbb{R}.$$

We found that the only possible values of the constant λ are $\lambda_n = -\frac{n^2\pi^2}{\ell^2}$ for $n \in \mathbb{N}$, and the corresponding function $X_n(x)$ are constant multiples of $\sin(n\pi x/\ell)$. The physical interpretation of the values of $|\lambda_n|$ is that they are frequencies, but frequencies are positive quantities. So, the spectral zeta function corresponding to this problem is therefore

$$\zeta_{[0,\ell]}(s) = \sum_{n\geq 1} \frac{\ell^{2s}}{(n\pi)^{2s}} = \frac{\ell^{2s}}{\pi^{2s}} \sum_{n\geq 1} \frac{1}{n^{2s}}.$$

Here we used $\zeta_{[0,\ell]}$ to indicate that the zeta function is for the interval $[0, \ell]$. For the sake of clarity, let us write the Riemann zeta function as

$$\zeta_R(s) = \sum_{n\geq 1} n^{-s},$$

then we have

$$\zeta_{[0,\ell]}(s) = \frac{\ell^{2s}}{\pi^{2s}} \zeta_R(2s).$$

This zeta function will be further explored in the exercises. In particular, although it is only apparently well defined for values of s such that the series converges, there is a way to meromorphically extend it to the entire complex plane. Furthermore, we will show that this extension is holomorphic in a neighborhood of the point $s = 0$. For a moment, let us throw mathematical rigour to the wind, and compute

$$\left.\frac{d}{ds}\zeta_R(s)\right|_{s=0} = \sum_{n\geq 1} -\log(n)n^0 = -\sum_{n\geq 1} \log(n),$$

and so

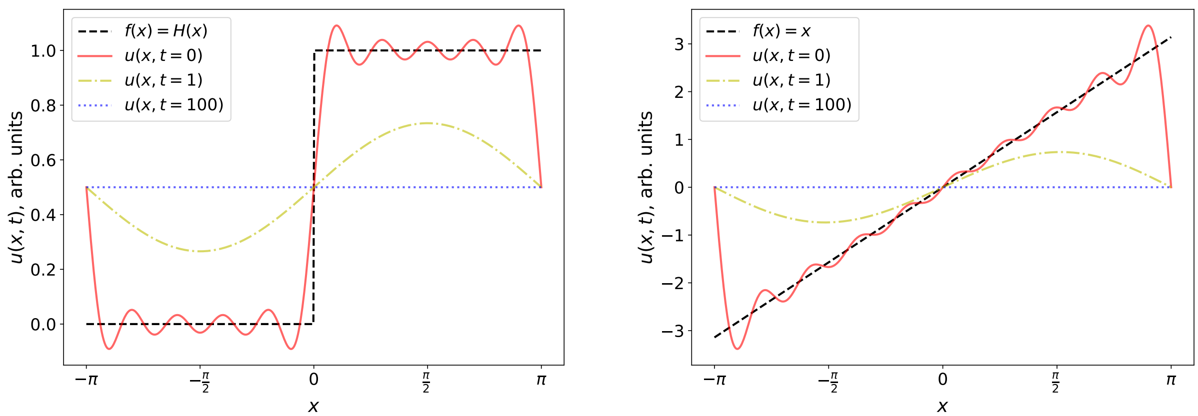
$$e^{-\zeta_R'(0)} \text{``}=\text{''}\, e^{\sum_{n\geq 1}\log(n)} = \prod_{n\geq 1} n.$$

Well, the right side is pure nonsense, so why are we doing this? Our differential operator provides us with an orthogonal basis for the infinite dimensional Hilbert space $\mathcal{L}^2(0, \ell)$. Note that it consists of sines, and this fits with our ability to expand functions in Fourier sine series, if we would like to do that! If we view the differential operator acting on the Hilbert space as an infinite dimensional matrix, then this orthogonal basis of sines diagonalizes the operator. The determinant of the operator should be equal to the product of all of its eigenvalues. Now, this product cannot be defined, because it is infinite. However, in the exercises we will prove that $\zeta_R'(0)$ *is* well-defined, and so it allows us to define the determinant of the operator as $\exp(-\zeta_{[0,\ell]}'(0))$. This procedure is known as *zeta function regularization.* According to Stephen Hawking, shown in Figure 4.13, this method allows one to obtain "an energy momentum tensor which is finite even on the horizon of a black hole" [8]. In the language of physics, the zeta regularization is a technique for obtaining finite values to path integrals for fields (including the gravitational field) on a curved spacetime background. In the language of mathematics, this is a technique for evaluating the determinants of differential operators. As I like to say 'physics is just mathematics in cooler words.'

It may be interesting to note that in [1], we proved that the zeta-regularized determinant, defined in this way, for the analogous problem on a rectangle of side lengths L and $1/L$ (thus normalized to have area equal to one) is uniquely minimized by the square, and it tends to zero if the rectangle collapses, that is if $L \to 0$. To prove this result, we used classical number theoretic equalities of Hardy and Ramanujan, as well as modern analytic number theory estimates for the partition function.

4.8.2 The Gibbs phenomenon

Assume that a function f is discontinuous at a point $x_0 \in (-\pi, \pi)$. Then, the Fourier series for f on any interval containing x_0 does *not* converge uniformly. The reason is that the Fourier series is comprised of continuous functions, and the uniform limit of continuous functions is continuous. Since the limit function, f, is *not* continuous, the convergence therefore cannot be uniform. If we compute the partial Fourier series of such a function, as shown in Figures 4.14a and 4.14b we will see high frequency oscillations; these oscillations demonstrate the lack of uniformity of the convergence. This is known as the Gibbs phenomenon.



(a) Initial data equal to the Heavyside function.

(b) Initial data equal to x.

Figure 4.14: The heat equation on a 'hot rod' with initial data that has discontinuities and therefore we observe the non-uniform convergence of the Fourier series to the initial data, known as the Gibbs phenomenon.

Figures 4.14a and 4.14b show the numerical approximation of the solution of the heat equation on a circular rod with initial data equal to the Heavyside function, and the function x, respectively. These figures were generously created by Eric Lindgren using the python code contained in §B.2. The Heavyside function is equal to one for $x > 0$ and equal to zero for $x < 0$. The heavyside function is discontinuous at zero, and when we extend it as a periodic function, it has discontinuities at all integer multiples of π. So, at all of these points, the Fourier series converges to the average of the left and right sided limits. Moreover, the series oscillates wildly and exhibits the Gibbs phenomenon; it does not converge uniformly to the initial data. Similarly, in Figure 4.14b, the initial data x, extended to a 2π periodic function has discontinuities at odd integer multiples of π. So, at all of these points, the Fourier series converges to the average of the left and right sided limits. Due to these discontinuities, we also see the Gibbs phenomenon occurring here; the Fourier series does not converge uniformly to the initial data.

4.9 Exercises

1. Compute the Fourier series of the function defined on $(-\pi, \pi)$

$$f(x) := \begin{cases} 0 & -\pi < x < 0 \\ \cos(x) & 0 < x < \pi \end{cases}.$$

2. Compute the Fourier series of the function defined on $(-\pi, \pi)$

$$f(x) := \cosh(x).$$

3. Compute the Fourier series of the function defined on $(-\pi, \pi)$

$$f(x) := \begin{cases} 1 & -a < x < a \\ -1 & 2a < x < 4a \\ 0 & \text{elsewhere in } (-\pi, \pi). \end{cases}.$$

Here one ought to assume that $0 < a < \pi$ for this to make sense.

4. Compute the Fourier series of the function defined on $(-\pi, \pi)$

$$f(x) = \cos^2(x).$$

5. Use the Fourier series of the function $f(x) = x(\pi - |x|)$, defined on $(-\pi, \pi)$ and extended to be 2π periodic on $\mathbb{R}$, to compute the sums:

$$\sum_{n\geq 1} \frac{1}{n^2} = \frac{\pi^2}{6}, \quad \sum_{n\geq 1} \frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{12}.$$

6. Use the Fourier series of the function $f(x) = e^{bx}$, defined on $(-\pi, \pi)$ and extended to be 2π periodic on $\mathbb{R}$, to compute the sum:

$$\sum_{n\geq 1} \frac{1}{n^2 + b^2} = \frac{\pi}{2b} \coth(b\pi) - \frac{1}{2b^2}.$$

7. [4, 2.5.4] Consider a vibrating string of length ℓ. Suppose that the string is plucked in the middle such that its initial displacement is

$$u(x, 0) = \begin{cases} \frac{2mx}{\ell} & 0 \leq x \leq \frac{\ell}{2} \\ \frac{2m(\ell - x)}{\ell} & \frac{\ell}{2} \leq x \leq \ell, \end{cases}$$

for some $m > 0$. Assume that its initial velocity is zero. Find the displacement $u(x, t)$ for subsequent times.

8. [4, 2.5.5] Suppose now that the initial displacement is $\frac{mx}{a}$ for $0 \leq x \leq a$ and $\frac{m(\ell - x)}{\ell - a}$ for $a \leq x \leq \ell$ for some $a \in (0, \ell)$. Find the displacement $u(x, t)$ assuming the initial velocity is zero.

9. Use the Fourier series for the function $f(x) = |\sin(x)|$ to compute the sum

$$\sum_{n=1}^{\infty} \frac{1}{4n^2 - 1} = \frac{1}{2}, \quad \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{4n^2 - 1} = \frac{\pi - 2}{4}.$$

10. Use the Fourier series for the function $f(x) = x(\pi - |x|)$ to compute the sum

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n - 1)^3} = \frac{\pi^3}{32}.$$

11. [4, 2.3.5] Let $f(x)$ be the periodic function such that $f(x) = e^x$ for $x \in (-\pi, \pi)$, and extended to be 2π periodic on the rest of $\mathbb{R}$. Let

$$\sum_{n\in\mathbb{Z}} c_n e^{inx}$$

be its Fourier series. Therefore, by the Theorem on Pointwise Convergence of Fourier Series

$$e^x = \sum_{n\in\mathbb{Z}} c_n e^{inx}, \quad \forall x \in (-\pi, \pi).$$

If we differentiate this series term-wise then we get $\sum inc_n e^{inx}$. On the other hand, we know that $(e^x)' = e^x$. So, then we should have

$$\sum inc_n e^{inx} = \sum c_n e^{inx} \implies c_n = inc_n \quad \forall n.$$

This is clearly wrong. Where is the mistake?

12. Determine the Fourier sine and cosine series of the function

$$f(x) = \begin{cases} x & 0 \leq x \leq \frac{\pi}{2} \\ \pi - x & \frac{\pi}{2} \leq x \leq \pi \end{cases}$$

13. Expand the function

$$f(x) = \begin{cases} 1 & 0 < x < 2 \\ -1 & 2 < x < 4 \end{cases}$$

in a cosine series on $[0, 4]$.

14. Expand the function e^x in a series of the form

$$\sum_{n \in \mathbb{Z}} c_n e^{2\pi inx}, \quad x \in (0, 1).$$

Hint: the functions $\{e^{2\pi inx}\}_{n \in \mathbb{Z}}$ are an orthogonal basis for $(0, 1)$. Use the standard formula for c_n on this interval: $c_n = \frac{\langle e^x, e^{2\pi inx} \rangle}{||e^{2\pi inx}||^2}$.

15. Define

$$f(t) = \begin{cases} t & 0 \leq t \leq 1 \\ 1 & 1 < t < 2 \\ 3 - t & 2 \leq t \leq 3 \end{cases}$$

and extend f to be 3-periodic on $\mathbb{R}$. Expand f in a Fourier series. Determine, in the form of a Fourier series, a 3-periodic solution to the equation

$$y''(t) + 3y(t) = f(t).$$

16. [4, 2.5.1] A rod 100 cm long is insulated along its length and at both ends. Assume that its initial temperature is $u(x, 0) = x$, such that x is measured in cm, u in centigrade, t in seconds, and $0 \leq x \leq 100$. Assume that the diffusivity coefficient $k = 1.1$ (this is about right if the rod is made of copper!). Find the temperature $u(x, t)$ for $t > 0$. What happens to $u(x, t)$ as $t \to \infty$?

17. Consider a rod also 100 cm long with $k = 0.1$ (this is about right if the rod is ceramic). Now the rod is initially at temperature 100 degrees and the ends are subsequently put into an ice bath at 0 degrees. Assume that there is no heat loss along the length of the rod. Find the temperature $u(x, t)$ at subsequent times. What happens to $u(x, t)$ as $t \to \infty$?

18. [4, 2.5.6] Consider a vibrating string of length ℓ. The string is struck in the middle, so that its initial displacement is zero *but* its initial velocity is $u_t(x, 0) = 1$ for $|x - \ell/2| < \delta$ and 0 elsewhere, for some $\delta > 0$. Find $u(x, t)$ for $t > 0$.

19. [4, 2.5.7] Suppose that the temperature at time t at a point on the surface of the earth is given by

$$u(0, t) = 10 - 7\cos(2\pi t) - 5\cos(2\pi 365t).$$

Here u is measured in centigrade, t is in years, and the coefficients are roughly correct for Seattle, Washington. Suppose that the diffusivity coefficient of the earth $k = 9.46m^2$ per year. Find $u(x, t)$ for $t > 0$. At what depth x do the daily variations in the temperature become less than one unit? What about the annual variation?

20. [4, 2.3.7] How many derivatives can you guarantee that the following functions have?

   (a) $\sum_{n\in\mathbb{Z}} \frac{e^{inx}}{n^{13.2}+2n^6-1}$

   (b) $\sum_{n\geq 0} \frac{\cos(nx)}{2^n}$

   (c) $\sum_{n\geq 0} \frac{\cos(2^n x)}{2^n}$

21. This exercise is about the Riemann zeta function and the zeta regularization process. To quote Stephen Hawking,

   > The zeta function can be applied to calculate the partition functions for thermal gravitons and matter quanta on black holes.

   Moreover, the Riemann zeta function is the central object in one of the most famous long-standing open problems in number theory in particular and mathematics in general: the Riemann hypothesis.

   (a) The Riemann zeta function is defined for $s \in \mathbb{C}$ to be:

   $$\zeta(s) := \sum_{n\geq 1} \frac{1}{n^s}.$$

   For which values of $s \in \mathbb{C}$ does this series converge? Explain how you can use Fourier series to compute $\zeta(s)$ when s is a positive even integer.

   (b) One of the most special functions in mathematics is the Gamma function, written using the capitalized Greek letter Γ,

   $$\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}dt.$$

   Compute $\Gamma(1)$. Show that the integral above can be defined for all $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 0$. Show that Γ satisfies the functional equation

   $$\Gamma(s+1) = s\Gamma(s).$$

   In this way, show that

   $$\lim_{s\to 0} \Gamma(s) = \lim_{s\to 0} \frac{\Gamma(s+1)}{s},$$

   and therefore Γ has a simple pole at $s = 0$. Use the functional equation to extend Γ meromorphically to $\mathbb{C}$ with simple poles at $0, -1, -2, \ldots$.

   (c) Show that the Riemann zeta function satisfies

   $$\zeta_R(s) = \sum_{n\geq 1} n^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \sum_{n\geq 1} e^{-nt}dt,$$

   for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$.

   (d) The series

   $$\sum_{n\geq 1} e^{-nt} = \sum_{n\geq 1} (e^{-t})^n$$

   is a convergent geometric series for any fixed $t > 0$. Show that it converges to

   $$\frac{e^{-t}}{1-e^{-t}} = \frac{1}{e^t - 1}, \quad t > 0.$$

Consequently, the Riemann zeta function satisfies

$$\zeta_R(s) = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \frac{1}{e^t - 1} dt.$$

The function $\frac{1}{e^t-1}$ has a simple pole at $t = 0$. Split the integral above into $\int_0^1 + \int_1^\infty$ and use the Laurent series expansion of $\frac{1}{e^t-1}$ together with the fact that $\Gamma(s)$ has a simple pole at $s = 0$ to show that we can meromorphically extend $\zeta_R(s)$, and it is holomorphic in a neighborhood of $s = 0$.

(e) Use the preceding exercise together with the calculation

$$\zeta_{[0,\ell]}(s) = \frac{\ell^{2s}}{\pi^{2s}} \zeta_R(2s)$$

to show that $\zeta_{[0,\ell]}(s)$ is also holomorphic in a neighborhood of $s = 0$.

(f) Search for the special value of $\zeta_R'(0)$. Use this to explain the rather funny notion that

$$\prod_{n \geq 1} n = \sqrt{2\pi}.$$

Chapter 5

PDEs in bounded regions of space: stay calm and carry on, tackling challenges one-by-one!

To solve homogeneous partial differential equations on bounded intervals that have homogeneous boundary conditions, we can follow this three-step-procedure:

1. Separation of variables (a means to an end),
2. Superposition principle (smash solutions together to make a supersolution),
3. Fourier series to find the coefficients obtained using the initial data ($\mathcal{L}^2$ scalar product and divide by the norm).

Mathematics is in many ways similar to martial arts, as we will see in this chapter. The first similarity we emphasize is that both mathematics and martial are *experiential* subjects. It takes active practice to master either discipline. Perhaps this is one of the reason they are called disciplines; because their mastery requires just that - discipline! We are just beginning, so as in martial arts, we are a white belt like in Figure 5.1.

5.1 The homogeneous wave equation in a bounded interval

We shall begin our training to advance to the next rank by solving the equation:

$$u_{tt} = u_{xx}, \quad t > 0, \quad x \in (-1, 1),$$

$$\begin{cases} u(0,x) &= 1 - |x| \\ u_t(0,x) &= 0 \\ u_x(t,-1) &= 0 \\ u_x(t,1) &= 0 \end{cases}$$

We use separation of variables, writing $u(x,t) = X(x)T(t)$. We substitute this into the PDE:

$$T''X = X''T.$$

Divide everything by XT to get

$$\frac{T''}{T} = \frac{X''}{X}.$$

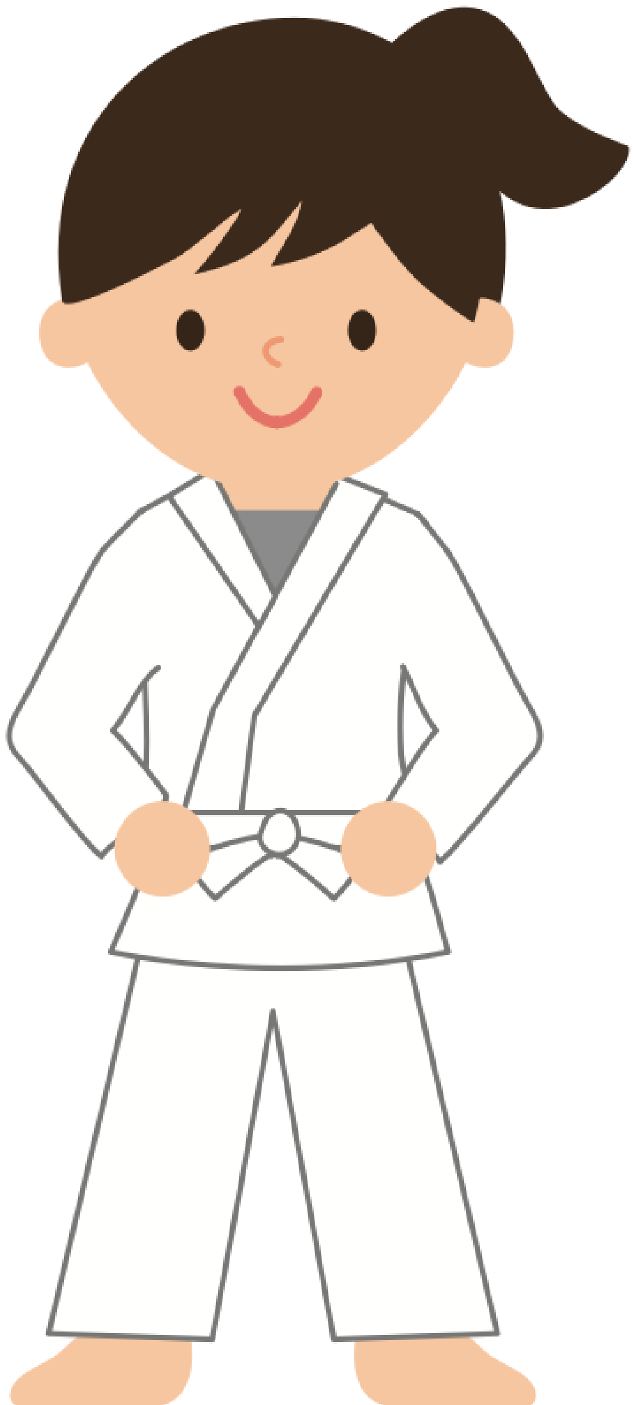


Figure 5.1: In many martial arts, beginners start fresh with the rank of white belt. According to legend, whereas you would wash your training attire, you would *not* wash your belt. You earned a black belt by training for so many years that your white belt eventually became black. I'm not entirely convinced by this, however. In many modern systems, the belts transition through a range of colors. In Tang Soo Do Moo Duk Kwan, for example, the original belt system was comprised of white, green, red, and midnight blue, representing winter, spring, summer, and fall. Similarly, these colors represent a bare, empty beginning in the winter, followed by the blossoming knowledge and skill in the spring, ripening over the summer, and maturity in the fall. Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

Since the two sides depend on different variables, they are both constant. We start with the X side because that is how we have learned to do this technique. The boundary conditions demand that

$$u_x(t,-1) = u_x(t,1) = 0 \implies X'(-1) = X'(1) = 0.$$

So, we have the equation

$$\frac{X''}{X} = \text{ constant, call it } \lambda.$$

Thus we are solving

$$X'' = \lambda X, \quad X'(-1) = X'(1) = 0.$$

Case 1: $\lambda = 0$: In this case, we have solved this equation before. One way to think about it is that the second derivative is like acceleration. If $X'' = 0$, it's saying X has constant acceleration. Therefore X can only be a linear function. Now, we have the boundary condition which says that $X'(-1) = X'(1) = 0$. So the slope of the linear function must be zero, hence X must be a constant function in this case. So, the only solutions in this case are the constant functions. Let's keep these in mind, because constant functions need not be zero, just constant.

Case 2: $\lambda > 0$: In this case, a general solution is of the form:

$$X(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}.$$

Let us assume that A and B are not both zero. The left boundary condition requires

$$A\sqrt{\lambda}e^{-\sqrt{\lambda}} - \sqrt{\lambda}Be^{\sqrt{\lambda}} = 0.$$

Since $\lambda > 0$ we can divide by $\sqrt{\lambda}$ to say that we must have

$$Ae^{-\sqrt{\lambda}} = Be^{\sqrt{\lambda}} \implies \frac{A}{B} = e^{2\sqrt{\lambda}}.$$

The right boundary condition requires

$$A\sqrt{\lambda}e^{\sqrt{\lambda}} - \sqrt{\lambda}Be^{-\sqrt{\lambda}} = 0.$$

Since $\lambda > 0$, we can divide by $\sqrt{\lambda}$, to make this:

$$Ae^{\sqrt{\lambda}} = Be^{-\sqrt{\lambda}} \implies e^{2\sqrt{\lambda}} = \frac{B}{A}.$$

Hence combining with the other boundary condition we get:

$$\frac{A}{B} = e^{2\sqrt{\lambda}} = \frac{B}{A} \implies A^2 = B^2 \implies A = \pm B \implies \frac{A}{B} = \pm 1.$$

Neither of these are possible because

$$e^{2\sqrt{\lambda}} > 1 \text{ since } 2\sqrt{\lambda} > 0.$$

So, we run amok under the assumption that A and B are not both zero. Hence, the only solution in this case requires $A = B = 0$. This is the waveless wave; the identically zero solution.

Case 3: $\lambda < 0$: In this case a general solution is of the form:

$$X(x) = a\cos(\sqrt{|\lambda|}x) + b\sin(\sqrt{|\lambda|}x).$$

To satisfy the left boundary condition we need

$$-a\sqrt{|\lambda|}\sin(-\sqrt{|\lambda|}) + b\sqrt{|\lambda|}\cos(-\sqrt{|\lambda|}) = 0 \iff a\sin(\sqrt{|\lambda|}) = -b\cos(\sqrt{|\lambda|}).$$

To satisfy the right boundary condition we need

$$-a\sqrt{|\lambda|}\sin(\sqrt{|\lambda|}) + b\sqrt{|\lambda|}\cos(\sqrt{|\lambda|}) = 0 \iff a\sin(\sqrt{|\lambda|}) = b\cos(\sqrt{|\lambda|}).$$

Hence we need

$$a\sin(\sqrt{|\lambda|}) = -b\cos(\sqrt{|\lambda|}) = b\cos(\sqrt{|\lambda|}). \tag{5.1.1}$$

We do not want both a and b to vanish. So, we need to have either

1. the sine vanishes, so we need $\sin(\sqrt{|\lambda|}) = 0$ which then implies that

$$\sqrt{|\lambda|} = n\pi, \quad n \in \mathbb{Z}$$

2. or the cosine vanishes so we need $\cos(\sqrt{|\lambda|}) = 0$ which then implies that

$$\sqrt{|\lambda|} = \left(n + \frac{1}{2}\right)\pi, \quad n \in \mathbb{Z}.$$

Note that these two cases are *mutually exclusive*. In case (1), by (5.1.1) this means that $b = 0$. In case (2), by (5.1.1) this means that $a = 0$. So, we have two types of solutions, which up to constant factor look like:

$$X_m(x) = \begin{cases} \cos(m\pi x/2) & m \text{ is even} \\ \sin(m\pi x/2) & m \text{ is odd} \end{cases}$$

In both cases,

$$\lambda_m = -\frac{m^2\pi^2}{4}.$$

We can now solve for the partner function, $T_m(t)$. The equation is

$$\frac{T_m''}{T_m} = \frac{X_m''}{X_m} = \lambda_m = -\frac{m^2\pi^2}{4}.$$

Therefore, we are in case 3 for the T_m function as well, so we know that

$$T_m(t) = a_m \cos\left(\frac{m\pi t}{2}\right) + b_m \sin\left(\frac{m\pi t}{2}\right).$$

Then we have for

$$u_m(x,t) = X_m(x)T_m(t), \quad \Box u_m = 0 \quad \forall m.$$

(Recall that $\Box = \partial_{tt} - \partial_{xx}$, that is the wave operator). Hence, our functions solve a homogeneous PDE, so we can use the superposition principle to smash them all together to make a super solution:

$$u(x,t) = \sum_{m\in\mathbb{Z}} u_m(x,t) = \sum_{m\in\mathbb{Z}} X_m(x)\left(a_m \cos\left(\frac{m\pi t}{2}\right) + b_m \sin\left(\frac{m\pi t}{2}\right)\right).$$

As we have observed before, we don't actually need all $m \in \mathbb{Z}$ because we can lump together terms with $\pm m$, so we can simplify this to

$$u(x,t) = \sum_{m\geq 0} u_m(x,t).$$

Note that the boundary condition will also be satisfied when we add all the solutions together, because $X_m'(\pm 1) = 0$ for all m, so we also get $\partial_x u_m(\pm 1, t) = X_m'(\pm 1)T_m(t) = 0$ for all m, and therefore the same holds for $u_x(\pm 1, t)$. Following the mantra TIDGLAS, the initial data goes last, and this is what we use to determine the coefficients in this last step. We determine these coefficients by expanding the initial data in a Fourier series.

The initial data is

$$\begin{cases} u(0,x) & = 1 - |x| \\ u_t(0,x) & = 0 \end{cases}$$

Let us plug $t = 0$ into our solution:

$$u(x,0) = \sum_{m\in\mathbb{N}} X_m(x)a_m.$$

We demand that this is the initial data, so we need

$$1 - |x| = \sum_{m \in \mathbb{N}} X_m(x) a_m.$$

It is a Fourier series on the right side!! We therefore just need to expand the function $1 - |x|$ in a Fourier series. If we think about the basis functions $\{X_m(x)\}_{m \geq 0}$ then

$$a_m = \frac{\langle 1 - |x|, X_m(x) \rangle}{||X_m||^2},$$

where

$$\langle 1 - |x|, X_m(x) \rangle = \int_{-1}^{1} (1 - |x|) \overline{X_m(x)} dx,$$

$$||X_m||^2 = \int_{-1}^{1} |X_m(x)|^2 dx.$$

On an exam, you are not actually required to compute these integrals!

Now, for the other coefficients (the b_n), we use the condition on the derivative:

$$u_t(x, 0) = \sum_{m \in \mathbb{N}} m_n \frac{m\pi}{2} X_m(x) = 0.$$

We know how to Fourier expand the zero function: its coefficients are all just zero. Hence, it suffices to take

$$b_m = 0 \forall m.$$

5.2 To deal with time independent inhomogeneities, don't worry mate, find a steady state!

Our three step plan works very well as long as there are no *inhomogeneities.* The heuristic way I think about inhomogeneities is that it is: *non-zero stuff that doesn't belong and can mess me up.* It is an enemy to be fought, so we gear up with our ninja garb and prepare to fight like this little ninja in Figure 5.2.

Roughly speaking, inhomogeneities are *non-zero stuff that we wish was not there.* If we see an inhomogeneity, this could be upsetting and cause us to worry. In Australia, a friendly way to refer to another person is 'mate.' So, when we start to worry because our problem has an inhomogeneity, we should just imagine a friendly Australian telling us, 'don't worry mate, find a steady state!' The reason is because when the inhomogeneity is *independent* of the time variable t, then we can use a method known as *steady state* to solve the inhomogeneity and reduce the problem to one where the inhomogeneous stuff has vanished!

5.2.1 Heat equation on an interval with an inhomogeneous time-independent boundary condition

We wish to solve the problem:

$$u_t - u_{xx} = 0, \quad 0 < x < 4, \quad t > 0,$$

$$u(x, 0) = v(x),$$
$$u_x(4, t) = 0,$$
$$u(0, t) = 20.$$

Figure 5.2: Would you rather fight one hundred duck sized horses or one horse sized duck? I would rather take on one tough opponent than try to deal with several simultaneously. Similarly, when we solve PDEs, it is best to tackle each inhomogeneity one-at-a-time. This is a commonly known martial arts strategy: divide and conquer!Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`.

Let us call this problem ♡. The boundary conditions are *not zero.* This means that we will have problems using superposition because unlike zero, adding up a bunch of 20s does not result in 20. Zero is special in this way; it is the only number you can add to itself as many times as you want and it stays the same. However, this inhomogeneous part of the boundary condition is the number 20, and that is constant, meaning it does not depend on time at all. It's just 20.[1] We can use a "steady state solution" to deal with this inhomogeneous boundary condition because the condition is independent of time. If the BC $u(0,t) = 20$ were instead $u(0,t) = 0$, then the BCs would be self adjoint BCs. So we want to make it so. Since the PDE is homogeneous, the

> Don't worry mate, find a steady state! Deal with non-self adjoint BCs which are independent of time by finding a steady state solution.

We want a function $f(x)$ which satisfies the equation

$$-f''(x) = 0,$$

and which gives us the bad BC

$$f(0) = 20.$$

[1]The show Space Ghost has a wonderfully funny episode featuring the Icelandic musician Björk as well as the British musician Tom York. A quote from the episode is '20! Yes!' `https://www.dailymotion.com/video/xliue`.

We have a nice homogeneous BC on the other side, so we don't want to mess that up, so we want

$$f'(4) = 0.$$

Then, the function

$$f(x) = ax + b.$$

We use the BCs to compute

$$f(0) = 20 \implies b = 20.$$

$$f'(4) = 0 \implies a = 0.$$

Figure 5.3: While visiting the Australian National University in Canberra, my friend and colleague Lashi Bandara invited me to help him do some volunteer work at a local wild animal refuge. We built a fence to expand the area available to the animals. I also got to meet some of the rescued animals including this little wombat. Keep our friends down under in mind when solving partial differential equations with the motto: *don't worry mate, find a steady state!*

When we use a steady state technique, we will need to take care of the initial conditions before proceeding. To help remember this:

Be careful with the steady state, don't leave things up to fate!

If we were to solve the problem:

$$u_t - u_{xx} = 0, \quad 0 < x < 4, \quad t > 0,$$

$$\begin{aligned} u(x,0) &= v(x), \\ u_x(4,t) &= 0, \\ u(0,t) &= 0. \end{aligned}$$

and add it to our steady state solution, then the initial condition would become $u(x,0) = v(x) + 20$ not $v(x)$. The steady state part always runs the risk of screwing up the initial conditions, because it does not depend on time. So, we need to be careful with the steady state, not just leave things up to fate! So, rather than searching for a solution with the original initial data, we search for a solution with the steady state subtracted from the initial data:

$$u_t - u_{xx} = 0, \quad 0 < x < 4, \quad t > 0,$$

$$\begin{aligned} u(x,0) &= v(x) - f(x), \\ u_x(4,t) &= 0, \\ u(0,t) &= 0. \end{aligned}$$

This is now a PDE that we know how to solve, so we call this problem ♡♡, because we like it even better than ♡. We use separation of variables to write $u = XT$ (just a means to an end).[2] Next, we get the equation

$$\frac{T'}{T} = \frac{X''}{X} = \lambda.$$

We solve the problem

$$X'' = \lambda X, \quad X(0) = 0 = X'(4).$$

Exercise 55. *Verify that the only solutions for the cases $\lambda \geq 0$ are solutions which are identically zero.*

We only get non-zero solutionsn for $\lambda < 0$. The solutions are of the form

$$a_n \cos(\sqrt{|\lambda_n|}x) + b_n \sin(\sqrt{|\lambda_n|}x).$$

The BC at 0 tells us that

$$a_n = 0.$$

The BC at 4 tells us that

$$\cos(\sqrt{|\lambda_n|}4) = 0 \implies \sqrt{|\lambda_n|}4 = \frac{2n+1}{2}\pi \implies \sqrt{|\lambda_n|} = \frac{2n+1}{8}\pi.$$

We then also get

$$\lambda_n = -\frac{(2n+1)^2\pi^2}{64}.$$

We shall deal with the coefficients at the very end. So, we set

$$X_n(x) = \sin(\sqrt{|\lambda_n|}x).$$

The partner function

$$\frac{T_n'}{T_n} = \lambda_n \implies T_n(t) = \alpha_n e^{\lambda_n t} = \alpha_n e^{-(2n+1)^2\pi^2 t/64}.$$

We put it all together writing

$$u(x,t) = \sum_{n\geq 1} T_n(t) X_n(x).$$

To make the IC, we need

$$u(x,0) = \sum_{n \geq 1} T_n(0) X_n(x) = v(x) - f(x).$$

Since

$$T_n(0) = \alpha_n,$$

[2]The French rap artist, M.C. Solaar, has a song called *La fin justifie les moyens,* meaning *the ends justify the means.* It is from the album Prose Combat, and both the song as well as the album are great! You can check out the song here: https://www.youtube.com/watch?v=qKCkgyZ78T4

we need

$$\sum_{n\geq 1} \alpha_n X_n(x) = v(x) - f(x).$$

So we want the coefficients to be the Fourier coefficients of $v - f$, thus

$$\alpha_n = \frac{\langle v-f, X_n\rangle}{||X_n||^2} = \frac{\int_0^4 (v(x)-f(x))\,\overline{X_n(x)}dx}{\int_0^4 |X_n(x)|^2 dx}.$$

Our full solution is

$$U(x,t) = u(x,t) + f(x) = 20 + \sum_{n\geq 1} T_n(t)X_n(x).$$

5.3 Fourier series with time dependent coefficient functions

What if the partial differential equation is inhomogeneous, and the inhomogeneous (non-zero) stuff depends on time? Then, whatever we use to deal with that will also necessarily depend on time. The trick will be to use a Fourier series that has *time dependent coefficients.* To show how this works, we consider the following problem:

$$u_t - u_{xx} = tx, \quad 0 < x < 4, \quad t > 0,$$

$$\begin{aligned} u(x,0) &= v(x), \\ u_x(4,t) &= 0, \\ u(0,t) &= 0. \end{aligned}$$

Non! Sacre bleu! Tabernac![3] This is an inhomogeneous PDE *and the inhomogeneity (tx) depends on time!* A steady-state solution cannot save us. What do we do? Our ninja might need two swords now like in Figure 5.4! The technique we will use here is to first solve the homogeneous problem, so that we obtain the basis functions for our Hilbert space, $\mathcal{L}^2$ of the interval on which we are working. Then, we will use these basis functions to construct our solution to the inhomogeneous problem by writing our solution as a Fourier series with *time-dependent coefficients.* Since the coefficients are time-dependent, we remember this technique by the following quote:

> Stay calm and carry on as the time moves along! Solve a time-dependent inhomogeneous PDE using a Fourier series with time-dependent coefficients!

We first solve the homogeneous problem. The reason we do this is because we will use the orthogonal basis functions to build the solution.

Exercise 56. *Use separation of variables to solve the homogeneous problem:*

$$w_t - w_{xx} = 0, \quad 0 < x < 4, \quad t > 0,$$

$$\begin{aligned} w(x,0) &= v(x) \\ w_x(4,t) &= 0 \\ w(0,t) &= 0. \end{aligned}$$

[3]This is how they curse in French Canada.

Figure 5.4: To deal with time dependent inhomogeneities we break out not just one sword but two! The idea is that first we will solve the homogeneous problem to obtain the basis functions for the interval on which we are working. That's like going at the problem with the first sword. Next, we take our second sword to the problem by writing our solution as a Fourier series with *time-dependent coefficients.* To remember this technique, we say *stay calm and carry on as the time moves along!* Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`.

Having done this, we obtain

$$\lambda_n = -\frac{(2n+1)^2\pi^2}{64}, \quad X_n(x) = \sin(\sqrt{|\lambda_n|}x).$$

$$T_n(t) = \alpha_n e^{\lambda_n t}.$$

$$\alpha_n = \frac{\langle v, X_n \rangle}{||X_n||^2} = \frac{\int_0^4 v(x)\overline{X_n(x)}dx}{\int_0^4 |X_n(x)|^2 dx},$$

and

$$w(x,t) = \sum_{n\geq 1} T_n(t) X_n(x).$$

Now, we look for a solution to this problem:

$$\phi_t - \phi_{xx} = tx, \quad 0 < x < 4, \quad t > 0,$$

$$\phi(x,0) = 0,$$
$$\phi_x(4,t) = 0,$$
$$\phi(0,t) = 0.$$

Idea!

Look for a solution of the form

$$\sum_{n\geq 1} c_n(t) X_n(x).$$

So, we keep our X_n from the homogeneous problem, and we look for different $c_n(t)$ which will now be *functions of* t. We want the function to satisfy

$$u_t - u_{xx} = tx,$$

so we put the series in the left side into this PDE:

$$\sum_{n\geq 1} c_n'(t)X_n(x) - c_n(t)X_n''(x) = tx.$$

We use the fact the $X_n'' = \lambda_n X_n$, so we want to solve

$$\sum_{n\geq 1} X_n(x)\left(c_n'(t) - c_n(t)\lambda_n\right) = tx.$$

Here is where we need an idea.

Idea!

Write out tx as a Fourier series in terms of X_n.

The t just goes along for the ride, and

$$tx = t\sum_{n\geq 1} a_n X_n(x),$$

where

$$a_n = \frac{\langle x, X_n\rangle}{||X_n||^2} = \frac{\int_0^4 xX_n(x)dx}{\int_0^4 |X_n|^2 dx}.$$

As usual, we do not need to compute these integrals.

So, we want:

$$\sum_{n\geq 1} X_n(x)\left(c_n'(t) - c_n(t)\lambda_n\right) = tx = \sum_{n\geq 1} tX_n(x)a_n.$$

We equate the coefficients of X_n:

$$\left(c_n'(t) - \lambda_n c_n(t)\right) = ta_n.$$

This is an ODE for $c_n(t)$. We also want the IC, $c_n(0) = 0$. The solution to the homogeneous ODE,

$$f' - \lambda_n f = 0 \implies f(t) = e^{\lambda_n t} \text{ times some constant factor.}$$

A particular solution to the inhomogeneous ODE is a linear function of the form:

$$A_n t + B_n \implies A_n - \lambda_n(A_n t + B_n) = a_n t \implies A_n = \frac{-a_n}{\lambda_n}, \quad B_n = \frac{A_n}{\lambda_n} = -\frac{a_n}{\lambda_n^2}.$$

So general solutions are of the form:

$$c_n(t) = C_n e^{\lambda_n t} - \frac{a_n}{\lambda_n}t - \frac{a_n}{\lambda_n^2}, \quad \text{for some constant } C_n.$$

To obtain the initial condition that $c_n(0) = 0$, we see that we need

$$C_n = \frac{a_n}{\lambda_n^2}.$$

Thus, we have found

$$c_n(t) = \frac{a_n}{\lambda_n^2}e^{\lambda_n t} - \frac{a_n}{\lambda_n}t - \frac{a_n}{\lambda_n^2}.$$

Therefore the solution we seek is

$$u(x,t) = \sum_{n\geq 1} c_n(t)X_n(x),$$

and the full solution to the original problem is

$$U(x,t) = w(x,t) + u(x,t).$$

5.4 The homogeneous wave equation inside a rectangle

Sometimes the equation is homogeneous, but in order to solve it, we will be forced to deal with inhomogeneities. The homogeneous wave equation inside a rectangle provides a good example of how and when that can happen. We wish to solve the homogeneous wave equation inside a rectangle:

$$\Box u = 0 \text{ inside a rectangle, } u(x,y,0) = f(x,y), \quad u_t(x,y,0) = 0,$$

$$u(x,y,t) = g(x,y) \text{ for } (x,y) \text{ on the boundary of the rectangle.}$$

We name this problem $\heartsuit$. Here we have an inhomogeneous boundary condition. So, to solve the problem, we break it into two smaller problems which we tackle *one at a time:* divide and conquer.

> Deal with time independent boundary conditions by finding a steady state solution. Don't worry mate, find a steady state!

So, we begin by looking for

$$\Phi(x,y)$$

to satisfy

$$\Box \Phi = 0 \text{ inside the rectangle,}$$

$$\Phi = g \text{ on the boundary of the rectangle.}$$

Since the physical problem doesn't care where in space the rectangle is sitting, let us put it so that its vertices are at $(0,0)$, $(0,B)$, $(A,0)$, (A,B). Let us call this problem $\heartsuit\heartsuit$.

Once we have found Φ, we will look for a solution w to solve

$$\Box w = 0 \text{ inside the rectangle,}$$

$$w(x,y,t) = 0 \text{ on the boundary of the rectangle,}$$

$$w(x,y,0) = f(x,y) - \Phi(x,y), \quad w_t(x,y,0) = 0.$$

Then, our solution to $\heartsuit$ will be

$$u(x,y,t) = w(x,y,t) + \Phi(x,y).$$

So, we look for Φ to solve $\heartsuit\heartsuit$.

> Deal with each inhomogeneous boundary component one at a time. Stay calm and carry on, tackling challenges one-by-one!

It is the same principle: divide and conquer. So, first, let us make nice zero boundary conditions on the sides, and just deal with the complicated boundary conditions on the top and bottom. Therefore we look for a function $\phi(x,y)$ which satisfies

$$\Box \phi = 0,$$

$$\phi(0,y) = \phi(A,y) = 0,$$

$$\phi(x,0) = g(x,0), \quad \phi(x,B) = g(x,B).$$

Idea!

Since the PDE is homogeneous and half of the BCs are good and homogeneous, use separation of variables.
We therefore write the PDE:

$$-X''Y - Y''X = 0 \implies -\frac{Y''}{Y} = \frac{X''}{X} = \lambda.$$

The BCs for X are $X(0) = X(A) = 0$. We have solved this problem. The solutions are, up to constant factors

$$X_n(x) = \sin\left(\frac{n\pi x}{A}\right), \quad \lambda_n = -\frac{n^2\pi^2}{A^2}.$$

The equation for the partner function is then:

$$-\frac{Y_n''}{Y_n} = \lambda_n \implies Y_n'' = \frac{n^2\pi^2}{A^2} Y_n.$$

A basis of solutions is given by real exponentials, or equivalently hyperbolic sines and cosines. Since our region contains 0, we have been given a hint that using the hyperbolic sines and cosines may be more simple. So, we follow that hint, with

$$Y_n(y) = a_n \cosh\left(\frac{n\pi y}{A}\right) + b_n \sinh\left(\frac{n\pi y}{A}\right).$$

Next we use superposition to create a super solution, which is legit because the PDE is homogeneous:

$$\phi(x,y) = \sum_{n\geq 1} X_n(x) Y_n(y).$$

To obtain the boundary conditions, we need

$$\phi(x,0) = g(x,0) = \sum_{n\geq 1} a_n X_n(x).$$

Hence, the coefficients

$$a_n = \frac{\langle g(x,0), X_n\rangle}{||X_n||^2} = \frac{\int_0^A g(x,0)\overline{X_n(x)}dx}{\int_0^A |X_n(x)|^2 dx}.$$

For the other BC, we need

$$\phi(x,B) = g(x,B) = \sum_{n\geq 1} X_n(x)\left(a_n \cosh\left(\frac{n\pi B}{A}\right) + b_n \sinh\left(\frac{n\pi B}{A}\right)\right).$$

Therefore we need

$$\left(a_n \cosh\left(\frac{n\pi B}{A}\right) + b_n \sinh\left(\frac{n\pi B}{A}\right)\right) = \frac{\langle g(x,B), X_n\rangle}{||X_n||^2}$$
$$= \frac{\int_0^A g(x,B) X_n(x) dx}{\int_0^A |X_n(x)|^2 dx}.$$

Solving for b_n we get

$$b_n = \frac{1}{\sinh(\frac{n\pi B}{A})}\left(\frac{\langle g(x,B), X_n\rangle}{||X_n||^2} - a_n \cosh\left(\frac{n\pi B}{A}\right)\right).$$

Next, we proceed similarly by searching for a function to fix up the BCs on the left and the right. Having dealt with the inhomogeneous BCs at the top and bottom, we set the BC there equal to zero. In that way, when we sum, we shall not mess up the function ϕ. So, we look for a solution to:

$$\Box\psi(x,y) = 0, \quad \psi(x,0) = \psi(x,B) = 0, \quad \psi(0,y) = g(0,y), \quad \psi(A,y) = g(A,y).$$

By symmetry, the solution will be given by

$$\sum_{n\geq 1} \widetilde{X_n}(y)\widetilde{Y_n(x)},$$

with

$$\widetilde{X_n(y)} = \sin\left(\frac{n\pi y}{B}\right),$$

and

$$\widetilde{Y_n(x)} = \widetilde{a_n}\cosh\left(\frac{n\pi x}{B}\right) + \widetilde{b_n}\sinh\left(\frac{n\pi x}{B}\right).$$

The coefficients come from the boundary conditions:

$$\widetilde{a_n} = \frac{\langle g(0,y), \widetilde{X_n}\rangle}{||\widetilde{X_n}||^2} = \frac{\int_0^B g(0,y)\widetilde{X_n(y)}dy}{\int_0^B |X_n(y)|^2 dy}.$$

The other one

$$\widetilde{b_n} = \frac{1}{\sinh\left(\frac{n\pi A}{B}\right)}\left(\frac{\langle g(A,y), \widetilde{X_n}\rangle}{||\widetilde{X_n}||^2} - \widetilde{a_n}\cosh\left(\frac{n\pi A}{B}\right)\right).$$

So, we have found

$$\psi(x,y) = \sum_{n\geq 1} \widetilde{X_n(y)}\widetilde{Y_n(x)}.$$

The full solution to this part of the problem is

$$\Phi(x,y) = \phi(x,y) + \psi(x,y).$$

Exercise 57. *Verify that this function satisfies both the PDE* $\Box\Phi = 0$ *as well as all of the boundary conditions.*

To complete the problem, we have only to solve the homogeneous wave equation with the lovely Dirichlet boundary condition and the initial condition with Φ subtracted. So, we are solving:

$$\Box u = 0, \quad u_t(x,y,0) = 0, \quad u(x,y,0) = f(x,y) - \Phi(x,y), \quad u = 0 \text{ on the boundary.}$$

Idea!

Since we have homogeneous PDE and BC, use separation of variables and superposition. We use separation of variables for t, x, and y. Write

$$u = TXY.$$

The PDE is

$$T''XY - X''TY - Y''TX = 0 \iff \frac{T''}{T} = \frac{X''}{X} + \frac{Y''}{Y} = \lambda.$$

Since we have nice homogeneous (Dirichlet) boundary conditions, we begin with the functions that depend on the position in the rectangle, that is X and Y.

Their equation is:

$$\frac{X''}{X} + \frac{Y''}{Y} = \lambda \implies \frac{X''}{X} = \lambda - \frac{Y''}{Y}.$$

OBS! The left and right sides depend on *different independent variables.* Hence, by the same reasoning that gave us λ, we get that

$$\frac{X''}{X} = \lambda - \frac{Y''}{Y} = \mu.$$

Let us solve for X first.[4] So, we are looking to solve:

$$X'' = \mu X, \quad X(0) = X(A) = 0.$$

We have solved this before. The solutions are up to constant factors:

$$X_n(x) = \sin\left(\frac{n\pi x}{A}\right) \quad \mu_n = -\frac{n^2\pi^2}{A^2}.$$

This gives the equation for Y,

$$\frac{Y''}{Y} = \lambda - \mu_n, \quad Y(0) = Y(B) = 0.$$

Let us briefly call

$$\nu = \lambda - \mu_n.$$

Then, this is just the same equation but with different names for things:

$$Y'' = \nu Y, \quad Y(0) = Y(B) = 0.$$

Up to constant factors, the solutions are

$$Y_m(y) = \sin\left(\frac{m\pi y}{B}\right) \quad \nu_m = -\frac{m^2\pi^2}{B^2}.$$

Since

$$\nu_m = \lambda - \mu_n \implies \lambda = \lambda_{n,m} = \nu_m + \mu_n = -\frac{m^2\pi^2}{B^2} - \frac{n^2\pi^2}{A^2}.$$

Recalling the equation for the partner function, T, we have

$$T_{n,m}(t) = a_{n,m}\cos(\sqrt{|\lambda_{n,m}|}t) + b_{n,m}\sin(\sqrt{|\lambda_{n,m}|}t).$$

Hence we write

$$u(x,y,t) = \sum_{n,m\geq 1} T_{n,m}(t)X_n(x)Y_m(y).$$

The initial condition

$$u_t(x,y,0) = 0 \implies b_{n,m} = 0 \forall n,m.$$

The other condition is that

$$u(x,y,0) = f(x,y) - \Phi(x,y) = \sum_{n,m\geq 1} a_{n,m}X_n(x)Y_m(y).$$

Hence we require

$$a_{n,m} = \frac{\langle f-\Phi, X_nY_m\rangle}{||X_nY_m||^2} = \frac{\int_{[0,A]\times[0,B]}(f(x,y)-\Phi(x,y))X_n(x)Y_m(y)dxdy}{\int_{[0,A]\times[0,B]}|X_n(x)Y_m(y)|^2dxdy}.$$

The full solution is then

$$u(x,y,t) + \Phi(x,y).$$

Remark 3. The numbers we obtained above

$$\lambda_{n,m} = -\frac{m^2\pi^2}{B^2} - \frac{n^2\pi^2}{A^2}$$

[4]In this case, we could solve for either X or Y first, it actually does not matter which you choose.

correspond to the resonant frequencies produced by a vibrating rectangle that has side lengths equal to A and B. It is interesting to compare this to the numbers we obtain for a vibrating string of length ℓ,

$$\mu_n = -\frac{n^2\pi^2}{\ell^2}.$$

These are the resonant frequencies produced by a vibrating string. As you can see they are all square integer multiples of

$$\mu_1 = -\frac{\pi^2}{\ell^2},$$

that corresponds to the ground tone or fundamental tone. Consequently, for a vibrating string, all of the higher harmonics, corresponding to μ_n for $n \geq 2$, are square integer multiples of the ground tone. This is the mathematical reason that vibrating strings sound lovely. On the other hand, if the rectangle is not a square, that is $A \neq B$, it is no longer true that the $\lambda_{n,m}$ are all multiples of

$$\lambda_{1,1} = -\frac{\pi^2}{B^2} - \frac{\pi^2}{A^2}.$$

For this reason, vibrating rectangles can sound rather awful. You can listen to something similar, for tori, here: http://www.toroidalsnark.net/som.html.

5.5 Mathematical physics and martial arts

According to Chinese legends, a true martial arts master cannot be defeated even when fighting blindfolded. One of the reasons is because they are able to hear everything: the location of the opponent, what type of weapon the opponent is using, and what exactly the opponent is doing at every moment in time. Is this possible? Can we investigate this using mathematical physics? The answer to the first question is, maybe. The answer to the second question is more encouraging: yes we can! For those of you who can read Mandarin, you might find our article interesting https://www.global-sci.com/intro/article_detail.html?journal=mc&article_id=13178.[5]

This question is related to Kac's famous article [9] titled *can one hear the shape of a drum?* Mathematically, this question is: *if two bounded domains in the plane have the same eigenvalue spectrum for the Laplace eigenvalue equation with the Dirichlet boundary condition, then are the domains the same shape?* The eigenvalues determine the sound of the domain, if it were the drumhead of a vibrating drum. Consequently, the question is, if we listen to two drumheads with perfect hearing, and they sound identical because all of the eigenvalues are identical, then are the shapes of the drumheads identical? Mathematically, this eigenvalue problem is to find all solutions u such that for some $\lambda \in \mathbb{C}$

$$u_{xx} + u_{yy} - \lambda u = 0, \text{ inside the domain, and } u = 0 \text{ on the boundary of the domain.} \tag{5.5.1}$$

This is a very difficult problem, because although we can prove general facts like that λ must actually be real and non-negative, we cannot calculate these numbers analytically. There are only a handful of examples for which we can calculate these eigenvalues analytically, and many of those examples will be done in this text. So, if we cannot calculate these eigenvalues, how on earth could we hope to answer Kac's question? Two domains have associated to them two unknown sets of numbers that are the same and somehow we want to know if the domains are the same shape? This seems impossible.

Indeed, it is certainly not easy. To answer it, we investigate quantities that are determined by these eigenvalues. The set of all eigenvalues of a domain is called its *spectrum,* and consequently the quantities that are determined by the eigenvalues are called *spectral invariants.* Of particular importance are geometric spectral invariants, that are the geometric features of a domain that are determined by the spectrum. Hermann Weyl discovered the first geometric spectral invariant in about 1912: the area of a domain is a

[5]My Chinese name is 罗茱莉.

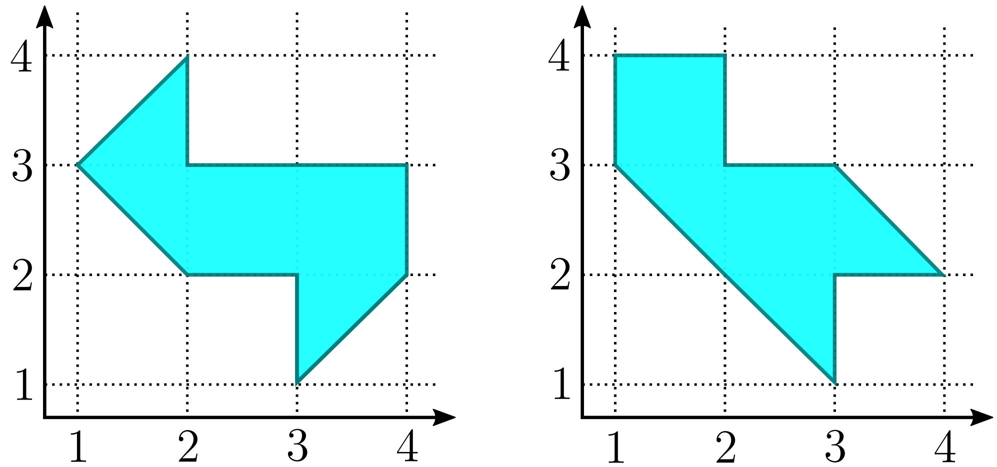


Figure 5.5: These two domains have exactly the same set of Laplace eigenvalues with the Dirichlet boundary condition. As one can see, they are not the same shape! This image is kindly contributed by Gottfrid Olsson, based on the example in [6] and [7].

spectral invariant [24] About a half century later, Åke Pleijel [19] proved that the perimeters is a spectral invariant. Shortly thereafter, in 1965, M. Kac wrote his paper. It took about a quarter century to solve the problem, which was achieved by Carol Gordon, David Webb, and Scott Wolpert in 1991 [6, 7]. The answer is no.

In contrast to a nice round drumhead, the "identical sounding drums," in Figure 5.5 both have corners. A natural question is therefore: can one hear the corners of a drum? This means, is it possible for two drums to sound the same, and one of them has a nicely rounded, but not necessarily circular, shape, whereas the other has at least one sharp corner? In other words, can one hear the corners of a drum? My co-authors and I have proved that the answer is *yes* [13, 16]. The sound produced by a drumhead with at least one sharp corner will always be different from the sound produced by any drumhead without corners. My co-author and I also proved that one can hear the symmetry of regular polygons in the sense that if an n-sided polygon has the same spectrum as a regular n-gon, then in fact that n-sided polygon is regular. These results require that the *entire* spectra be equal, and it can be shown that these sets of eigenvalues are always infinite. In terms of the physical interpretation, that means one would need perfect hearing. We proved that if one is only listening to convex n-sided polygons, then in fact, one can detect regular n-gons through a *finite* collection of eigenvalues [12]. In this sense, one could realistically expect to be able to distinguish a regular n-gon by listening to the sound it makes as the head of a vibrating drum.

Returning to our Kung Fu master, these results indicate that if there could be some truth to the Chinese folklore, and perhaps one day it may even be possible to mathematically prove that a Kung Fu master with sufficiently good hearing can detect the size and shape of their opponent and their weapon! One of the most famous martial artists of all time, Bruce Lee, also majored in philosophy at the University of Washington, but unfortunately he finished his studies there long before I began studying mathematics. Could Bruce Lee, shown in Figure 5.6 recognize his opponent's weapon from only its sound? We will never know, but we will continue to investigate this question from a mathematical physics perspective!

5.6 Exercises

1. [4, 4.2.1] Suppose the end $x = 0$ of a rod of length ℓ is held at temperature zero while the end at $x = \ell$ is insulated. This means that the boundary conditions are $u(0, t) = 0$, and $u_x(\ell, t) = 0$.

   (a) Find a series expansion for the temperature $u(x, t)$ given the initial temperature $f(x) = u(x, 0)$.

Figure 5.6: Bruce Lee, born 李振藩 but more commonly known as 李小龙 was an actor, director, and one of the most influential martial artists of all time. His philosophy was to draw from different combat disciplines, rather than limiting himself to a single style. This is one of the reasons that Bruce Lee is often credited with paving the way for modern mixed martial arts. He lived from 1940–1972. This image is public domain.

(b) What is $u(x,t)$ when $f(x) = 50$?

2. [4, 4.2.5] Solve:
$$u_t = ku_{xx} + e^{-2t}\sin(x),$$
with
$$u(x,0) = u(0,t) = u(\pi,t) = 0.$$

3. A function is 2 periodic with $f(x) = (x+1)^2$ for $|x| < 1$. Expand $f(x)$ in a Fourier series. Search for a 2 periodic solution to the equation
$$2y'' - y' - y = f(x).$$

4. [4, 4.2.6] Solve:
$$u_t = ku_{xx} + Re^{-ct}, \quad R, c > 0,$$
$$u(0,t) = 0, \quad u_x(\ell,t) = 0, \quad u(x,0) = f(x).$$
Physically this is heat flow in a rod which has a chemical reaction in it such that the reaction produced inside the rod dies out over time. The end of the rod at $x = 0$ is held at temperature zero, while the right end of the rod is insulated, and the initial temperature inside the rod is $f(x)$.

5. [4, 4.3.5] Find the general solution of
$$u_{tt} = c^2 u_{xx} - a^2 u,$$
$$u(0,t) = u(l,t) = 0,$$
with arbitrary initial conditions. Physically, this is a model for a string vibrating in an elastic medium where the term $-a^2u$ represents the force of reaction of the medium on the string.

6. [4, 4.2.7] We have a radioactive rod of length ℓ, that is to say it generates heat within itself at a constant rate R, so the equation the temperature function $u(x,t)$ satisfies is
$$u_t = ku_{xx} + R.$$
Assume that it is insulated at both ends and along its length, and initially it is at temperature zero, so the boundary and initial conditions are
$$u(x,0) = u_x(0,t) = u_x(\ell,t) = 0.$$
Show that there is no steady state solution of $u_t = ku_{xx} + R$ with $u_x(0,t) = u_x(\ell,t) = 0$. Why does that make sense from a physics perspective? Then solve this problem!

7. Solve the problem:
$$u_{xx} + u_{yy} = y, \quad 0 < x < 2, \quad 0 < y < 1$$
$$u(x,0) = 0, \quad u(x,1) = 0$$
$$u(0,y) = y - y^3, \quad u(2,y) = 0.$$

8. Solve the problem
$$u_{xx} + u_{yy} + 20u = 0, \quad 0 < x < 1, \quad 0 < y < 1,$$
$$u(0,y) = u(1,y) = 0$$
$$u(x,0) = 0, \quad u(x,1) = x^2 - x.$$

9. [4, 4.4.1] Solve the equation

$$u_{xx} + u_{yy} = 0$$

inside the square $0 \leq x, y \leq l$, subject to the boundary conditions:

$$u(x, 0) = u(0, y) = u(l, y) = 0, \quad u(x, l) = x(l - x).$$

10. Expand the function $\cos(x)$ in a sine series on the interval $(0, \pi/2)$. Use the result to compute

$$\sum_{n \geq 1} \frac{n^2}{(4n^2 - 1)^2}.$$

11. [4, 4.2.2] Solve:

$$u_t = ku_{xx}, \quad u(x, 0) = f(x),$$

$$u(0, t) = C \neq 0, \quad u_x(l, t) = 0.$$

12. [4, 4.3.1] Show that the function

$$b_n(t) := \frac{\ell}{n\pi c} \int_0^t \sin \frac{n\pi c(t - s)}{\ell} \beta_n(s) ds$$

solves the differential equation:

$$b_n''(t) + \frac{n^2 \pi^2 c^2}{\ell^2} b_n(t) = \beta_n(t),$$

as well as the initial conditions $b_n(0) = b_n'(0) = 0$.

13. Show that if there is a complex-valued solution to (5.5.1), then its real and imaginary parts are also solutions.

14. Show that in fact the λ in (5.5.1) is real valued and non-negative.

15. [4, 4.2.7] We have a rod of length ℓ, in which a chemical reaction is occurring that produces heat over time, but the amount of heat is decreasing over time, so the equation the temperature function $u(x, t)$ satisfies is

$$u_t = ku_{xx} + Re^{-ct}.$$

Assume that it is insulated at both ends and along its length, and initially it is at temperature zero, so the boundary and initial conditions are

$$u(x, 0) = u_x(0, t) = u_x(\ell, t) = 0.$$

Determine $u(x, t)$ for $t > 0$.

16. [4, 4.3.2] One end of an elastic bar of length ℓ is held at $x = 0$, and the other end is stretched from its natural position to $(1 + b)\ell$. So, an arbitrary point x in the bar is moved to $(1 + b)x$, and its displacement from equilibrium is bx. At time $t = 0$ the ends of the bar are released; thus $u(x, 0) = bx$, and $u_t(x, 0) = 0$. Here the boundary conditions are that $u(0, t) = u_x(\ell, t) = 0$, corresponding to the left endpoint being fixed, and the right endpoint being free. This could also represent a column of air that is closed at one end and open at the other, like in a musical instrument, like a clarinet. Find the displacement $u(x, t)$ at times $t > 0$.

17. [4, 4.3.3] A horizontally stretched string is so heavy that gravity effects the wave equation, so that it becomes

$$u_{tt} = \mathfrak{c}^2 u_{xx} - g,$$

where g is the acceleration due to gravity. Assume the ends of the string are fixed. Find a steady state solution. Call it $\phi(x)$. Suppose initially that $u(x,0) = u_t(x,0) = 0$. Find the solution u, and show that it can be expressed as

$$u(x,t) = \phi(x) - \frac{1}{2}\left[\Phi(x+ct) + \Phi(x-ct)\right],$$

where Φ is the odd 2ℓ periodic extension of ϕ.

18. [4, 4.3.6] In real-life vibrating strings, the vibrations damp out because the strings are not perfectly elastic. This can be modeled by the modified wave equation

$$u_{tt} = \mathfrak{c}^2 u_{xx} - 2ku_t,$$

where the term $-2ku_t$ represents the frictional forces causing the dampening. Find the solution for general initial data subject to the boundary condition $u(0,t) = u(\ell,t) = 0$, in which the ends are fixed. Something interesting happens when the value of $k > 0$ is either $k < \pi\mathfrak{c}/\ell$ compared to $k \geq \pi\mathfrak{c}/\ell$...

19. [4, 4.3.8] The total energy of a vibrating string at time t is up to a constant factor given by

$$E(t) = \int_0^\ell [u_t(x,t)^2 + \mathfrak{c}^2 u_x(x,t)^2] dx.$$

The first term is kinetic energy, and the second term is the potential energy, and here we assume u is real-valued. If the string has fixed ends, derive a series expression for E dependent on the Fourier expansion of the initial data. Use this to demonstrate conservation of energy; $E(t)$ is independent of t.

20. [4, 4.4.3] Here we consider the Neumann problem in a square. This could arise as part of solving the heat equation in a square that has three sides which are insulated, and one side from which it is receiving or losing heat. We therefore wish to solve:

$$u_{xx} + u_{yy} = 0, \quad 0 < x, y < \ell, \quad u_x(0,y) = u_x(\ell,y) = u_y(0,y), \quad u_y(x,\ell) = f(x).$$

Show that a solution exists only if $\int_0^\ell f(x)dx = 0$, and that in this case the solution contains an arbitrary constant. What could that mean, physically?

21. [4, 4.4.5] Consider an annulus, that in polar coordinates is the set $\{(r,\theta) : r_0 \leq r \leq 1\}$ for some $0 < r_0 < 1$. Assume that the inner side is insulated, and the outer side is held at temperature $u(1,\theta) = f(\theta)$. Find the steady-state temperature. What is the solution if $f(\theta) = 1 + 2\sin\theta$?

22. [4, 4.4.7] Solve the Dirichlet problem in $S = \{(r,\theta) : 0 \leq r_0 \leq r \leq 1, 0 \leq \theta \leq \beta\}$ with $u(r_0,\theta) = u(1,\theta) = 0$, $u(r,0) = g(r)$, $u(r,\beta) = h(r)$. By the Dirichlet problem, here we wish that $u_{xx} + u_{yy} = 0$ inside S. (It is probably best to re-write the equation using polar coordinates!)

Chapter 6

Sturm-Liouville problems: need to solve a PDE? An SLP might be the key!

The process of separating variables to solve partial differential equations resulted in solving the problem:

$$X''(x) = \lambda X(x), \text{ subject to some boundary conditions.}$$

This is an example of a general type of problem that is called *a regular Sturm-Liouville problem.* For example, if we consider the interval $[-\pi, \pi]$, the functions which satisfy

$$X''(x) = \lambda X(x), \quad X(-\pi) = X(\pi)$$

are

$$X(x) = X_n(x) = e^{inx}, \quad n \in \mathbb{Z}$$

or equivalently

$$\{1, \sin(nx), \cos(nx)\}_{n \geq 1}$$

with the corresponding

$$\lambda_n = n^2.$$

If we introduce the differential operator

$$\Delta := \frac{d^2}{dx^2},$$

then the equation we solved is

$$\Delta X - \lambda X = 0,$$

or equivalently letting

$$\mu := -\lambda, \quad \Delta X + \mu X = 0.$$

This is an example of a Sturm-Liouville problem. A Sturm-Liouville problem is an *eigenvalue* problem for a differential operator. The goal is to find all eigenvalues, μ, such that there exists a corresponding eigenfunction, X, that satisfies the equation. This is like an infinite dimensional version of finding the eigenvalues of a matrix. Instead of a matrix, we have a differential operator. Instead of a finite dimensional vector space, we have an infinite dimensional vector space whose elements are functions. Here we take a brief foray into the general subject of Sturm-Liouville problems, exploring their theory, and learning how to solve them. SLPs often arise as a step in solving a partial differential equation using separation of variables. Consequently, we can imagine that an SLP allows us to slip and slide to the solution, with the analogy further inspired by the fact that pronouncing SLP as though it were a word sounds rather like 'slip.' Let's follow the Panda in Figure 6.1 to the solution! Our motto for this chapter is:

Figure 6.1: Since SLPs are often a step in solving a PDE using separation of variables, we can imagine that they are a technique that allows us to SLP, pronounced like slip without the 'i' to the solution. Need to solve a PDE? An SLP might be the key! Slp, slide, and glide, like this ice skating panda bear, and the solution will have no place to hide. Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

Slp, slide, and glide; the solution has no place to hide!

6.1 Regular Sturm-Liouville Problems

Let L be a linear, second order ordinary differential operator. So, we can write

$$L(f) = r(x)f''(x) + q(x)f'(x) + p(x)f(x).$$

Above, r, q, and p are specified real valued functions. As a simple example, take $r(x) = 1$, and $q(x) = p(x) = 0$. Then we have

$$L(f) = f''(x).$$

We are working with functions defined on an interval $[a, b]$ which is a *finite* interval. So, the Hilbert space in which everything is happening is $\mathcal{L}^2$ on that interval. Like with matrices, we can think about the *adjoint* of the operator L. The adjoint by definition satisfies

$$\langle Lf, g\rangle = \langle f, L^* g\rangle,$$

where we are using L^* to denote the adjoint operator. Whatever it is. On the left side, we know what everything is, so we write it out by definition of the scalar product

$$\langle Lf, g\rangle = \int_a^b L(f)\overline{g(x)}dx = \int_a^b \left(r(x)f''(x) + q(x)f'(x) + p(x)f(x)\right)\overline{g(x)}dx.$$

Integrating by parts, we get

$$= (r\bar{g})f'\big|_a^b - \int_a^b (r\bar{g})'f' + (qg)f\big|_a^b - \int_a^b (q\bar{g})'f + \int_a^b pf\bar{g}$$

$$= (r\bar{g})f' + (q\bar{g})f\big|_a^b - \int_a^b \left[(r\bar{g})'f' + (q\bar{g})'f - pf\bar{g}\right].$$

We integrate by parts once more on the $(r\bar{g})'f'$ term to get

$$= (r\bar{g})f' - (r\bar{g})'f + (q\bar{g})f)\big|_a^b + \int_a^b (r\bar{g})''f - (q\bar{g})'f + fp\bar{g}.$$

So, if the boundary conditions are chosen to make the stuff evaluated from a to b (these are called the boundary terms in integration by parts) vanish, then we could define

$$L^*g = (rg)'' - (qg)' + pg,$$

since then

$$\langle Lf, g\rangle = \int_a^b (r\bar{g})''f - (q\bar{g})'f + fp\bar{g} = \langle f, L^*g\rangle.$$

Here we use that r, q and p are real valued functions, so $\bar{r} = r$, $\bar{q} = q$, and $\bar{p} = p$. For the spectral theorem to work, we will want to have

$$L = L^*.$$

When this holds, we say that L *is formally self-adjoint.* So, we need

$$Lf = L^*f \iff rf'' + qf' + pf = (rf)'' - (qf)' + pf.$$

We write the things out:

$$rf'' + qf' + pf = (rf' + r'f)' - qf' - q'f + pf \iff rf'' + qf' = rf'' + 2r'f' + r''f - qf' - q'f$$

$$\iff qf' = 2r'f' + r''f - qf' - q'f \iff (2q - 2r')f' + (r'' - q')f = 0.$$

To ensure this holds for all f, we set the coefficient functions equal to zero:

$$2q - 2r' = 0 \implies q = r', \quad q' = r''.$$

Well, that just means that $q = r'$. So, we need L to be of the form

$$Lf = rf'' + r'f' + pf = (rf')' + pf.$$

The boundary terms should also vanish, so we want:

$$(r\bar{g})f' - (r\bar{g})'f + (q\bar{g})f)\big|_a^b = (r\bar{g})f' - (r\bar{g})'f + (r'\bar{g})f\big|_a^b = 0,$$

$$\iff r\bar{g}f' - r'\bar{g}f - r\bar{g}'f + r'\bar{g}f\big|_a^b = 0 \iff r\bar{g}f' - r\bar{g}'f\big|_a^b = 0$$

$$\iff r(\bar{g}f' - \bar{g}'f)\big|_a^b = 0.$$

So, we would like to guarantee that r, f, and g satisfy

$$r(\bar{g}f' - \bar{g}'f)\big|_a^b = 0.$$

Writing this out we get:

$$r(b)\left(\bar{g}(b)f'(b)-\bar{g}'(b)f(b)\right)-r(a)\left(\bar{g}(a)f'(a)-\bar{g}'(a)f(a)\right)=0 \iff$$

$$r(b)\left(\bar{g}(b)f'(b)-\bar{g}'(b)f(b)\right)=r(a)\left(\bar{g}(a)f'(a)-\bar{g}'(a)f(a)\right).$$

This is how we get to the definition of a regular SLP on an interval $[a,b]$. It is specified by

1. a formally self-adjoint operator
$$L(f)=(rf')'+pf,$$
where r and p are real valued, r, r', and p are continuous, and $r>0$ on $[a,b]$.

2. self-adjoint boundary conditions:
$$B_i(f)=\alpha_i f(a)+\alpha_i' f'(a)+\beta_i f(b)+\beta_i' f'(b)=0, \quad i=1,2.$$
The self adjoint condition further requires that the coefficients $\alpha_i, \alpha_i', \beta_i, \beta_i'$ are such that for all f and g which satisfy these conditions
$$r(\bar{g}f'-\bar{g}'f)\big|_a^b=0.$$

3. a positive, continuous function w on $[a,b]$.

The SLP is to find all solutions to the BVP

$$L(f)+\lambda wf=0, \quad B_i(f)=0, \quad i=1,2.$$

The eigenvalues are all numbers λ for which there exists a corresponding non-zero eigenfunction f so that together they satisfy the above equation, and f satisfies the boundary condition.

The formulation of SLPs is somewhat tedious and probably seems rather strange, cumbersome, and unintuitive. You may feel the same way towards these problems that I feel towards Swedish sidewalks in the winter: frustrated and clumsy! It is not so easy to master that slp, slide, glide walking technique discussed in Figure 6.2. Similarly, when faced with new and complicated appearing mathematics one may initially feel frustrated and clumsy. This is a normal part of the learning process. Although we cannot prove the Adult Spectral Theorem below, we can understand it, and we can compare it to the spectral theorem from linear algebra for hermitian matrices. Let A be an $n\times n$ matrix with possibly complex entries. Then, A acts on the Hilbert space $\mathbb{C}^n$ via matrix multiplication:

$$A: v\in\mathbb{C}^n \to Av\in\mathbb{C}^n.$$

This is a linear operator. We can apply the spectral theorem to A if it satisfies

$$\langle Av,w\rangle=\langle v,Aw\rangle, \quad \forall v,w\in\mathbb{C}^n. \tag{6.1.1}$$

If this is true, then the spectral theorem says that there is a orthonormal basis of $\mathbb{C}^n$ that consists of eigenvectors of A, that is vectors so that

$$Av_k=\lambda_k v_k.$$

Then, with respect to these basis vectors the action of A is given by multiplication by a *diagonal* matrix, in the sense that if we write

$$x\in\mathbb{C}^n=\sum_{k=1}^n \hat{x}_k v_k, \quad \hat{x}_k=\langle x,v_k\rangle,$$

then

$$Ax=\sum_{k=1}^n \lambda_k\hat{x}_k v_k,$$

Figure 6.2: In the Swedish winter, young people become particularly adept at walking over ice. There is a certain technique that involves walking a bit and then sliding on the soles of one's shoes. It looks really awesome; you can see people using this technique to speed walk in the winter wearing ordinary sneakers. These folks are slping, sliding, and gliding towards their destinations! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

and the right side is equal to the product of x with the diagonal matrix that has the eigenvalues of A on the diagonal! You know how matrix multiplication works in general right? Across and down, across and down, lots and lots and lots of calculations. If the matrix is diagonal, however, then matrix multiplication becomes just like multiplying numbers, way fewer calculations and way easier! This is one of the reasons the spectral theorem is awesome. However, a crucial fact is that the matrix A must satisfy (6.1.1).

Here, we are working with infinite dimensional Hilbert spaces, like $\mathcal{L}^2(a,b)$ for a bounded interval (a,b). We also have linear operators that act on our Hilbert space, like $\frac{d^2}{dx^2}$, or more generally, the L in the definition of a regular SLP. There is a spectral theorem for these operators as well! The condition that replaces (6.1.1) for L is *precisely* the rather lengthy and strange-seeming definition of *self-adjoint boundary condition.* This is precisely what guarantees that we have an analogous equation satisfied by L, namely

$$\langle Lf, g\rangle = \langle f, Lg\rangle$$

because

$$\langle LF, g\rangle = \int_a^b (L(f(x))\overline{g(x)}dx.$$

If we compare this to

$$\langle f, Lg\rangle = \int_a^b f(x)\overline{L(g(x)}dx,$$

the necessary and sufficient condition for these to be equal for all f and g in our Hilbert space is precisely the definition of a self-adjoint boundary condition. This is why we have that condition!

Theorem 58 (Spectral Theorem for regular SLPs)**.** *For every regular Sturm-Liouville problem as above, there is an orthonormal basis of L^2_w consisting of eigenfunctions $\{\phi_n\}_{n\in\mathbb{N}}$ with eigenvalues $\{\lambda_n\}_{n\in\mathbb{N}}$. We have*

$$\lim_{n\to\infty} \lambda_n = \infty.$$

Here, L^2_w is the weighted Hilbert space consisting of (the almost everywhere-equivalence classes of measurable) functions on the interval $[a, b]$ which satisfy

$$\int_a^b |f(x)|^2 w(x)dx < \infty,$$

and the scalar product is

$$\langle f, g\rangle_w = \int_a^b f(x)\overline{g(x)}w(x)dx.$$

We are not equipped to prove this fact, but roughly speaking, it is proven by deeper investigation of Hilbert spaces, an area of math known as *functional analysis.* The proof is in many ways similar to the proof of the spectral theorem for linear (matrix) operators on finite dimensional Hilbert spaces, like $\mathbb{C}^n$. As a corollary to this theorem however, we can prove the fact that we have been using thus far without proof.

Corollary 59. *The functions*

$$\left\{e^{inx}\right\}_{n\in\mathbb{Z}}$$

are an orthogonal basis for the Hilbert space $\mathcal{L}^2(-\pi, \pi)$.

Proof: These functions satisfy a regular SLP. This SLP is to find all constants λ and functions f such that

$$f'' + \lambda f = 0,$$

and f is 2π periodic. The operator L is just the operator

$$L(f) = f''.$$

The function $r = 1$, $p = 0$, and the weight is just 1. The boundary conditions are thus:

$$f(-\pi) - f(\pi) = 0, \quad f'(-\pi) - f'(\pi) = 0.$$

We can check that this is 'self-adjoint' by plugging it into the required condition. Assume that some totally arbitrary f and g satisfy this condition, so that $g(-\pi) - g(\pi) = 0$ also. Then

$$(\bar{g}f' - \bar{g}'f)|_{-\pi}^{\pi} = \bar{g}(\pi)f'(\pi) - \bar{g}'(\pi)f(\pi) - \bar{g}(-\pi)f'(-\pi) + \bar{g}'(-\pi)f(-\pi) = 0.$$

By our ODE theory, we can already say that all solutions (up to constant factors) to this problem are

$$f_n(x) = e^{inx}, \quad \lambda_n = n^2.$$

Now, by the Spectral Theorem for regular SLPs, we know that these are an orthogonal basis (they can be normalized if we so desire).

6.2 Useful cute facts about SLPs

Although we cannot prove the Adult Spectral Theorem, we can and indeed shall prove the following useful facts about SLPs.

Theorem 60 (The facts of SLPing and sliding)**.** *Let f and g be eigenfunctions for a regular SLP in an interval $[a, b]$ with weight function $w(x) > 0$. Let λ be the eigenvalue for f and μ the eigenvalue for g. Then:*

1. $\lambda \in \mathbb{R}$ *och* $\mu \in \mathbb{R}$;
2. *If* $\lambda \neq \mu$, *then:*
$$\int_a^b f(x)\overline{g(x)}w(x)dx = 0.$$

Proof: By definition we have $Lf + \lambda wf = 0$. Moreover, L is self-adjoint, which similar to matrices guarantees that
$$\langle Lf, f\rangle = \langle f, Lf\rangle.$$

By being an eigenfunction,
$$Lf = -\lambda wf.$$

So combining these facts:
$$\langle Lf, f\rangle = \langle -\lambda wf, f\rangle = -\lambda\langle wf, f\rangle$$
$$= \langle f, Lf\rangle = \langle f, -\lambda wf\rangle = -\overline{\lambda}\langle f, wf\rangle.$$

Since w is real valued,
$$\langle wf, f\rangle = \int_a^b w(x)f(x)\overline{f(x)}dx = \int_a^b |f(x)|^2 w(x)dx,$$
$$\langle f, wf\rangle = \int_a^b f(x)\overline{w(x)f(x)}dx = \int_a^b |f(x)|^2 w(x)dx.$$

Since $w > 0$ and f is an eigenfunction,
$$\int_a^b |f(x)|^2 w(x)dx > 0.$$

So, the equation
$$-\lambda\langle wf, f\rangle = -\lambda\int_a^b |f(x)|^2 w(x)dx = -\overline{\lambda}\langle f, wf\rangle = -\overline{\lambda}\int_a^b |f(x)|^2 w(x)dx$$

implies
$$\lambda = \overline{\lambda}.$$

For the second part, we use basically the same argument based on self-adjointness:
$$\langle Lf, g\rangle = \langle f, Lg\rangle.$$

By assumption
$$\langle Lf, g\rangle = -\lambda\langle wf, g\rangle = -\lambda\int_a^b w(x)f(x)\overline{g(x)}dx.$$

Similarly,

$$\langle f, Lg\rangle = \langle f, -\mu w g\rangle = -\overline{\mu}\langle f, wg\rangle = -\mu\langle f, wg\rangle = -\mu \int_a^b f(x)\overline{g(x)}w(x)dx,$$

since $\mu \in \mathbb{R}$ and $w(x)$ is real. So we have

$$-\lambda \int_a^b w(x)f(x)\overline{g(x)}dx = -\mu \int_a^b f(x)\overline{g(x)}w(x)dx.$$

If the integral is non-zero, then it forces $\lambda = \mu$ which is false. Thus the integral must be zero.

6.3 Do you need to solve a PDE? An SLP might be the key!

Consider the heat equation on a bounded interval subject to Newton's law of cooling at the endpoints:

$$u_t - u_{xx} = 0, \quad u_x(0,t) = \alpha u(0,t), \quad u_x(l,t) = -\alpha u(l,t), \quad u(x,0) = f(x).$$

Above, we assume that

$$\alpha > 0, \quad f \in \mathcal{L}^2.$$

These boundary conditions are based on Newton's law of cooling: the temperature gradient across the ends is proportional to the temperature difference between the ends and the surrounding medium. It is a homogeneous PDE, so we have good chances of being able to solve it using separation of variables. Thus, we write

$$u(x,t) = X(x)T(t) \implies T'(t)X(x) - X''(x)T(t) = 0 \implies \frac{T'}{T} = \frac{X''}{X}.$$

This means both sides are equal to a constant. Call it λ. We start with the x side, because we have more information about that due to the BCs. Are they self-adjoint BCs? Let's check! In the definition of SLP, we are looking for X to satisfy

$$\frac{X''}{X} = \lambda \iff X'' = \lambda X \iff X'' - \lambda X = 0.$$

OBS! The relationship between the constant we have named λ from the PDE has the *opposite* sign as the corresponding term in an SLP. So, the SLP would look like

$$X'' + \Lambda X = 0 \quad \Lambda = -\lambda.$$

The r and w are both 1 in the definition of SLP, and the p is 0. The $a = 0$ and $b = l$. So, we need to check that if f and g satisfy

$$f'(0) = \alpha f(0), \quad g'(l) = -\alpha g(l)$$

then

$$(\bar{g}f' - \bar{g}'f)\big|_0^l = 0.$$

We plug it in

$$\bar{g}(l)f'(l) - \bar{g}'(l)f(l) - \bar{g}(0)f'(0) + \bar{g}'(0)f(0)$$
$$= -\bar{g}(l)\alpha f(l) + \alpha\bar{g}(l)f(l) - \bar{g}(0)\alpha f(0) + \alpha\bar{g}(0)f(0) = 0.$$

Yes, the BC is a self-adjoint BC. So, the SLP theorem says there exists an $\mathcal{L}^2$ OB of eigenfunctions. What are they? We check the cases.

$$X'' = \lambda X.$$

What if $\lambda = 0$? Then

$$X(x) = ax + b.$$

To get

$$X'(0) = \alpha X(0) \implies a = \alpha b \implies b = \frac{a}{\alpha}.$$

Next,

$$X'(l) = -\alpha X(l) \implies a = -\alpha\left(al + \frac{a}{\alpha}\right) = -a(\alpha l + 1).$$

Presumably $a \neq 0$ because if $a = 0$ the whole solution is just 0. So, we can divide by it and we get

$$\implies 1 = -(\alpha l + 1) \implies \alpha l = -2.$$

Since $l > 0$ and $\alpha > 0$, this is impossible. So, no non-zero solutions for $\lambda = 0$.

Next we try $\lambda > 0$. Then the solution looks like

$$X(x) = ae^{\sqrt{\lambda}x} + be^{-\sqrt{\lambda}x}$$

or equivalently, we can use sinh and cosh, to write

$$X(x) = a\cosh(\sqrt{\lambda}x) + b\sinh(\sqrt{\lambda}x).$$

We try out the BCs. They require

$$X'(0) = \alpha X(0) \iff a\sqrt{\lambda}\sinh(0) + b\sqrt{\lambda}\cosh(0) = \alpha\left(a\cosh(0) + b\sinh(0)\right)$$

$$\iff b\sqrt{\lambda} = \alpha a \implies b = \frac{\alpha a}{\sqrt{\lambda}}.$$

We check out the other BC:

$$X'(l) = -\alpha X(l) \iff a\sqrt{\lambda}\sinh(\sqrt{\lambda}l) + \alpha a\cosh(\sqrt{\lambda}l) = -\alpha\left(a\cosh(\sqrt{\lambda}l) + \frac{\alpha a}{\sqrt{\lambda}}\sinh(\sqrt{\lambda}l)\right).$$

$$\iff a\sqrt{\lambda}\sinh(\sqrt{\lambda}l) + \frac{\alpha^2 a}{\sqrt{\lambda}}\sinh(\sqrt{\lambda}l) = -2\alpha a\cosh(\sqrt{\lambda}l)$$

If $a = 0$ the whole solution is zero, so we presume that is not the case and divide by a. Then this requires

$$\frac{\sinh(\sqrt{\lambda}l)}{\cosh(\sqrt{\lambda}l)} = \frac{-2\alpha}{\sqrt{\lambda} + \alpha^2/\sqrt{\lambda}}.$$

The left side is positive, but the right side is negative. $\lightning$

Thus, we finally try $\lambda < 0$. Then the solution looks like

$$X(x) = a\cos(\sqrt{|\lambda|}x) + b\sin(\sqrt{|\lambda|}x).$$

To get

$$X'(0) = \alpha X(0) \implies b\sqrt{|\lambda|} = \alpha a \implies b = \frac{\alpha a}{\sqrt{|\lambda|}}.$$

Next we need

$$X'(l) = -\alpha X(l)$$

$$\implies -a\sqrt{|\lambda|}\sin(\sqrt{|\lambda|}l) + \frac{\alpha a}{\sqrt{|\lambda|}}\sqrt{|\lambda|}\cos(\sqrt{|\lambda|}l) = -\alpha\left(a\cos(\sqrt{|\lambda|}l) + \frac{\alpha a}{\sqrt{|\lambda|}}\sin(\sqrt{|\lambda|}l)\right).$$

Presumably $a \neq 0$ because if that is the case then the whole solution is 0. So, we may divide by a, and we need

$$2\alpha\cos\sqrt{|\lambda|} = \sin(\sqrt{|\lambda|}l)\left(\sqrt{|\lambda|} - \frac{\alpha^2}{\sqrt{|\lambda|}}\right).$$

This is equivalent to

$$\frac{2\alpha}{\sqrt{|\lambda|} - \frac{\alpha^2}{\sqrt{|\lambda|}}} = \tan(\sqrt{|\lambda|}l)$$

$$\iff \frac{2\alpha\sqrt{|\lambda|}}{|\lambda| - \alpha^2} = \tan(\sqrt{|\lambda|}l).$$

Well, that's pretty weird, but according to the SLP theory, the sequence

$$\{\lambda_n\}_{n\geq 1} \text{ and } \{X_n(x)\}_{n\geq 1}, \quad X_n(x) = a_n\left(\cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x)\right)$$

of eigenvalues and corresponding eigenfunctions is an orthogonal basis of $\mathcal{L}^2$. Here since we are solving a PDE, it is most convenient to leave the coefficients simply as a_n and solve for them according to the initial conditions of the PDE.

The partner functions

$$T_n(t) \text{ satisfy } T_n'(t) = \lambda_n T_n(t) \implies T_n(t) = e^{\lambda_n t}.$$

Here it is good to note that the $\lambda_n < 0$ and tend to $-\infty$ as $n \to \infty$ which follows from the Adult Spectral Theorem, because in the SLP terminology,

$$\Lambda_n = -\lambda_n \to \infty \implies \lambda_n \to -\infty.$$

So, for heat, that is realistic. We build the solution using superposition because the PDE is linear and homogeneous, so

$$u(x,t) = \sum_{n\geq 1} T_n(t)X_n(x).$$

Since we wish this to be equal to the initial data at $t = 0$, we demand

$$u(x,0) = \sum_{n\geq 1} a_n\left(\cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x)\right) = f(x).$$

By the SLP theory, the functions above form an OB, so we can expand our initial data function in terms of this OB. To do this we compute

$$a_n = \frac{\langle f(x), \cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x)\rangle}{||\cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x)||^2},$$

where

$$\langle f(x), \cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x)\rangle = \int_0^l f(x)\overline{(\cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x))}dx,$$

$$||\cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x)||^2 = \int_0^l |\cos(\sqrt{|\lambda_n|}x) + \frac{\alpha}{\sqrt{|\lambda_n|}}\sin(\sqrt{|\lambda_n|}x)|^2 dx.$$

Having completed this example, we are one level better at slping, sliding, and gliding to the solutions of our problems!

Figure 6.3: Skiing also involves a lot of slping, sliding, and gliding. We can imagine that as we practice solving SLPs, we are mathematically becoming more and more adept at slping, sliding, and gliding to the solution! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

6.4 SLP, slide, and glide; the solution has no place to hide!

SLPs may come from solving a PDE, but to avoid overcomplicating things, sometimes one needs to solve an SLP by itself, not necessarily as part of the solution. Here we hone our slp, slide, and glide technique by working out the following example:

$$(xf')' + \lambda x^{-1} f = 0, \quad f(1) = f(b) = 0, \quad b > 1.$$

In this example the function $r(x) = x$, and the function $p(x) = 0$, whilst the weight function $w(x) = x^{-1}$. Let us consider three cases for λ.

Case $\lambda = 0$: If $\lambda = 0$, then the equation becomes

$$xf'' + f' = 0,$$

which we can re-arrange to

$$\frac{f''}{f'} = -\frac{1}{x}.$$

The left side is the derivative of $\log(f')$. So, integrating both sides (saving the constant for later):

$$\log(f') = -\log(x).$$

Exponentiating both sides we get

$$f' = \frac{1}{x} \implies f(x) = A\log(x) + B,$$

for some constants A and B. The boundary conditions demand that

$$f(1) = 0 \implies B = 0.$$

The other boundary condition demands that

$$f(b) = 0 \implies A = 0, \text{ since } b > 1 \text{ so } \log(b) > 0.$$

We are left with the zero function. That is never an eigenfunction. So $\lambda = 0$ is not an eigenvalue for this SLP.

Case $\lambda > 0$**:** If $\lambda > 0$, we observe that the equation we have is something called an Euler equation. (Or we look up the ODE section of Beta and search for this type of ODE, and see that Beta tells us this is an Euler equaiton). Consequently, we look for solutions of the form

$$f(x) = x^{\nu}.$$

The differential equation we wish to solve is:

$$xf'' + f' + \lambda x^{-1} f = 0 \implies x^2 f'' + xf' + \lambda f = 0,$$

so substituting $f(x) = x^{\nu}$, this becomes

$$x^2(\nu)(\nu - 1)x^{\nu-2} + x\nu x^{\nu-1} + \lambda x^{\nu} = 0.$$

This simplifies to:

$$x^{\nu}\left(\nu^2 - \nu + \nu + \lambda\right) = 0 \implies \nu^2 = -\lambda.$$

Since $\lambda > 0$, this means

$$\nu = \pm i\sqrt{\lambda}.$$

So, a basis of solutions is $x^{i\sqrt{|\lambda|}}$ and $x^{-i\sqrt{\lambda}}$. Note that

$$x^{\pm i\sqrt{\lambda}} = e^{\pm i\sqrt{\lambda}\log(x)}.$$

By Euler's formula, an equivalent basis of solutions is

$$\cos(\sqrt{\lambda}\log(x)), \quad \sin(\sqrt{\lambda}\log(x)).$$

Hence in this case our solution is of the form:

$$f(x) = A\cos(\sqrt{\lambda}\log(x)) + B\sin(\sqrt{\lambda}\log(x)).$$

The boundary conditions demand that

$$f(1) = 0 \implies A = 0.$$

The second boundary condition demands that

$$B\sin(\sqrt{\lambda}\log(b)) = 0.$$

Since we do not seek the zero function, we presume that $B \neq 0$ and thus require

$$\sin(\sqrt{\lambda}\log(b)) = 0 \implies \sqrt{\lambda}\log(b) = n\pi, \quad n \in \mathbb{N}.$$

We therefore have countably many eigenfunctions and eigenvalues, which we may index by the natural numbers, writing

$$\lambda_n = \frac{n^2\pi^2}{(\log b)^2}, \quad f_n(x) = \sin\left(\frac{n\pi\log(x)}{\log(b)}\right).$$

Nice.

The last case to consider is **case** $\lambda < 0$**:** We proceed similarly as above and obtain that a basis of solutions is

$$x^{\pm\sqrt{|\lambda|}}.$$

Write our solution as

$$f(x) = Ax^{\sqrt{|\lambda|}} + Bx^{-\sqrt{|\lambda|}}.$$

The boundary conditions demand that:

$$f(1) = 0 \implies A + B = 0 \implies B = -A.$$

The next boundary condition demands that:

$$f(b) = Ab^{\sqrt{|\lambda|}} - Ab^{-\sqrt{|\lambda|}} = 0 \implies A = 0 \text{ or } b^{\sqrt{|\lambda|}} = b^{-\sqrt{|\lambda|}} \implies b^{2\sqrt{|\lambda|}} = 1 \implies \sqrt{|\lambda|} = 0 \lightning.$$

Thus the only way for the boundary conditions to be satisfied is if the eigenfunction is the zero function, but this is not an eigenfunction! Hence no negative λ solutions.

The Adult Spectral Theorem tells us that these rather peculiar functions

$$\{f_n(x)\}_{n \geq 1}$$

are an orthogonal basis for $\mathcal{L}^2_{1/x}(1, b)$. This means that for any $g \in \mathcal{L}^2_{1/x}(1, b)$, we can expand it as a Fourier series with respect to this basis. The coefficients will be

$$\frac{\langle g, f_n \rangle_{1/x}}{||f_n||^2_{1/x}}, \quad \langle g, f_n \rangle_{1/x} = \int_1^b g(x)\overline{f_n(x)}x^{-1}dx, \quad ||f_n||^2_{1/x} = \int_1^b |f_n(x)|^2 x^{-1} dx.$$

If the function we wish to expand is specified, we could compute these integrals.

6.5 Practice makes perfect: slp, slide, and glide

Consider the problem

$$(x^2 f')' + \lambda f = 0, \quad f(1) = f(b) = 0, \quad b > 1.$$

Here we have $r(x) = x^2$ and $w(x) = 1$. The equation is:

$$2xf' + x^2 f'' + \lambda f = 0.$$

We shall consider the three cases for λ.

Case $\lambda = 0$: In this case the equation simplifies to

$$x^2 f'' + 2xf' = 0 \implies \frac{f''}{f'} = -\frac{2}{x} \implies (\log(f'))' = -\frac{2}{x} \implies \log(f') = -2\log x \implies f' = e^{-2\log x} = x^{-2}.$$

So, this gives us a solution of the form

$$f(x) = -A\frac{1}{x} + B.$$

Let us verify the boundary conditions. We require $f(1) = 0$ so this means

$$-A + B = 0 \implies B = A.$$

We also require $f(b) = 0$ so this means

$$-A\frac{1}{b} + B = 0 = \frac{-A}{b} + A \implies \frac{A}{b} = A \implies b = 1 \text{ or } A = 0.$$

So since $b > 1$ the only solution here is the zero function which is not an eigenfunction.

Case $\lambda > 0$: We consider the fact that this is an Euler equation, so we look for solutions of the form $f(x) = x^\nu$. Then the equation looks like:

$$x^2(\nu)(\nu - 1)x^{\nu-2} + 2x(\nu)x^{\nu-1} + \lambda x^\nu = 0 \iff x^\nu \left(\nu^2 - \nu + 2\nu + \lambda\right) = 0$$

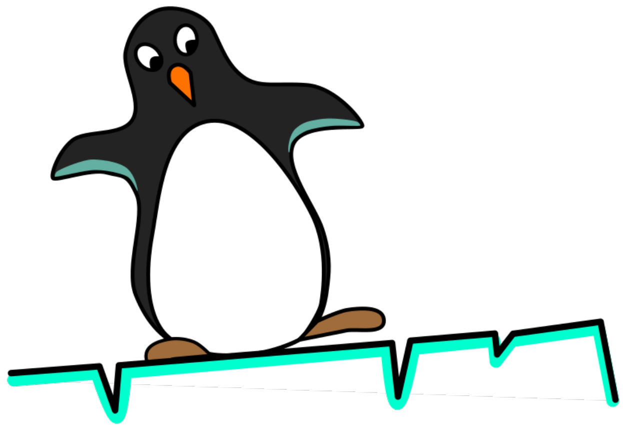


Figure 6.4: I am not skilled at walking on ice and am very far from mastering the Swedish technique of combining walking and sliding. My technique for managing ice is the penguin-strategy. Attempt to keep my center of gravity directly over my feet, and take short little steps. When you start learning to slp, slide, and glide to the solution you might start out with the penguin walk, and that's okay! Practice makes perfect! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

so we need ν to satisfy:

$$\nu^2 + \nu + \lambda = 0.$$

This is a quadratic equation, so we solve it:

$$\nu = -\frac{1}{2} \pm \sqrt{\frac{1}{4} - \lambda}.$$

So, actually the cases $\lambda > 0$ and $\lambda < 0$ really should split up into whether $\lambda = \frac{1}{4}$ or is larger or smaller. If $\lambda = \frac{1}{4}$, then we are only getting one solution this way, $x^{-1/2}$. To get a second solution we multiply by $\log x$.

Exercise 61. *Plug the function $x^{-1/2} \log x$ into the SLP for the value $\lambda = \frac{1}{4}$. Verify that it satisfy the equation.*

Now, let's see if such a function will satisfy the boundary conditions. We need

$$Ax^{-1/2} + Bx^{-1/2} \log(x)\Big|_{x=1} = 0 \implies A = 0.$$

We also need

$$Bb^{-1/2} \log(b) = 0, \quad b > 1 \implies B = 0.$$

So we only get the zero solution in this case.

When $\lambda < \frac{1}{4}$, solutions are of the form

$$Ax^{\nu_+} + Bx^{\nu_-}, \quad \nu_\pm = -\frac{1}{2} \pm \sqrt{\frac{1}{4} - \lambda}.$$

Exercise 62. *Check the boundary conditions. Verify that they are satisfied if and only if $A = B = 0$.*

Finally we consider $\lambda > \frac{1}{4}$. Then we have

$$\nu_\pm = -\frac{1}{2} \pm i\sqrt{\lambda - \frac{1}{4}} \implies f(x) = \frac{A}{\sqrt{x}} x^{i\sqrt{\lambda-1/4}} + \frac{B}{\sqrt{x}} x^{-i\sqrt{\lambda-1/4}}.$$

Using Euler's formula, this is equivalently expressed as

$$\frac{\alpha}{\sqrt{x}}\cos(\sqrt{\lambda-1/4}\log x)+\frac{\beta}{\sqrt{x}}\sin(\sqrt{\lambda-1/4}\log x).$$

Due to the boundary condition at $x=1$ we must have $\alpha=0$. So to obtain the other boundary condition, we need

$$\sin(\sqrt{\lambda-1/4}\log b)=0 \implies \sqrt{\lambda-1/4}\log b=n\pi, \quad n\in\mathbb{N}.$$

Hence

$$\lambda=\lambda_n=\frac{1}{4}+\frac{n^2\pi^2}{(\log b)^2}, \quad f_n(x)=x^{-1/2}\sin\left(\frac{n\pi\log x}{\log b}\right).$$

Note that in general we are not bothering to normalize our eigenfunctions because it is rather tedious and not fundamental to our learning experience in this subject.

6.6 Sturm-Liouville problems in mathematical physics

There is a tradition of referring to mathematical results, like the Cauchy-Schwarz inequality, that were obtained either in a joint work or independently, by two mathematicians, separating their surnames with a dash. Like Cauchy-Schwarz. Similarly, Sturm-Liouville problems are named after the two mathematicians, Jacques Charles Franç ois Sturm and Joseph Liouville. In 1837, they published a joint article titled *Analyse d'un Mémoire sur le développement des fonctions en séries, dont les différents terms sont assujettis à satisfaire à une même équation différentielle linéaire contenant un paramètre variable.* The translation is approximately *Analysis of a memoire on the expansion of functions in series in which the different terms are required to satisfy the same linear differential equation that contains one variable parameter.* I don't know about you, but I love that title! It is so descriptive and basically summarizes the fundamental result of the theory, the Adult Spectral Theorem. We should perhaps rename the theorem: Sturm & Liouville Spectral Theorem. A small detail here, is that I think we should change the dash-between-names-tradition, because in modern times, many people have surnames with a dash in them, like my co-author Susanne Menden-Deuer, and my mathematical brother, Jesse Gell-Redman. I suggest that we use & instead. Like Sturm & Liouville. A small observation is that unlike a typical math publication, in which the authors names are listed alphabetically, here we have Sturm's name first. The reason for this could be that he was working independently on this theory throughout the 1830s before teaming up with Liouville, so we might consider him as having contributed a bit more to the theory. For those interested in the history of the subject, check out [2]!

Sturm & Liouville problems are ubiquitous in mathematical physics. Sturm & Liouville were working on the problem of heat conduction through a metal bar and realized that the technique for solving this particular problem could be generalized and used for solving a large class of PDEs. For example in the equation $(p(x)u')'+(q(x)-\lambda r(x))u=0$, the unknown function u could be either a physical quantity or a quantum mechanical wave function. Indeed, the one-dimensional time-independent Schrödinger equation is a SLP! Another important example in physics is the two-body system equation, that reads

$$(Lu')'+Lu=\frac{1}{L}.$$

This system can be derived from Newtonian mechanics and describes the evolution of the system under torque. Such a system can describe for example planetary movements and be useful if one is interested in traveling through space. As a kid, I imagined that traveling in space is just like driving a car but you can travel in three dimensions not just two, like in Figure 6.5. Later I realized that is totally wrong, because of the influence of gravity due to massive things like planets, stars, and black holes. So it's more like driving a 3D car surrounded by scary flaming and/or frozen giant magnets that are moving in mysterious patterns. If there are only two giant objects nearby, then their trajectories in space could be understood by solving this two-body problem, and therefore extremely useful for planning the path of your spaceship. Although SLPs

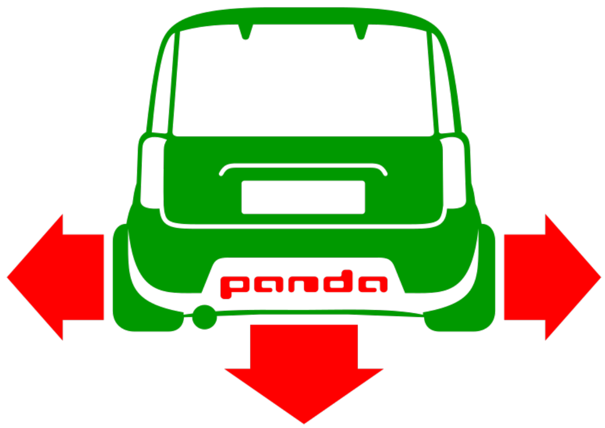


Figure 6.5: Why is space travel so difficult? Isn't it just like driving a car that can move in 3 dimensions instead of 2? Nope! It is a lot more complicated than that! To design and operate a vehicle suitable for space travel requires a *lot* of mathematics! Many of the mathematical equations that need to be understood in order to avoid crashing your space craft can be investigated with the help of... Sturm & Liouville problems! Image license and source: Creative Commons Zero 1.0 Public Domain License `openclipart.org`.

have been studied since the 1830s, there is still quite a lot of research on them in modern times. For example, [14] is a research article on the two-body system and its solutions via SLP methods that was published in 2020!

6.7 Exercises

1. (EO 23) Determine the eigenvalues and eigenfunctions of the SLP:
$$f'' + \lambda f = 0, \quad 0 < x < a,$$
$$f(0) - f'(0) = 0, \quad f(a) + 2f'(a) = 0.$$

2. [4, 3.5.4] Find the eigenvalues and normalized eigenfunctions for the problem
$$f'' + \lambda f = 0, \quad f'(0) = 0, \quad f(\ell) = 0,$$
on $[0, \ell]$.

3. (EO 24) Determine the eigenvalues and eigenfunctions of the SLP:
$$-e^{-4x} \frac{d}{dx}\left(e^{4x} \frac{du}{dx}\right) = \lambda u, \quad 0 < x < 1,$$
$$u(0) = 0, \quad u'(1) = 0.$$

4. [4, 3.5.7] Find all solutions f on $[0, 1]$ and all corresponding λ to the equation:
$$f'' + \lambda f = 0, \quad f(0) = 0, \quad f'(1) = -f(1).$$

5. [4, 4.2.9] Let $k(x)$ be a smooth positive function on $[0, \ell]$. Let $b > 0$ be a constant. Solve the boundary value problem
$$u_t = (kux)_x + f(x,t), \quad u(0,t) = u(\ell,t) = u(x,0) = 0,$$
in terms of the eigenvalues $\{\lambda_n\}$ and eigenfunctions $\{\phi_n\}$ of the SLP:
$$(kf')' + \lambda f = 0, \quad f(0) = f(\ell) = 0.$$

6. [4, 3.5.1] Under what condition on the constants c and c' are the boundary conditions $f(b) = cf(a)$ and $f'(b) = c'f'(a)$ self adjoint for the operator $L(f) = (rf')' + pf$ on $[a, b]$?

7. [4, 3.5.2] Show that the SLP has no negative eigenvalues if $\alpha > 0 > \beta$ and exactly one negative eigenvalue if $\beta > \alpha > 0$ or $0 > \beta > \alpha$, for the SLP

$$f'' + \lambda f = 0, \quad f'(0) = \alpha f(0), \quad f'(\ell) = \beta f(\ell).$$

Verify that the boundary conditions are self-adjoint.

8. [4, 3.5.3] Find the eigenvalues and eigenfunctions of the problem

$$f'' + \lambda f = 0, \quad f(0) = 0, f'(\ell) = 0.$$

Verify that the boundary conditions are self-adjoint.

9. [4, 3.5.5] Find the eigenvalues and eigenfunctions for the problem $f'' + \lambda f = 0$, $f'(0) = \alpha f(0)$, $f'(\ell) = \beta f(\ell)$, assuming that $\alpha = 0$.

10. [4, 3.5.6] Find the eigenvalues and eigenfunctions for the problem $f'' + \lambda f = 0$, $f'(0) = \alpha f(0)$, $f'(\ell) = \beta f(\ell)$, assuming that $\beta = 0$.

11. [4, 3.5.8] Sturm & Liouville are not limited to second order problems. For example, consider $L(f) := \frac{d^4}{dx^2} f(x)$ on the interval $[0, \ell]$. Show that the eigenvalues for the equation $L(f) - \lambda f = 0$, subject to any self-adjoint boundary conditions, are all real, and that the eigenfunctions corresponding to different eigenvalues are orthogonal in $\mathcal{L}^2(0, \ell)$. Show that there is an orthogonal basis of eigenfunctions.

12. [4, 3.5.10] Find the eigenvalues and eigenfunctions for the problem

$$(xf')' + \lambda x^{-1} f = 0, \quad f(1) = f(b) = 0, \quad (b > 1).$$

13. [4, 3.5.11] Find the eigenvalues and eigenfunctions for the problem

$$(x^2 f')' + \lambda f = 0, \quad f(1) = f(b) = 0, \quad (b > 1).$$

14. [4, 4.2.8] Solve

$$u_t = ku_{xx}, \quad u_x(0, t) = 0, \quad u_x(\ell, t) + bu(\ell, t) = 0, \quad u(x, 0) = 100,$$

where $b > 0$ is a fixed constant. What is the physical interpretation?

15. [4, 4.2.10] Assume that the rod on which we are studying the conduction of heat is not insulated along its length. This is described by the equation

$$u_t = ku_{xx} - hu,$$

for a positive constant h.

   (a) Show that u satisfies this equation if and only if $u(x, t) = e^{-ht} v(x, t)$ where v satisfies the standard heat equation.
   (b) Suppose that both ends are insulated, and that the initial temperature is $f(x) = x$. Solve for $u(x, t)$.
   (c) Suppose instead that one end is held at temperature 0, and the other is held at temperature 100, and that the initial temperature is zero. Solve for $u(x, t)$.

Chapter 7

Bessel functions are loads of fun; their zeros describe a vibrating drum!

Why do drums sound the way they do? This is actually a question that even today we do not completely understand. You'll soon find out why...

7.1 Drums and Bessel funs

A heavy metal band would not be complete without drums; see Figure 7.1. How would you describe the sound of a drum? To me, it is a somewhat cloudy sound, not as clear as the sound of a vibrating string. When we solved the mathematical problem that describes a vibrating string, we found the space part of the solutions

$$X_n(x) = \sin(n\pi x/\ell),$$

for a string of length ℓ whose ends are held fixed. The resonant frequencies of the string are the numbers

$$\frac{n^2\pi^2}{\ell^2}.$$

These numbers are obtained by requiring $X_n(0) = X_n(\ell) = 0$. Consequently, these numbers are the squares of the zeros of the sine function. When we solve the analogous problem to describe a vibrating drum, the boundary of the drumhead is fixed to a rigid material. It doesn't move. The interior of the drum vibrates up and down, however. We will see that the resonant frequencies of the drum are squares of zeros of Bessel

Figure 7.1: How would a heavy metal band sound without drums? Can you imagine it? Drums are crucial for the rhythm of the music, for 'keeping the beat!' One of my good friends is a drummer in a punk band, and you can check out their music here on Spotify: https://open.spotify.com/artist/4bXqTVV44kO06WxWJBGoZT Warning for explicit lyrics! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

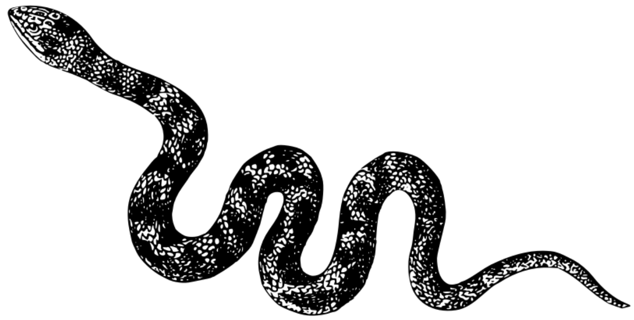


Figure 7.2: The graph of a sine or cosine function looks rather like a slithering snake, as shown here. We can imagine that the sine and cosine got together one night and had baby snakes. Those baby snakes continued reproducing, getting a bit weird due to inbreeding, but nonetheless still resembling the sine and cosine. We call all of these descendants of sine and cosine the Bessel functions, or Bessel funs, for *fun!* After all, we shall see that the zeros of Bessel funs describe the sounds of beating drums! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

functions. Bessel functions bear many similarities to trigonometric functions. Like trigonometric functions, they admit a series expansion, and this expansion looks somewhat similar to the series expansion of sine and cosine,

$$J_\nu(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\Gamma(n+\nu+1)} \left(\frac{x}{2}\right)^{\nu+2n}.$$

This is the definition of the Bessel function of order ν, written J_ν. The Γ we have already met in Chapter 4, but just as a reminder

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt, \quad s \in \mathbb{C} \text{ with } \operatorname{Re}(s) > 0.$$

We will prove in the exercises of this chapter that

$$\Gamma(n+1) = n! = \prod_{k=1}^{n} k, \quad \forall n \in \mathbb{N}.$$

Consequently, the series expression for $J_\nu(x)$ is similar to that of both sine and cosine, since

$$\sin(x) = \sum_{n\geq 0} \frac{(-1)^n x^{2n+1}}{(2n+1)!}, \quad \cos(x) = \sum_{n\geq 0} \frac{(-1)^n x^{2n}}{(2n)!}.$$

The sum alternates sign, due to the changing powers of (-1), and the powers of x in the sum increase by 2, and the denominator has factorials. So, one should think of Bessel functions as being like an infinite family of trigonometric functions, because there are not just two of them, but uncountably many of them, because the order ν can be any complex (or real) number! Perhaps we can think of it like this: sine and cosine are two snakes, because their graphs look snake-like. They got together and had baby snakes, and those snakes continued to multiply and unfortunately inbreed caused them to get kind of weird, a bit different from the original sine and cosine. Some of the snakes live in the trees, others live in the ocean, and there is a whole wealth of different snakes (Bessel functions) found all over the world, but they all are somewhat similar to their ancestors, the sine and cosine. Let's give them a nickname: rather than calling them Bessel functions, let's call them *Bessel funs,* and check out Figure 7.2.

7.1.1 The mathematics of a vibrating drum

We shall solve the initial value problem for a vibrating drum. We begin by mathematicizing the drumhead as a circular membrane. Since it is a drumhead, the boundary is attached to the rest of the drum, so the boundary does not vibrate, it remains fixed. We think of the drumhead as being instantaneously still at the moment when we hit it. Consequently, the height on the drum at a point $z = (x, y)$ and time t satisfies:

$$u_{tt} - u_{xx} - u_{yy} = 0, \quad x^2 + y^2 \leq L^2, \quad \begin{cases} u(x, y, t) = 0 & (x, y) \text{ on the boundary} \\ u_t(x, y, 0) = 0 & \\ u(x, y, 0) = f(x, y) & \end{cases}.$$

To solve this problem, we see that it is homogeneous, and it is also occurring in a bounded region of the plane. So we see if we can use separation of variables. For this we first separate the time and space variables. So our equation is

$$T''(t)S(x, y) - S_{xx}(x, y)T - S_{yy}(x, y)T = 0.$$

We divide everything by TS, move things around, and get

$$\frac{T''}{T} = \frac{S_{xx} + S_{yy}}{S}.$$

Since each side depends on a different variable, we have the equation

$$\frac{S_{xx} + S_{yy}}{S} = \lambda = \frac{T''}{T}.$$

Which side to solve first? Remember our mantra of TIDGLAS: the initial data goes last. So we solve for the space part of the solution first, not the time part. Consequently we seek a solution to:

$$S_{xx} + S_{yy} = \lambda S.$$

Expressing the boundary using x and y it is:

$$x^2 + y^2 = L^2.$$

This is not something of the form "variable equals value." It is more complicated. The reason is because the natural coordinate system for a disk is *not* the square Cartesian coordinates. The natural coordinate system is the polar coordinate system.

Exercise 63. *Show that the differential operator*

$$\partial_{xx} + \partial_{yy}$$

in polar coordinates (r, θ) *becomes*

$$\partial_{rr} + r^{-1}\partial_r + r^{-2}\partial_{\theta\theta}.$$

Hint: use the chain rule!

In terms of polar coordinates the boundary is at $r = L$. This is the type of expression we usually have for a boundary. The function S should vanish at $r = L$. Moreover, we are on a disk. So, the function S at θ and $\theta + 2k\pi$ should be the same for all $k \in \mathbb{Z}$. Let us separate variables, writing $S = R(r)\Theta(\theta)$. Then our equation becomes

$$R''\Theta + r^{-1}R'\Theta + r^{-2}\Theta''R = \lambda R\Theta, \quad R(L) = 0, \quad \Theta(\theta + 2k\pi) = \Theta(\theta).$$

Let's get the different variables cordoned off to different sides of the equation. So, we first divide by $R\Theta$:

$$\frac{R''}{R} + r^{-1}\frac{R'}{R} + r^{-2}\frac{\Theta''}{\Theta} = \lambda.$$

Multiply everything by r^2 to liberate the term with Θ from any r dependence:

$$r^2\frac{R''}{R} + r\frac{R'}{R} + \frac{\Theta''}{\Theta} = r^2\lambda \iff r^2\frac{R''}{R} + r\frac{R'}{R} - r^2\lambda = -\frac{\Theta''}{\Theta}.$$

Each side depends on a different variable, so they are both constant. Since we have the lovely periodicity condition for Θ, and its equation is more simple, let us look for its solution first. We have

$$-\frac{\Theta''}{\Theta} = \text{ constant } = \mu, \quad \Theta(\theta + 2k\pi) = \Theta(\theta).$$

So, we are looking for a 2π periodic function which has Θ'' equal to a constant times Θ. The only functions which have this are sines and cosines! Equivalently, we may use complex exponentials. So, we may choose to use

$$\{\sin(nx), \cos(nx)\}_{n \in \mathbb{N}_0}, \text{ or } \{e^{inx}\}_{n \in \mathbb{Z}}.$$

Either of these will do the job. The numbers

$$\mu = \mu_n = -n^2.$$

So, now let us take the value of μ_n and use it to find the partner function R_n. It satisfies

$$r^2 \frac{R_n''}{R_n} + r\frac{R_n'}{R_n} - r^2\lambda = -\frac{\Theta_n''}{\Theta_n} = - - n^2 = n^2.$$

Re-arranging the equation, we get

$$r^2 R_n'' + rR_n' - r^2\lambda R_n - n^2 R_n = 0. \tag{7.1.1}$$

This is quite close to Bessel's equation.

Definition 64. The differential equation

$$x^2 u''(x) + xu'(x) + (x^2 - \alpha^2)u(x) = 0, \quad \alpha \in \mathbb{C}$$

is Bessel's equation. The differential equation

$$x^2 u''(x) + xu'(x) - (x^2 + \alpha^2)u(x) = 0,$$

is the modified Bessel equation.

So, let's try to relate our equation (7.1.1). The main differences are: λ factor attached to r^2 term and different signs. Let us consider first the case in which $\lambda < 0$. Then $-\lambda > 0$. So, let us write

$$R_n(r) = F_n(x), \quad x = r\sqrt{|\lambda|} \implies R_n'(r) = F_n'(x)\sqrt{|\lambda|}.$$

So we also have

$$rR_n'(r) = \frac{x}{\sqrt{|\lambda|}} R_n'(r) = \frac{x}{\sqrt{|\lambda|}} F_n'(x)\sqrt{|\lambda|} = xF_n'(x).$$

Similarly we get

$$r^2 R_n''(r) = x^2 F_n''(x).$$

Moreover, since $\lambda < 0$,

$$-r^2\lambda = x^2.$$

So for the function F_n the differential equation (7.1.1) is

$$x^2 F_n''(x) + xF_n'(x) + x^2 F_n(x) - n^2 F_n(x).$$

This is

$$x^2 F_n''(x) + xF_n'(x) + (x^2 - n^2)F_n(x) = 0.$$

This is Bessel's equation! The solution in this case is given by the function

$$F_n(x) = J_n(x) \implies R_n(r) = J_n(r\sqrt{|\lambda|}).$$

What should $\sqrt{|\lambda|}$ be? This comes from the boundary condition. We need

$$R_n(L) = 0 \implies J_n(L\sqrt{|\lambda|}) = 0 \implies L\sqrt{|\lambda|} \text{ is a number where } J_n \text{ vanishes.}$$

Theorem 65. *The Bessel function J_n has infinitely many zeros along the real axis. We may therefore write $\{z_{n,m}\}_{m\geq 1}$ to indicate the m^{th} positive zero of the Bessel function J_n.*

Consequently, we require

$$L\sqrt{|\lambda|} = z_{n,m} \quad \text{for some } m \geq 1.$$

This shows that (recalling $\lambda < 0$ in this case)

$$\lambda = \lambda_{n,m} = -\frac{z_{n,m}^2}{L^2}.$$

Exercise 66. *Consider the case $\lambda > 0$. Do a similar change of variables to (7.1.1) to show that in this case we obtain the modified Bessel equation:*

$$x^2 F_n''(x) + xF_n'(x) - (x^2 + n^2)F_n(x) = 0.$$

Check the literature to see that the solutions are the modified Bessel functions, I_n and K_n. Verify in the literature that the functions $K_n(x) \to \infty$ when $x \to 0$. So, these do not yield physically viable solutions to the wave equation because there is no reason for our drum to go off to infinity at the center point. Verify that the functions $I_n(x)$ do not have any positive real zeros, so there is no way to obtain the boundary condition $R_n(L) = 0$. Hence, these too can be discarded.

So, with the exercise, we are able to conclude that only the case $\lambda < 0$ yields physically viable solutions. Equipped with this knowledge, we may return to our equation for the time dependent function.

$$\frac{T_{n,m}''}{T_{n,m}} = \lambda_{n,m} = -\frac{z_{n,m}^2}{L^2} \implies T_{n,m}(t) = a_{n,m}\cos(z_{n,m}t/L) + b_{n,m}\sin(z_{n,m}t/L).$$

The coefficients shall be determined by our initial conditions. Using superposition to create a super solution we have

$$u(t, r, \theta) = \sum_{n\geq 0, m\geq 1} (a_{n,m}\cos(z_{n,m}t/L) + b_{n,m}\sin(z_{n,m}t/L))\, J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta)).$$

The time derivative should vanish when $t = 0$, which means that the coefficients

$$b_{n,m} = 0 \quad \forall n, m.$$

The other condition is

$$u(0, r, \theta) = \sum_{n\geq 0, m\geq 1} a_{n,m} J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta)) = f(r, \theta).$$

So, we would like to have a sort of Fourier expansion in terms of these Bessel functions and sines and cosines. We will have a theorem which says that indeed this is true. Thus

$$a_{n,m} = \frac{\langle f, J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta))\rangle}{||J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta))||^2}.$$

Here since we are doing things on a disk and using polar coordinates, our scalar products are:

$$\langle f, J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta))\rangle = \int_0^L \int_0^{2\pi} f(r, \theta)\overline{J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta))}\, r\,dr\,d\theta,$$

and

$$||J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta))||^2 = \int_0^L \int_0^{2\pi} |J_n(z_{m,n}r/L)(\cos(n\theta) + \sin(n\theta))|^2 r\,dr\,d\theta.$$

We therefore have mathematically described the vibrating drum as in Figure 7.3.

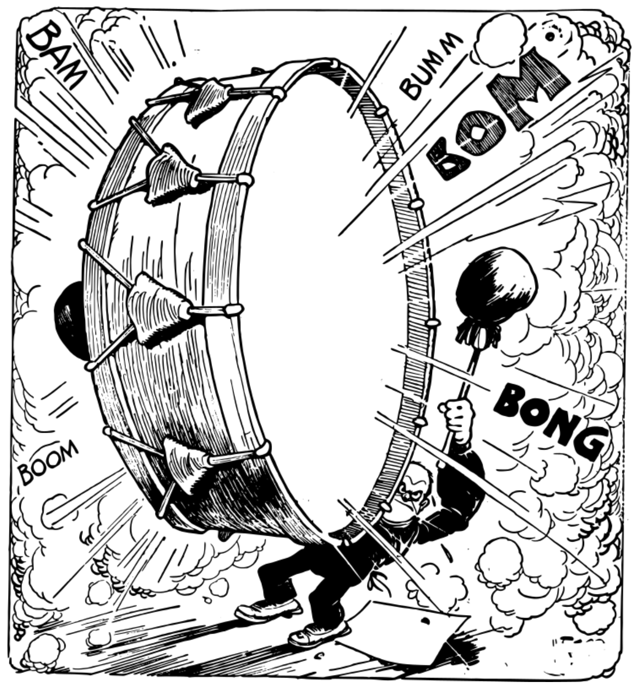


Figure 7.3: There is a rock song by Dire Straights, *Money for Nothing,* https://open.spotify.com/playlist/37i9dQZF1E8zzE1OlcU3xt. Warning for explicit lyrics! Apparently some of the lyrics were from a conversation the singer overheard while in an appliance store. The workers were commenting on musicians in music videos on MTV. The song lyrics include 'I shoulda learned to play the drums... Bangin' on the bongoes like a chimpanzee. That ain't workin' that's the way you do it Get your money for nothin and your chicks for free.' The song lyrics continue to contrast being a musician with working at an appliance store, 'We gotta install microwave ovens, custom kitchen deliveries. We gotta move these refrigerators. We gotta move these colour TVs.' It is a great song and has a cute video. This picture looks rather like it's a chimpanzee banging on the drum. Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

7.2 A series solution to Bessel's equation

Let us write Bessel's equation in this way:

$$x^2 f'' + xf' + (x^2 - \nu^2)f = 0.$$

Assume that f has a series expansion (we will later see that this assumption luckily works out - if it didn't - we'd just have to keep trying other methods). Then we write

$$f(x) = \sum_{j \geq 0} a_j x^{j+b}.$$

Stick it into the ODE:

$$x^2 \sum_{j \geq 0} a_j (j+b)(j+b-1) x^{j+b-2} + x \sum_{j \geq 0} a_j (j+b) x^{j+b-1} + (x^2 - \nu^2) \sum_{j \geq 0} a_j x^{j+b} = 0.$$

Pull the factors of x inside the sum:

$$\sum_{j \geq 0} a_j (j+b)(j+b-1) x^{j+b} + \sum_{j \geq 0} a_j (j+b) x^{j+b} + \sum_{j \geq 0} a_j x^{j+b+2} - \nu^2 a_j x^{j+b} = 0.$$

To make the sum vanish, it will certainly suffice to make all the individual terms in the sum vanish. So, we investigate the powers of x and their coefficients. The lowest power is $x^{0+b} = x^b$. We collect together all the terms that have a factor of x^b, obtaining

$$a_0 \left(b(b-1) + b - \nu^2 \right) x^b.$$

Since we need this to vanish for *all* x, we need

$$a_0 = 0 \text{ or } b^2 - \nu^2 = 0 \implies b = \pm \nu.$$

The next power of x is x^{b+1}. Collecting all the terms that have x^{b+1}, we would like to guarantee that

$$a_1 \left((1+b)(1+b-1) + (1+b) - \nu^2 \right) x^{b+1} = 0.$$

Let's simplify what's in the parentheses, so we need

$$a_1 \left((1+b)^2 - \nu^2 \right) = 0.$$

So, here are our options:

1. Let $b = \nu$, set $a_1 = 0$, and be free to choose a_0 OR
2. Let $(1+b) = \nu$, set $a_0 = 0$, and be free to choose a_1.

If we think about it, the second option is rather like doing the first one for $\nu - 1$ instead of ν. So, the two options are basically equivalent, but the first one is a bit more simple, so that is what we choose to do. We set $b = \nu$, $a_1 = 0$, and we shall choose $a_0 \neq 0$ later.

What happens with the higher terms? Once $j \geq 2$ the term with $a_j x^{j+b+2}$ gets involved. Let's group the terms in the series in a nice way:

$$\sum_{j \geq 0} x^{j+b} a_j \left((j+b)(j+b-1) + (j+b) - \nu^2 \right) + a_j x^{j+b+2} = 0.$$

This is

$$\sum_{j \geq 0} x^{j+b} a_j \left((j+b)^2 - \nu^2 \right) + a_j x^{j+b+2} = 0.$$

We figured out how to make the terms with the powers x^b and x^{b+1} vanish. For the higher powers, the coefficient of

$$x^{j+b+2} \text{ is } a_{j+2}\left((j+2+b)^2-\nu^2\right)+a_j.$$

Therefore, we need

$$a_{j+2}\left((j+2+b)^2-\nu^2\right)=-a_j \implies a_{j+2}=-\frac{a_j}{(j+2+b)^2-\nu^2)}.$$

Recalling that we picked $b=\nu$, this means

$$a_{j+2}=-\frac{a_j}{(j+2+\nu)^2-\nu^2},$$

so we are not dividing by zero which is a good thing. Equivalently, for $j\geq 2$, we have

$$a_j=-\frac{a_{j-2}}{(j+\nu)^2-\nu^2}=-\frac{a_{j-2}}{j^2+2\nu j}=-\frac{a_{j-2}}{j(j+2\nu)}.$$

We therefore see that since we picked $a_1=0$, all of the odd terms are zero. On the other hand, for the even terms, we can figure out what these are using induction. I claim that

$$a_{2k}=\frac{(-1)^k a_0}{2^{2k}k!(1+\nu)(2+\nu)\dots(k+\nu)}.$$

To begin we check the base case which has $k=1$:

$$a_2=-\frac{a_0}{2(2+2\nu)}=-\frac{a_0}{4(1+\nu)}=\frac{(-1)^1a_0}{2^{2(1)}1!(1+\nu)}.$$

So the formula is correct. We next assume that it holds for k and verify using what we computed above that it works for $k+1$. We have for $j=2k+2$,

$$a_{2k+2}=-\frac{a_{2k}}{(2k+2)(2k+2+2\nu)}.$$

We insert the expression for a_{2k} by the induction assumption that the formula holds for k:

$$a_{2k+2}=-\frac{(-1)^k a_0}{(2k+2)(2k+2+2\nu)2^{2k}k!(1+\nu)(2+\nu)\dots(k+\nu)}.$$

We note that

$$(2k+2)(2k+2+2\nu)=4(k+1)(k+1+\nu)=2^2(k+1)(k+1+\nu).$$

So

$$a_{2k+2}=-\frac{(-1)^k a_0}{2^{2(k+1)}(k+1)k!(1+\nu)(2+\nu)\dots(k+\nu)(k+1+\nu)}.$$

Finally we note that

$$(k+1)k!=(k+1)!.$$

So,

$$a_{2k+2}=-\frac{(-1)^k a_0}{2^{2(k+1)}(k+1)!(1+\nu)(2+\nu)\dots(k+\nu)(k+1+\nu)}.$$

This is the formula for $k+1$, so it is indeed correct. Before we proceed, we recall the Γ function

$$\Gamma(s):=\int_0^\infty t^{s-1}e^{-t}dt, \quad s\in\mathbb{C}, \quad \mathrm{Re}(s)>1.$$

Exercise 67. *Use integration by parts to show that*

$$s\Gamma(s) = \Gamma(s+1).$$

Next, show that $\Gamma(1) = 1$. *Use induction to show that* $\Gamma(n+1) = n!$ *for* $n \geq 1$.

Since $\Gamma(1) = 1$, *this is the reason we define*

$$0! := 1.$$

Moreover, viewing Γ *as an extension of the factorial function to real numbers, we can compute silly expressions like*

$$\pi! = \Gamma(\pi+1), \quad e! = \Gamma(e+1), \quad i! = \Gamma(i+1).$$

Use the so-called functional equation $s\Gamma(s) = \Gamma(s+1)$ *to show that* Γ *extends to a meromorphic function whose only poles occur at the points* 0 *and the negative integers.*

So, motivated by the form of the coefficients, the tradition is to choose

$$a_0 = \frac{1}{2^\nu \Gamma(\nu+1)}.$$

Therefore coefficient

$$a_{2k} = \frac{(-1)^k}{2^{2k+\nu} k!(1+\nu)(2+\nu)\ldots(k+\nu)\Gamma(\nu+1)} = \frac{(-1)^k}{2^{2k+\nu} k! \Gamma(k+\nu+1)}.$$

This is because

$$(\nu+1)\Gamma(\nu+1) = \Gamma(\nu+2).$$

Next

$$(\nu+2)\Gamma(\nu+2) = \Gamma(\nu+3).$$

We continue all the way to

$$(\nu+k)\Gamma(\nu+k) = \Gamma(\nu+k+1).$$

We have therefore arrived at the definition of the Bessel function of order ν,

$$J_\nu(x) := \sum_{k \geq 0} \frac{(-1)^k \left(\frac{x}{2}\right)^{2k+\nu}}{k! \Gamma(k+\nu+1)}.$$

For the special case $\nu = n \in \mathbb{N}$, the Bessel function is defined for good reason via

$$J_{-n}(x) = (-1)^n J_n(x).$$

The Weber Bessel function is defined for $\nu \notin \mathbb{N}$ to be

$$Y_\nu(x) = \frac{\cos(\nu\pi) J_\nu(x) - J_{-\nu}(x)}{\sin(\nu\pi)}.$$

The second linearly independent solution to Bessel's equation is then defined for $n \in \mathbb{N}$ to be

$$Y_n(x) := \lim_{\nu \to n} Y_\nu(x),$$

and this is well defined. If you are curious about Bessel functions, there are books by Olver [18], Watson [23], and Lebedev [11] to name a few. I am very fond of Watson's book [23] and use it in my research. What is most important about Y_n is that it blows up when $x \to 0$. That's okay. Since $J_n(x) \to 0$ as $x \to 0$, for $n \geq 1$, this shows that Y_n and J_n are certainly linearly independent! Hence they indeed form a basis of solutions to the Bessel equation.

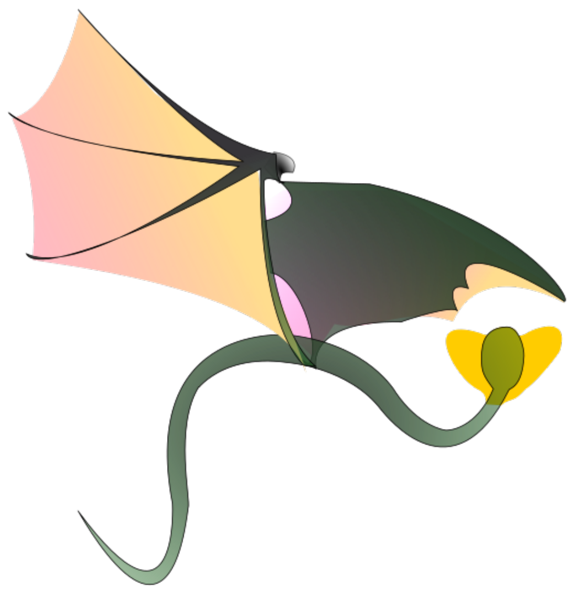


Figure 7.4: The Bessel funs are the descendants of the sine and cosine snakes. They have evolved over the years so that while they still resemble their ancestors, they can live in the sea, in trees, and even fly through the air! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

7.3 Properties of Bessel funs

The Bessel functions evolved from the original sine and cosine, so they bear resemblance in many ways but also have their own distinct features. With our snake analogy, we could imagine that some of the Bessel funs evolved to have wings as shown in Figure 7.4.

Recall the recurrence relationship we have between the original sine and cosine snakes:

$$\sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x).$$

The Bessel funs have inherited something similar to this trait, known as recurrence formulas.

Theorem 68 (Recurrence Formulas)**.** *For all x and ν*

$$(x^{-\nu}J_\nu(x))' = -x^{-\nu}J_{\nu+1}(x)$$
$$(x^{\nu}J_\nu(x))' = x^{\nu}J_{\nu-1}(x)$$
$$xJ_\nu'(x) - \nu J_\nu(x) = -xJ_{\nu+1}(x)$$
$$xJ_\nu'(x) + \nu J_\nu(x) = xJ_{\nu-1}(x)$$
$$xJ_{\nu-1}(x) + xJ_{\nu+1}(x) = 2\nu J_\nu(x)$$
$$J_{\nu-1}(x) - J_{\nu+1}(x) = 2J_\nu'(x)$$

Proof: Can you guess what we do? That's right - use the definition!!!! First,

$$x^{-\nu}J_\nu(x) = \sum_{n\geq 0} \frac{(-1)^n \frac{x^{2n}}{2^{2n+\nu}}}{n!\Gamma(n+\nu+1)}.$$

Take the derivative of the sum termwise. This is totally legitimate because this series converges locally uniformly in $\mathbb{C}$. So, we compute

$$\sum_{n\geq 1} \frac{(-1)^n 2n \frac{x^{2n-1}}{2^{2n+\nu}}}{n!\Gamma(n+\nu+1)} = \sum_{m\geq 0} \frac{(-1)^{m+1} 2(m+1) \frac{x^{2m+1}}{2^{2m+2+\nu}}}{(m+1)!\Gamma(m+2+\nu)}.$$

Above we re-indexed the sum by defining $n = m+1$. Next we do some simplifying around

$$= -\sum_{m\geq 0} \frac{(-1)^m \frac{x^{2m+1}}{2^{2m+1+\nu}}}{m!\Gamma(m+2+\nu)} = -x^{-\nu} \sum_{m\geq 0} \frac{(-1)^m \frac{x^{2m+1+\nu}}{2^{2m+1+\nu}}}{m!\Gamma(m+2+\nu)} = -x^{-\nu}J_{\nu+1}(x).$$

Next we compute similarly the derivative of $x^\nu J_\nu$ is

$$\sum_{n\geq 0} \frac{(-1)^n (2n+2\nu) \frac{x^{2n+2\nu-1}}{2^{2n+\nu}}}{n!\Gamma(n+\nu+1)}.$$

We factor out a 2 to get

$$\sum_{n\geq 0} \frac{(-1)^n (n+\nu) \frac{x^{2n+2\nu-1}}{2^{2n+\nu-1}}}{n!\Gamma(n+\nu+1)}.$$

Note that

$$\Gamma(n+\nu+1) = (n+\nu)\Gamma(n+\nu) \implies \frac{(n+\nu)}{\Gamma(n+\nu+1)} = \frac{1}{\Gamma(n+\nu)}.$$

So, above we have

$$\sum_{n\geq 0} \frac{(-1)^n \frac{x^{2n+2\nu-1}}{2^{2n+\nu-1}}}{n!\Gamma(n+\nu)} = x^\nu J_{\nu-1}(x).$$

To do the third one it is basically expanding out the first one:

$$(x^{-\nu}J_\nu(x))' = -\nu x^{-\nu-1}J_\nu + x^{-\nu}J_\nu' = -x^{-\nu}J_{\nu+1}.$$

Multiply through by $x^{\nu+1}$ to get

$$-\nu J_\nu + xJ_\nu' = -xJ_{\nu+1}.$$

We do similarly in the second formula:

$$\nu x^{\nu-1}J_\nu + x^\nu J_\nu' = x^\nu J_{\nu-1}.$$

Multiply by $x^{-\nu+1}$ to get

$$\nu J_\nu + xJ_\nu' = xJ_{\nu-1}.$$

Next, to get the fifth formula, subtract the third formula from the fourth. Finally, to get the sixth formula, add the third formula to the fourth.

7.3.1 The generating function for the Bessel functions

Euler's equation relates the original sine and cosine snakes to the exponential function,

$$e^{ix} = \cos(x) + i\sin(x), \quad x \in \mathbb{R}, \quad i = \sqrt{-1}.$$

This is one way to obtain the series expansions for the sine and cosine, using the fact that

$$e^z = \sum_{n\geq 0} \frac{z^n}{n!}, \quad \forall z \in \mathbb{C}.$$

The Bessel funs have inherited a similar relationship to the exponential function, but of course since they evolved for infinitely many generations from the original sine and cosine, the simple relationship between the Bessel funs and the exponential function also evolved into something a bit more complicated. This relationship between the exponential function and the Bessel funs is known as *the generating function for the Bessel funs.* We call it that, because this expression is one of the first ways that the Bessel funs were defined, namely, the n^{th} coefficient function for the Laurent expansion of the function $e^{\frac{x}{2}\left(z-\frac{1}{z}\right)}$. Expanding in terms of z, that is a series with z^n for $n \in Z$, the coefficient of z^n will be some function of x. We will prove below that it is none other than the n^{th} Bessel fun! So, one way to obtain the Bessel funs is to solve the Bessel differential equation. An alternative way is to investigate the Laurent series of $e^{\frac{x}{2}\left(z-\frac{1}{z}\right)}$, and determine the coefficients of z^n.

Theorem 69 (Generating function for the Bessel funs)**.** *For all x and for all $z \neq 0$, the Bessel functions, J_n satisfy*

$$\sum_{n=-\infty}^{\infty} J_n(x)z^n = e^{\frac{x}{2}(z-\frac{1}{z})}.$$

Proof: We begin by writing out the familiar Taylor series expansion for the exponential functions

$$e^{xz/2} = \sum_{j\geq 0} \frac{\left(\frac{xz}{2}\right)^j}{j!},$$

and

$$e^{-x/(2z)} = \sum_{k\geq 0} \frac{\left(\frac{-x}{2z}\right)^k}{k!}.$$

These converge beautifully, absolutely and uniformly for z in compact subsets of $\mathbb{C} \setminus \{0\}$. So, since we presume that $z \neq 0$, we can multiply these series and fool around with them to try to make the Bessel functions pop out... Thus, we write

$$e^{xz/2}e^{-x/(2z)} = \sum_{j \geq 0} \frac{\left(\frac{xz}{2}\right)^j}{j!} \sum_{k \geq 0} \frac{\left(\frac{-x}{2z}\right)^k}{k!} = \sum_{j,k \geq 0} (-1)^k \left(\frac{x}{2}\right)^{j+k} \frac{z^{j-k}}{j!k!}. \tag{7.3.1}$$

Here is where the one and only clever idea enters into this proof, but it's rather straightforward to come up with it.

Idea!

We would like a sum with $n = -\infty$ to ∞. So we look around into the above expression on the right, hunting for something which ranges from $-\infty$ to ∞. The only part which does this is $j - k$, because each of j and k range over 0 to ∞. Thus, we keep k as it is, and we let $n = j - k$. Then $j + k = n + 2k$, and $j = n + k$. However, now, we have $j! = (n + k)!$, but this is problematic if $n + k < 0$. There were no negative factorials in our original expression! So, to remedy this, we use the equivalent definition via the Gamma function,

$$j! = \Gamma(j + 1), \quad k! = \Gamma(k + 1).$$

Moreover, we observe that in (7.3.1), $j!$ and $k!$ are for j and k non-negative. We also observe that

$$\frac{1}{\Gamma(m)} = 0, \quad m \in \mathbb{Z}, \quad m \leq 0.$$

Hence, we can write

$$e^{xz/2}e^{-x/(2z)} = \sum_{n=-\infty}^{\infty} \sum_{k=0}^{\infty} (-1)^k \left(\frac{x}{2}\right)^{n+2k} \frac{z^n}{\Gamma(n+k+1)k!}.$$

This is because for all the terms with $n + k + 1 \leq 0$, which would correspond to $(n + k)!$ with $n + k < 0$, those terms ought not to be there, but indeed, the $\frac{1}{\Gamma(n+k+1)}$ causes those terms to vanish!

Now, by definition,

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{x}{2}\right)^{n+2k}}{k!\Gamma(k+n+1)}.$$

Hence, we have indeed see that

$$e^{xz/2}e^{-x/(2z)} = \sum_{n=-\infty}^{\infty} J_n(x)z^n.$$

7.3.2 Integral representation of the Bessel functions

Let $z = e^{i\theta}$ for $\theta \in \mathbb{R}$. Then the theorem on the generating function for the Bessel functions says

$$\sum_{n \in \mathbb{Z}} J_n(x)z^n = e^{\frac{xz}{2} - \frac{x}{2z}}.$$

So, we use the fact that

$$\frac{1}{e^{i\theta}} = e^{-i\theta},$$

together with this formula to see that

$$\sum_{n\in\mathbb{Z}} J_n(x)e^{in\theta} = e^{\frac{x}{2}(e^{i\theta}-e^{-i\theta})}.$$

By Euler's formula,

$$\sum_{n\in\mathbb{Z}} J_n(x)e^{in\theta} = e^{ix\sin\theta} = \cos(x\sin\theta) + i\sin(x\sin\theta).$$

Therefore, the left side is the Fourier expansion of the function on the right. OMG!!! Hence, the Bessel functions are actually *Fourier coefficients* of this function! So,

$$J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ix\sin\theta}e^{-in\theta}d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(x\sin\theta - n\theta) + i\sin(x\sin\theta - n\theta)d\theta.$$

Note that

$$\sin(x\sin(-\theta) - n(-\theta)) = \sin(-x\sin\theta - n(-\theta)) = -\sin(x\sin\theta - n\theta).$$

So the sine part is odd and integrates to zero. We therefore have

$$J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(x\sin\theta - n\theta)d\theta.$$

This formula can be super useful. For example, we see that the Bessel functions have yet another property similar to their ancestors, the sine and cosine. They satisfy $|J_n(\theta)| \le 1 \forall x$.

7.4 Applications to solving PDEs in circular type regions

We shall now see how to generalize our Bessel function techniques to solve problems on pieces of circular sectors. Consider a circular sector of radius ρ and opening angle α. In the eyes of polar coordinates, this is a rectangle, $[0,\rho]\times[0,\alpha]$. That is, this set in $\mathbb{R}^2$ is in polar coordinates

$$\{(r,\theta)\in\mathbb{R}^2 : 0\le r\le\rho, \text{ and } 0\le\theta\le\alpha\}.$$

This is much the same as how we describe a rectangle using *rectangular* coordinates, (x,y).

The homogeneous heat equation is:

$$\partial_t u + \Delta u = 0, \quad \Delta = -\partial_{xx} - \partial_{yy}.$$

The homogeneous wave equation is:

$$u_{tt} + \Delta u = 0.$$

If we have neat and tidy (self-adjoint) boundary conditions, we can use separation of variables. Writing our function as $T(t)S(x,y)$, we obtain the equations:

$$\text{heat equation } T'S + T\Delta S = 0 \iff \frac{\Delta S}{S} = -\frac{T'}{T} = \text{ constant.}$$

$$\text{wave equation } T''S + T\Delta S = 0 \iff \frac{\Delta S}{S} = -\frac{T''}{T} = \text{ constant.}$$

So we see that in both cases we need to solve an equation of the form

$$\Delta S = \lambda S, \quad \lambda \text{ is a constant.}$$

After we solve this, we can then continue with solving both the heat equation and the wave equation.

7.4.1 Dirichlet boundary condition on a circular sector

Let's assume that we have the Dirichlet boundary condition on the boundary of the circular sector. So, we are looking for a function S which is zero on the boundary.

The boundary condition in polar coordinates is:

$$r = \rho, \quad \theta = 0, \quad \theta = \alpha.$$

So, it makes a lot more sense to use these coordinates. To proceed, we need to *write the operator using polar coordinates also!* We have previously computed in an exercise that in polar coordinates, the operator is:

$$\Delta = -\partial_{rr} - r^{-1}\partial_r - r^{-2}\partial_{\theta\theta}.$$

Let us try to solve $\Delta S = \lambda S$ in the circular sector using separation of variables. So, we have

$$R(r) \text{ and } \Theta(\theta).$$

The first one only depends on the r coordinate, whereas the second one only depends on the θ coordinate. Now, our PDE is:

$$-R''(r)\Theta(\theta) - r^{-1}R'(r)\Theta(\theta) - r^{-2}\Theta''(\theta)R(r) = \lambda R(r)\Theta(\theta).$$

First, we multiply everything by r^2, then we divide it all by ΘR to get

$$\frac{-r^2R'' - rR'}{R} - \frac{\Theta''}{\Theta} = \lambda \implies \frac{-r^2R'' - rR'}{R} - \lambda r^2 = \frac{\Theta''}{\Theta}.$$

Since the two sides depend on different variables, they are both constant. It turns out that the Θ side is much easier to deal with, so we look at solving it:

$$\frac{\Theta''}{\Theta} = \mu, \quad \Theta(0) = \Theta(\alpha) = 0.$$

We have solved such an equation a few times before. There are no non-zero solutions for $\mu > 0$. For $\mu < 0$ solutions are, up to constant factors,

$$\Theta_m(\theta) = \sin\left(\frac{m\pi\theta}{\alpha}\right), \quad \mu_m = -\frac{m^2\pi^2}{\alpha^2}.$$

As a consequence, we get the equation for R,

$$\frac{-r^2R'' - rR'}{R} - \lambda r^2 = \mu_m.$$

We multiply this equation by R, obtaining

$$-r^2R'' - rR' - \lambda r^2 R = \mu_m R.$$

This is equivalent to

$$r^2R'' + rR' + (\lambda r^2 + \mu_m)R = 0.$$

We make a small clever change of variables. Let

$$x = \sqrt{\lambda}r, \quad f(x) := R(r), \quad r = \frac{x}{\sqrt{\lambda}}.$$

Then by the chain rule

$$R'(r) = \sqrt{\lambda}f'(x), \quad R''(r) = \lambda f''(x).$$

So, the equation becomes

$$\left(\frac{x^2}{\lambda}\right)\lambda f''(x)+\frac{x}{\sqrt{\lambda}}\sqrt{\lambda}f'(x)+(x^2+\mu_m)f(x)=0.$$

This simplifies, recalling that $\mu_m=-m^2\pi^2/\alpha^2$,

$$x^2f''(x)+xf'(x)+(x^2-m^2\pi^2/\alpha^2)f(x)=0. \tag{7.4.1}$$

This is the definition of Bessel's equation of order $\frac{m\pi}{\alpha}$. Consequently, a solution to this equation is

$$J_{m\pi/\alpha}(x)=J_{m\pi/\alpha}(\sqrt{\lambda}r).$$

To satisfy the boundary condition, we would like

$$J_{m\pi/\alpha}(\sqrt{\lambda}\rho)=0.$$

So, $\sqrt{\lambda}\rho$ should be a point at which this Bessel function vanishes. We have a useful fact about these zeros.

Theorem 70. *The Bessel function $J_{m\pi/\alpha}$ has infinitely many positive zeros which can be indexed as*

$$\{z_{m,k}\}_{k\geq 1},$$

where $z_{m,k}$ is the k^{th} positive zero.

Consequently, we shall have

$$J_{m\pi/\alpha}(z_{m,k}r/\rho),\quad \lambda_{m,k}=\frac{z_{m,k}^2}{\rho^2}.$$

We therefore have the collection of functions

$$S_{m,k}(\theta,r)=\sin(m\pi\theta/\alpha)J_{m\pi/\alpha}\left(\frac{z_{m,k}r}{\rho}\right).$$

Now we may obtain the time part of the solution.

Let us look for a solution to the homogeneous heat equation which satisfies

$$u(r,\theta,0)=f(r,\theta).$$

Then, the partner functions T shall be given by:

$$\frac{\Delta S}{S}=-\frac{T'}{T}=\lambda_{m,k}\implies T_{m,k}(t)=A_{m,k}e^{-\lambda_{m,k}t}.$$

By superposition our full solution is therefore

$$\text{solution to heat equation: } u(r,\theta,t)=\sum_{m,k}A_{m,k}e^{-\lambda_{m,k}t}S_{m,k}(r,\theta).$$

Let us look for a solution to the homogeneous wave equation which satisfies

$$w(r,\theta,0)=g(r,\theta),\quad w_t(r,\theta,0)=0.$$

$$\frac{\Delta S}{S}=-\frac{T''}{T}=\lambda_{m,k}\implies T_{m,k}(t)=a_{m,k}\cos(z_{m,k}t/\rho)+b_{m,k}\sin(z_{m,k}t/\rho).$$

By superposition our full solution is therefore

$$\text{solution to wave eqn: } w(r,\theta,t)=\sum_{m,k}(a_{m,k}\cos(z_{m,k}t/\rho)+b_{m,k}\sin(z_{m,k}t/\rho))S_{m,k}(r,\theta).$$

To determine the coefficients, we shall use the following theorem, that shows that the Bessel functions are an orthogonal basis for $\mathcal{L}^2$ on a sector.

Theorem 71. *The set of functions*

$$\sin(m\pi\theta/\alpha) J_{m\pi/\alpha}\left(\frac{z_{m,k}r}{\rho}\right), \quad k \geq 0, \quad m \geq 1$$

are an orthogonal basis for $\mathcal{L}^2$ on the sector of radius ρ and opening angle α. Above, $z_{m,k}$ is the k^{th} positive zero of $J_{m\pi/\alpha}$.

Consequently, for the heat equation we demand

$$u(r,\theta,0) = \sum_{m,k} A_{m,k} S_{m,k}(r,\theta) = f(r,\theta),$$

which shows us that the coefficients should be

$$A_{m,k} = \frac{\langle f, S_{m,k}\rangle}{||S_{m,k}||^2},$$

where

$$\langle f, S_{m,k}\rangle = \int_0^\alpha \int_0^\rho f(r,\theta)\overline{S_{m,k}(r,\theta)} r dr d\theta,$$

and

$$||S_{m,k}||^2 = \int_0^\alpha \int_0^\rho |S_{m,k}(r,\theta)|^2 r dr d\theta.$$

For the wave equation we demand

$$w(r,\theta,0) = \sum_{m,k} a_{m,k} S_{m,k}(r,\theta) = g(r,\theta) \implies a_{m,k} = \frac{\langle g, S_{m,k}\rangle}{||S_{m,k}||^2}.$$

The second condition tells us what the other coefficients should be:

$$w_t(r,\theta,0) = \sum_{m,k} z_{m,k}/\rho b_{m,k} S_{m,k}(r,\theta) = 0 \implies b_{m,k} = 0 \forall m,k.$$

The following theorem guarantees that we can use the Bessel functions as an orthogonal basis when working in polar coordinates. Note that in polar coordinates, $dxdy$ becomes $rdrd\theta$. So, integration with respect to theta is just the usual integration, but now when we integrate with respect to the radial variable, we integrate rdr, so we are using a weighted $\mathcal{L}^2$ space as discussed in the previous chapter. The first case of the theorem below will correspond to the Dirichlet boundary condition, while the second case will correspond to the Neumann boundary condition.

Theorem 72 (Bessel functions as an orthogonal basis)**.** *Assume that $b > 0$. Let $w(x) = x$.*

1. *Let $\{\lambda_k\}$ be the positive zeros of $J_\nu(x)$, and let $\phi_k(x) := J_\nu(\lambda_k x/b)$. Then $\{\phi_k\}$ is an orthogonal basis for $\mathcal{L}^2_w(0,b)$ and moreover*
$$||\phi_k||_w^2 = \frac{b^2}{2} J_{\nu+1}(\lambda_k)^2.$$

2. *Let*
$$\{z_k\}_{k\geq 1}$$
be the non-negative zeros of $zJ_\nu'(z) = 0$. Let
$$\psi_k(x) = J_\nu(z_k x/b), \quad \nu > 0,$$

and in case $\nu = 0$, define further $\psi_0(x) = 1$. (If $\nu \neq 0$, then this case is omitted). Note that this guarantees $\psi_k'(b) = \frac{z_k}{b} J_\nu'(z_k) = \frac{1}{b}\left(zJ_\nu'(z)\right) = 0$. Then $\{\psi_k\}_{k\geq 1}$ for $\nu \neq 0$ is an orthogonal basis for $\mathcal{L}^2_w(0,b)$. For $\nu = 0$, $\{\psi_k\}_{k\geq 0}$ is an orthogonal basis for $\mathcal{L}^2_w(0,b)$. Moreover

$$||\psi_k||_w^2 = \frac{b^2(z_k^2 - \nu^2)}{2z_k^2} J_\nu(z_k)^2, \quad k \geq 1, \quad ||\psi_0||_w^2 = \frac{b^{2\nu+2}}{2\nu+2}.$$

3. *Let $\{\mu_k\}_{k\geq 1}$ be the positive solutions of $\mu J_\nu'(\mu) + cJ_\nu(\mu) = 0$ for a constant $c > 0$. Then for $\varphi_k(x) = J_\nu(\mu_k x/b)$, $\{\varphi_k\}_{k\geq 1}$ is an orthogonal basis for $\mathcal{L}^2_w(0,b)$.*

7.4.2 Bessel functions with Neumann boundary condition

Let us name the sector

$$\Sigma.$$

We wish to solve the heat equation for a sector with an insulated boundary, corresponding to the Neumann boundary condition

$$u_t + \Delta u = 0, \quad \text{inside } \Sigma,$$

$$u(r,\theta,0) = v(r,\theta) \text{ inside } \Sigma$$

the outward pointing normal derivative of $u = 0$ on the boundary of Σ.

We do the same procedure as before, separating variables, and following the same steps as for the Dirichlet boundary condition. We arrive at the equation for the Θ part:

$$\Theta'' = \mu\Theta, \quad \Theta'(0) = \Theta'(\alpha) = 0.$$

You can do the exercise to show that the only solutions are for $\mu < 0$, and to satisfy the boundary conditions, up to constant multiples

$$\Theta_m(\theta) = \cos(m\pi/\alpha), \quad \mu_m = -\frac{m^2\pi^2}{\alpha^2}, \quad m \geq 0.$$

Then, we again arrive at the Bessel equation of order $m\pi/\alpha$ for the function R. So, we get that

$$R_m(r) = J_{\nu_m}(\sqrt{\lambda} r), \quad \nu_m = m\pi/\alpha.$$

The boundary condition for R_m is that

$$R_m'(\rho) = 0.$$

So, this means we need

$$\sqrt{\lambda} J_{\nu_m}'(\sqrt{\lambda}\rho) = 0.$$

In other words, $\sqrt{\lambda}$ needs to be a solution of the equation

$$xJ_{\nu_m}'(\rho x) = 0.$$

If z_k is a solution to

$$xJ_{\nu_m}'(x) = 0,$$

then

$$z_k J_{\nu_m}'(z_k) = 0 \implies \frac{z_k}{\rho} J_{\nu_m}'(z_k\rho/\rho) = 0.$$

So, to satisfy the boundary condition, we need

$$\sqrt{\lambda} = \frac{z_k}{\rho} \implies \sqrt{\lambda} J_{\nu_m}'(\sqrt{\lambda}\rho) = 0.$$

Really, z_k also depends on m, so that is why we write $z_{m,k}$ to mean the k^{th} positive solution of the equation

$$xJ'_{\nu_m}(x) = 0.$$

Our function

$$R_{m,k}(r) = J_{\nu_m}(z_{m,k}r/\rho).$$

This also shows that

$$\lambda_{m,k} = \frac{z_{m,k}^2}{\rho^2}.$$

Now, we recall the equation for the partner function, T,

$$T'_{m,k}(t) = -\lambda_{m,k}T_{m,k}(t).$$

So, up to constant factors,

$$T_{m,k}(t) = e^{-\lambda_{m,k}t}.$$

To apply the theorem, we note that

$$\nu_m = m\pi/\alpha > 0 \forall m \in \mathbb{N}.$$

Therefore taking $c = 0$ in the theorem, $c \geq -\nu_m$ for all m. The theorem then tells us that the set

$$\{R_{m,k}(r)\}_{k\geq 1} = \{J_{\nu_m}(z_{m,k}r/\rho)\}_{k\geq 1}$$

is an orthogonal basis for $\mathcal{L}^2(0, \rho)$ with respect to integrating against rdr. We also know that the $\Theta_m(\theta)$ functions are an orthogonal basis for $\mathcal{L}^2(0, \alpha)$ with respect to integrating against $d\theta$. Consequently, the entire collection

$$S_{m,k}(r, \theta) = \Theta_m(\theta)R_{m,k}(r)$$

is an orthogonal basis for $\mathcal{L}^2(\Sigma)$. This is because integrating on $\mathcal{L}^2(\Sigma)$ in polar coordinates is integrating

$$\int_\Sigma v(r, \theta)rdrd\theta = \int_0^\rho \int_0^\alpha v(r, \theta)rdrd\theta.$$

So, the theorem says that we can expand the initial data in a Fourier series with respect to the orthogonal basis functions $S_{m,k}$. We therefore write the solution

$$u(r, \theta, t) = \sum_{m,k} \widehat{v_{m,k}} T_{m,k}(t)S_{m,k}(r, \theta),$$

where

$$\widehat{v_{m,k}} = \frac{\int_\Sigma v(r, \theta)S_{m,k}(r)rdrd\theta}{||S_{m,k}||^2}$$

$$= \frac{\int_0^r \int_0^\theta \sin(m\pi\theta/\alpha)J_{m\pi/\alpha}(z_{m,k}r/\rho)v(r, \theta)rdrd\theta}{\int_0^r \int_0^\theta \sin(m\pi\theta/\alpha)^2 J_{m\pi/\alpha}(z_{m,k}r/\rho)^2 rdrd\theta}.$$

7.5 Bessel funs in mathematical physics and music

According to my PhD supervisor, these books about special functions, including [18], [23], [11] to name a few. were written by poorly paid mathematics professors seeking to earn extra money. Many governments allocate a large amount of their budget to their military. Vehicles and weapons all obey the laws of physics, and therefore optimizing their performance involves many mathematical physics calculations as well as solutions of PDEs. Since Bessel functions arise in many of these solutions, governments will pay to have books written that can be used in solving the PDEs needed to make their vehicles and weapons work best. So, doing all these calculations was one way for poor mathematicians to make an extra buck.

Quoting [3],

In Europe, there was an efforescence in the development of music in the seventeenth century. This was complemented by empirical and experimental studies of the physical as well as the aesthetic properties of musical instruments and their interdependence.

Inspired by music, mathematicians sought to solve the equations that describe the vibration of strings as on stringed instruments, and columns of air, as in wind instruments, and the vibration of a circular membrane like a drumhead. Isn't it interesting to note that musical instruments, that were developed independently without communication or collaboration, often share common features? It seems that stringed instruments are found everywhere, so there must be some reason that human ears like the sounds they make. Similarly, drums with round drumheads were developed on every continent that supports human life. For some reason we like the sounds they make. This is beautiful and mysterious because there is no known analytical formula for the zeros of the Bessel functions! There are numerous results like asymptotic formulas, recursive formulas, and approximations for the zeros of the Bessel functions, but there is no simple formula for these zeros like we have for the zeros of $\sin(x)$ and $\cos(x)$. The zeros of the sine are numbers of the form $k\pi$ for integers $k \in \mathbb{Z}$. The zeros of the cosine are numbers of the form $(k + 1/2)\pi$ for integers $k \in \mathbb{Z}$.

Figure 7.5: The desire to understand music lead to amazing scientific discoveries in mathematics and mathematical physics. Contributors include: the Bernoulli family, Count Riccati, G. W. Leibniz (the general notion of the product rule in differentiation is often called Leibniz's rule), E. C. J. von Lommel, M. Mersenne (after whom Mersenne primes are named), J. Saveur (who coined the term acoustics), B. Taylor, C. Huygens (after whom Huygens's principle is named), Euler, W. K. Clifford (after whom Clifford algebras are named), J. le rond d'Alembert (who is credited with obtaining the wave equation in the form we know it today), J. L. L. L. Lagrange (for whom the optimization technique of Lagrange multipliers in multivariable calculus is named as well as the notion of a Lagrangian in the study of PDEs), P. S. Laplace (for whom the Laplace transform is named), M. A. Parseval (that name should sound familiar!), Fourier, C. Maxwell (for whom Maxwell's equations in physics are named), S. D. Poisson (for whom the Poisson relation is named), Bessel and Hankel (both names are often used for generalized Bessel functions). Music inspired a wealth of mathematics and mathematical physics discoveries. The arts are not only intrinsically valuable, but also can lead to great advances in science!

Could it be that what is missing, in order to obtain a concise formula for the zeros of Bessel functions, is a *new* special numerical value? Since there are infinitely many Bessel functions, J_ν of order ν, we might need infinitely many special numerical values, like a family of π_ν, where the Bessel function of order ν has a closed expression in terms of the numerical value π_ν. This π_ν would probably be a transcendental number, similar to π and e. In other words, it would be neither rational, nor would it be obtained as a solution of a polynomial equation with rational coefficients. Now, I'm probably not the first person to have this idea, so please don't set off and try to find the magical π_ν that is the key to expressing the zeros of J_ν in a simple way, because I am pretty sure people just as smart as we are (and possibly smarter) have already had this idea and tried. Watson [23] wrote hundreds of pages on Bessel functions and did countless calculations with them. If Watson couldn't figure out π_ν, we don't stand a chance. Nonetheless, it is an interesting concept to dream about, the magical π_ν that is the key to describing the zeros of J_ν as well as the mystical explanation for why human ears from all different areas of the Earth enjoy the sounds of stringed instruments, wind instruments, and drums with circular drumheads.

7.6 Exercises

1. (Eö 28) Solve:
$$\begin{cases} u_t - u_{xx} = t\sin(x), \quad 0 < x < 1, \quad t > 0, \\ u(0,t) = u(1,t) = 0 \\ u(x,0) = \sin(2\pi x). \end{cases}$$

2. [4, 5.5.2] A circular cylinder of radius ρ is at the constant temperature A. At time $t = 0$ it is tightly wrapped in a sheath of the same material of thickness δ, thus forming a cylinder of radius $\rho + \delta$. The sheath is initially at temperature B, and its outside surface is maintained at temperature B. If the ends of the new, enlarged cylinder are insulated, find the temperature inside at subsequent times.

3. (EO 30) Solve the problem:
$$u_{xx} + u_{yy} = 0$$
in the region in polar coordinates $0 < \theta < \frac{\pi}{4}$, $1 < r < 2$, with the boundary conditions
$$\begin{cases} u = 0 \text{ for } r = 1, \quad u_r = 0 \text{ for } r = 2 \\ u = 0 \text{ for } \theta = 0, \quad u = r - 1 \text{ for } \theta = \frac{\pi}{4}. \end{cases}$$

4. (EO 52) Solve
$$\begin{cases} u_{xx} + 1 = \frac{1}{4}u_{tt} & 0 < x < 2, \, t > 0 \\ u(0,t) = 0, & u(x,0) = x - x^2, \\ u(2,t) = -2, & u_t(x,0) = 0. \end{cases}$$

5. (EO 53) Solve
$$u_{xx} + u_{yy} = 0, \quad r = \sqrt{x^2 + y^2} < 1, \quad u(r = 1, \theta) = \sin^2\theta + \cos\theta.$$

6. [4, 5.5.4] A cylindrical uranium rod of radius 1 generates heat within itself at a constant rate a (think radioactive material). Its ends are insulated and its circular surface is immersed in a cooling bath at temperature zero. Thus
$$u_t = u_{rr} + r^{-1}u_r + r^{-2}u_{\theta\theta} + a, \quad u(1,t) = 0.$$
First find the steady state temperature $v(r)$ in the rod. Then find the temperature in the rod if its initial temperature is zero.

7. [4, 5.2.4] Demonstrate the identity:
$$\int_0^x sJ_0(s)ds = xJ_1(x), \quad \int_0^x J_1(s)ds = 1 - J_0(x).$$

8. [4, 5.5.1] A cylinder of radius b is initially at the constant temperature A. Find the temperatures in it at subsequent times if its ends are insulated and its circular surface obeys Newton's law of cooling, $u_r + cu = 0$, $(c > 0)$.

9. [4, 5.5.5] Solve the problem

$$u_{rr} + r^{-1}u_r + r^{-2}u_{\theta\theta} + u_{zz} = 0 \text{ in } D = \{(r, \theta, z) : 0 \le r \le b, 0 \le z \le l\}$$
$$u(r, \theta, 0) = 0, \quad u(r, \theta, l) = g(r, \theta), \quad u(b, \theta, z) = 0.$$

10. [4, 5.5.6] Find the steady-state temperature in the cylinder $0 \le r \le 1$, $0 \le z \le 1$ when the circular surface is insulated, the bottom is kept at temperature 0, and the top is kept at temperature $f(r)$.

11. [4, 5.2.11] Show that for all real x,

$$J_0(x)^2 + 2\sum_{n=1}^{\infty} J_n(x)^2 = 1.$$

(Hint: Parsevals equation). Deduce that $|J_0(x)| \le 1$ and $|J_n(x)| \le 2^{-1/2}$ for $n > 0$.

12. (The interlacing theorem) For $\nu \in \mathbb{R}$ prove that between every two positive zeros of J_ν there is a zero of $J_{\nu+1}$. (Hint: Use Rolle's theorem and the recurrence formulas).

13. [4, 5.3.6] Let $f(x) = x^{1/2}J_\nu(x)$. Show that f satisfies

$$f'' + f = (\nu^2 - 1/4)x^{-2}f.$$

14. Show that f from the preceding problem satisfies

$$\int_{2\pi n}^{(2n+1)\pi} (1/4 - \nu^2)x^{-2}f(x)\sin(x)dx = -\left(f((2n+1)\pi) + f(2n\pi)\right), \quad n \in \mathbb{N}.$$

15. For f as in the preceding two exercises, assume now that $-\frac{1}{2} < \nu < \frac{1}{2}$. Show that f must vanish somewhere in the interval $[2n\pi, (2n+1)\pi]$.

16. [4, 5.3.7] Use the preceding exercises to show that J_ν has infinitely many positive zeros when $-\frac{1}{2} < \nu < \frac{1}{2}$. Show that $J_{1/2}(x) = \sqrt{2/(\pi x)}\sin(x)$ and use this to conclude that $J_{1/2}$ also has infinitely many positive zeros. Use this together with the interlacing theorem to show that J_ν has infinitely many positive zeros for all real ν.

17. [4, 5.3.8] Let j_ν denote the smallest positive zero of J_ν. Show that $j_{\nu-1} < j_\nu$ for all $\nu \ge 1$.

18. [4, 5.2.9] Show that for all x,

$$J_0(x) + 2\sum_{n=1}^{\infty} J_{2n}(x) = 1, \quad \sum_{n=1}^{\infty}(2n-1)J_{2n-1}(x) = \frac{x}{2}.$$

19. [4, 5.2.10] Show that for each fixed x,

$$\lim_{n\to\infty} n^k J_n(x) = 0, \quad \forall k.$$

20. [4, 5.2.12] Use the formula

$$J_0(x) = \frac{2}{\pi}\int_0^{\pi/2} \cos(x\sin\theta)d\theta$$

to show that J_0 satisfies Bessel's equation of order zero.

21. [4, 5.2.13] Show that for $n \geq 0$, $n \in \mathbb{Z}$

$$J_{2n}(x) = (-1)^n \frac{2}{\pi} \int_0^{\pi/2} \cos(x\cos\theta)\cos(2n\theta)d\theta,$$

$$J_{2n+1}(x) = (-1)^n \frac{2}{\pi} \int_0^{\pi/2} \sin(x\cos\theta)\cos((2n-1)\theta)d\theta.$$

22. (Poisson's integral for J_ν). Show that if $\mathrm{Re}(\nu) > -\frac{1}{2}$,

$$J_\nu(x) = \frac{x^\nu}{2^\nu \pi^{1/2}\Gamma(\nu+1/2)} \int_{-1}^{1} (1-t^2)^{\nu-1/2} e^{ixt} dt.$$

(Hint: expand e^{ixt} in a Taylor series and integrate termwise. This is okay because the series converges uniformly on compact subsets of $\mathbb{C}$.) Do a search for the so-called 'beta function' and use it to deduce that

$$J_\nu(x) = \frac{x^\nu}{2^\nu \pi^{1/2}\Gamma(\nu+1/2)} \int_{-\pi/2}^{\pi/2} e^{ix\sin\theta}\cos^{2\nu}(\theta)d\theta.$$

23. [4, 5.5.3] A cylindrical core of radius 1 is removed from a block of material whose temperature increases linearly from left to right. Thus if the cylinder occupies the region $x^2 + y^2 \leq 1$, the initial temperature is $ax + b$ for some constants a and b. Find the subsequent temperatures in the core if its ends are completely insulated.

24. Assume the same setup as in the preceding problem, but now assume that the ends are insulated, and the circular surface is maintained at temperature zero. Find the subsequent temperatures in the cylindrical core.

25. [4, 5.5.8] Analyze the vibrations of an elastic solid cylinder occupying the region $0 \leq r \leq 1$, $0 \leq z \leq 1$ in cylindrical coordinates if its top and bottom are held fixed, its circular surface is free, and the initial velocity u_t is zero. That is, find general solutions of

$$u_{tt} = \mathfrak{c}^2(u_{rr} + r^{-1}u_r + r^{-2}u_{\theta\theta} + u_{zz}), \quad u(r,\theta,0,t) = u(r,\theta,1,t) = u_r(1,\theta,z,t) = u_t(r,\theta,z,0) = 0.$$

Chapter 8

There's not just one orthogonal base: the *best* base is found with help from space!

In the previous chapter, we discovered that Bessel functions are orthogonal bases of functions in circular regions and pieces of circles. Their zeros can be used to describe the sound of a vibrating drum, because the resonant frequencies of a vibrating drum with a circular drumhead are equal to the squares of the zeros of the Bessel functions! When solving PDEs in circular type regions, the Bessel functions appear, and we use them as an orthogonal base. This is an example of a more general phenomenon, and our motto for this chapter.

There's not just one orthogonal base: the *best* base is found with help from space!

The spatial region, and its geometry, will lead you to the best orthogonal base to use to solve your PDE. If we think about the historical origins of Bessel functions, we discovered them because we wished to understand the vibrations of a *circular membrane.* So far, we have been working with regions of space that include:

1. bounded intervals (one dimensional),
2. a circle (one dimensional),
3. rectangles (two dimensional),
4. annuli (two dimensional),
5. a disk (two dimensional),
6. cylinders (three dimensional),
7. pieces of disks, like sectors (two dimensional).

We have seen how the process of solving PDEs like the heat and wave equation often leads to a set of functions which comprise an orthogonal basis for $\mathcal{L}^2$ or a weighted $\mathcal{L}^2$ space. These basis functions generally come from separation of variables. When we solve the "space" part of the PDE, we very often end up solving a type of SLP. The easiest examples are:

$$f'' = \lambda f, \quad f(a) = 0 = f(b), \text{ for f defined on the interval, } [a, b]$$

$$f'' = \lambda f, \quad f'(a) = 0 = f'(b), \text{ for f defined on the interval, } [a, b]$$

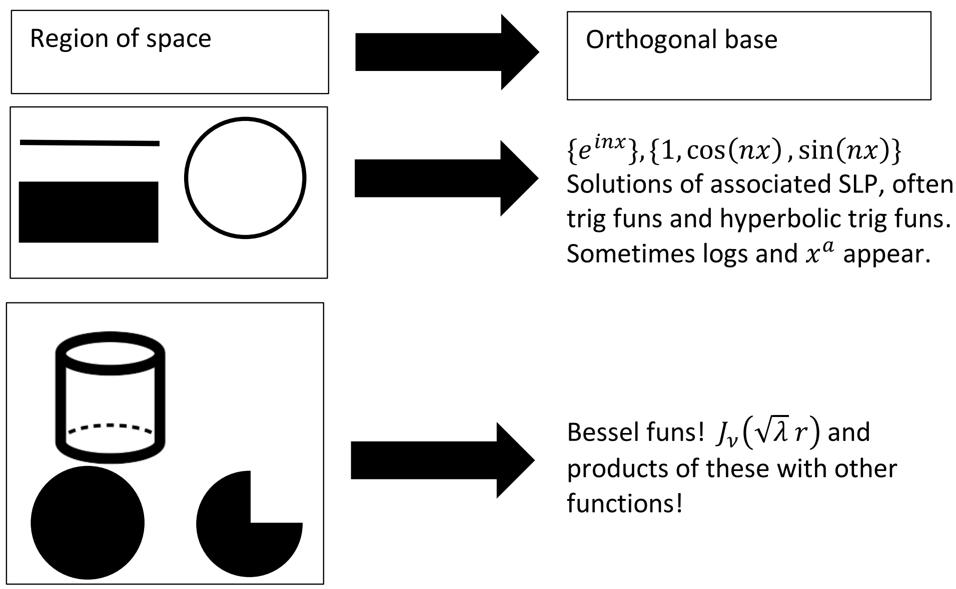


Figure 8.1: The region of space where the physical phenomenon is occurring, like the vibrating string on a guitar, or the vibrating head of a drum, will lead us to a specific choice of orthogonal base for our Hilbert space. The Hilbert space is $\mathcal{L}^2$ of the region where the problem is occurring, and so it is natural that the Hilbert space base depends on space! Now, there are *lots* of bases on any Hilbert space, but if we work carefully paying attention to all the details of our problem, this process will lead us to a particular orthogonal base that can be used to solve our problems, as well as to understand the physical phenomena. Here we will learn about many more orthogonal bases!

Figure 8.2: Here is a picture of Bora Bora, one of the French Polynesian islands. The particular functions that are required to solve a partial differential equation always depend on the *geometry* of where we are solving the equation. We have considered circles, intervals, cylinders, rectangles, and pieces of circles. Imagine what could happen if our geometric context is something as complicated as a real island, like Bora Bora! Image license and source: Creative Commons Zero 1.0 Public Domain License openclipart.org.

$$f'' = \lambda f, \quad f(a) = 0 = f'(b), \text{ for f defined on the interval, } [a, b]$$

$$f'' = \lambda f, \quad f'(a) = 0 = f(b), \text{ for f defined on the interval, } [a, b].$$

A more challenging example comes from solving the heat and wave equations on a circular sector. There, when we did separation of variables, we got the nice type of SLP above for the angular variable (θ), and we got a more complicated SLP for the radial variable. This turned into a Bessel equation. We used the initial data to determine the coefficients in our series expansion, by writing the initial data as a Fourier-Bessel type series. To solve problems in these regions of space, we used a suitable orthogonal base. In this way, the region of space guides us to use the *best* orthogonal base to solve our problem as depicted in Figure 8.1.

In *other* geometric settings, this same process will lead to *other* special functions. We can compare this to navigating different cities. Whereas cities in the US are generally built with a grid-like structure, other cities around the world are built with a different structure, for example Paris, Beijing, and Moscow all have a ring-like structure. Finding our way around those cities is different. Similarly, the orthogonal base that guides our way in a Hilbert space is not always the same. So, now we may be solving PDEs in more exotic geometric settings, like French Polynesia, as shown with a glimpse of Bora Bora in Figure 8.2. Hence, more exotic functions will play the role of the SLP part of the problem. Three such types of functions are *the French polynomials:*

1. The Legendre polynomials arise from using spherical coordinates to solve the wave and heat equations on a three-dimensional sphere.

2. The Hermite polynomials arise from using parabolic coordinates to solve the wave and heat equations in a parabolic shaped region.

3. The Laguerre polynomials arise from the quantum mechanics of the hydrogen atom.

These are all examples of a more general concept: *orthogonal polynomials.*

8.1 General theory of orthogonal polynomials

Constructing an orthogonal basis that is comprised of polynomials is useful not only for solving PDEs but also for obtaining best approximations. We begin this process by proving that we can express arbitrary polynomials in terms of a sequence of polynomials of our choosing.

Proposition 73. *Assume that $\{p_n\}_{n\in\mathbb{N}}$ is a sequence of polynomials such that p_n is of degree n for each n. Assume that $p_0 \neq 0$. Then for each $k \in \mathbb{N}$, any polynomial of degree k is a linear combination of $\{p_j\}_{j=0}^{k}$.*

Proof: The proof is by induction. If q_0 is a polynomial of degree 0, then we may simply write

$$q_0 = \frac{q_0}{p_0} p_0.$$

This is okay because p_0 is degree zero, so it is a constant, and $p_0 \neq 0$, so the coefficient q_0/p_0 is also a constant. Assume that we have verified the proposition for all $0, 1, \ldots k$. We wish to show that it holds for $k+1$. So, let q be a polynomial of degree $k+1$. This means that

$$q(x) = ax^{k+1} + l.o.t. \quad l.o.t. \text{ means lower order terms}$$

has

$$a \neq 0.$$

Moreover, since p_{k+1} is of degree $k+1$ (not of a lower degree), it is of the form

$$p_{k+1} = bx^{k+1} + l.o.t., \quad b \neq 0.$$

So, let us consider

$$q(x) - \frac{a}{b} p_{k+1}(x) = p(x) \text{ which is degree } k.$$

By induction, p is a linear combination of $p_0, \ldots, p_k$. Therefore

$$q(x) = \frac{a}{b} p_{k+1} + \sum_{j=0}^{k} c_j p_j,$$

for some constants $\{c_j\}_{j=0}^{k}$.

Next we prove that polynomials that are orthogonal with respect to $\mathcal{L}^2$ on a bounded interval are an orthogonal base!

Proposition 74. *Let $\{p_k\}_{k=0}^{\infty}$ be a set of polynomials such that each p_k is of degree k, and $p_0 \neq 0$. Moreover, assume that they are $\mathcal{L}^2$ orthogonal on a finite bounded interval $[a,b]$. Then these polynomials comprise an orthogonal basis of $\mathcal{L}^2$ on the interval $[a,b]$.*

Proof: Assume that some $f \in \mathcal{L}^2$ on the interval is orthogonal to all of these polynomials. Therefore by the preceding proposition, f is orthogonal to *all* polynomials. To see this, note that if p is a polynomial of degree n, then there exist numbers $c_0, \ldots, c_n$ such that

$$p = \sum_{j=0}^{n} c_j p_j \implies \langle f, p \rangle = \sum_{j=0}^{n} \overline{c_j} \langle f, p_j \rangle = 0.$$

We shall use the fact that continuous functions are dense in $\mathcal{L}^2$. Therefore given $\varepsilon > 0$, there exists a continuous function, g, such that

$$||f - g|| < \frac{\varepsilon}{2(||f|| + 1)}.$$

Next, we use the Stone-Weierstrass Theorem which says that all continuous functions on bounded intervals can be approximated by polynomials. Therefore, there exists a polynomial p such that

$$||g - p|| < \frac{\varepsilon}{2(||f|| + 1)}.$$

Finally, we compute

$$||f||^2 = \langle f, f\rangle = \langle f - g + g - p + p, f\rangle = \langle f - g, f\rangle + \langle g - p, f\rangle + \langle p, f\rangle$$

$$= \langle f - g, f\rangle + \langle g - p, f\rangle.$$

By the Cauchy-Schwarz inequality,

$$||f||^2 \leq ||f - g||||f|| + ||g - p||||f|| < \frac{||f||\varepsilon}{2(||f|| + 1)} + \frac{||f||\varepsilon}{2(||f|| + 1)} < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows that $||f|| = 0$. Hence by the three equivalent conditions to be an orthogonal basis, we have that the polynomials are an orthogonal basis of $\mathcal{L}^2$ on the interval.

Figure 8.3: Paris has a much more complicated street layout compared to the standard grid-style of cities in the US. Fortunately, if you speak French at least, people are pretty helpful in giving directions! This is an old photo with my parents visiting Paris. I hope you all will be able to visit someday too, and that you'll be able to navigate this place with an orthogonal base!

One of the main applications we have for orthogonal polynomials is using them to obtain best approximations of functions in the Hilbert space $\mathcal{L}^2$ on a bounded interval.

8.1.1 Best approximations

To use orthogonal polynomials to obtain best approximations, we will need the following slight generalization of the best approximation theorem.

Theorem 75 (Best approximation generalization)**.** *Let $\{\phi_n\}_{n\in\mathbb{N}}$ be an orthonormal set in a Hilbert space, H. If $f \in H$,*

$$||f - \sum_{n\in\mathbb{N}} \langle f, \phi_n\rangle \phi_n|| \le ||f - \sum_{n\in\mathbb{N}} c_n\phi_n||, \quad \forall \{c_n\}_{n\in\mathbb{N}} \in \ell^2,$$

and $=$ holds $\iff$ $c_n = \langle f, \phi_n\rangle$ holds $\forall n \in \mathbb{N}$. More generally, let $\{\phi_n\}_{n=0}^N$ be an orthogonal, non-zero set in a Hilbert space H. Then,

$$||f - \sum_{n=0}^N \frac{\langle f, \phi_n\rangle}{||\phi_n||^2}\phi_n|| \le ||f - \sum_{n=0}^N c_n\phi_n||, \quad \forall \{c_n\}_{n=0}^N \in \mathbb{C}^{N+1}.$$

Equality holds if and only if

$$c_n = \frac{\langle f, \phi_n\rangle}{||\phi_n||^2}, \quad n = 0, \ldots, N.$$

Proof: The first part of the statement is just a repetition of the best approximation theorem. We will use that to prove the second part. Define $\psi_n = 0$ for $n > N$. Next define

$$\psi_n = \frac{\phi_n}{||\phi_n||}, \quad n = 0, \ldots, N.$$

We repeat the argument in the proof of the best approximation theorem using $\{\psi_n\}_{n\in\mathbb{N}}$ instead of ϕ_n.

$$||f - \sum_{n\in\mathbb{N}} c_n\psi_n||^2 = ||f - \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n + \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n - \sum_{n\in\mathbb{N}} c_n\psi_n||^2$$

$$= ||f - \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n||^2 + ||\sum_{n\in\mathbb{N}} \hat{f}_n\psi_n - \sum_{n\in\mathbb{N}} c_n\psi_n||^2 + 2\,\mathrm{Re}\langle f - \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n, \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n - \sum_{n\in\mathbb{N}} c_n\psi_n\rangle.$$

The scalar product

$$\langle f - \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n, \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n - \sum_{n\in\mathbb{N}} c_n\psi_n\rangle = \langle f, \sum_{n\in\mathbb{N}} (\hat{f}_n - c_n)\Psi_n\rangle - \sum_{n\in\mathbb{N}} \hat{f}_n\langle\psi_n, \sum_{m\in\mathbb{N}} (\hat{f}_m - c_m)\Psi_n\rangle.$$

By the orthogonality and definition of Ψ_n, and the definition of $\hat{f}_n$,

$$= \sum_{n\in\mathbb{N}} \hat{f}_n\overline{(\hat{f}_n - c_n)} - \sum_{n\in\mathbb{N}} \hat{f}_n \sum_{m\in\mathbb{N}} \overline{(\hat{f}_m - c_m)}\langle\psi_n, \psi_m\rangle$$

$$= \sum_{n\in\mathbb{N}} \hat{f}_n\overline{(\hat{f}_n - c_n)} - \sum_{n\in\mathbb{N}} \hat{f}_n\overline{(\hat{f}_n - c_n)} = 0.$$

Therefore

$$||f - \sum_{n\in\mathbb{N}} c_n\psi_n||^2 = ||f - \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n||^2 + ||\sum_{n\in\mathbb{N}} \hat{f}_n\psi_n - \sum_{n\in\mathbb{N}} c_n\psi_n||^2$$

$$= ||f - \sum_{n=0}^N \hat{f}_n\psi_n||^2 + \sum_{n=0}^N |\hat{f}_n - c_n|^2 \le ||f - \sum_{n=0}^N \hat{f}_n\psi_n||^2,$$

with equality if and only if $c_n = \hat{f}_n$ for all n. Since

$$\sum_{n=0}^N \hat{f}_n\psi_n = \sum_{n=0}^N \frac{\langle f, \phi_n\rangle}{||\phi_n||^2}\phi_n,$$

this completes the proof.

Figure 8.4: This is a photograph I took of the Forbidden City in Beijing. Somewhat similar to Paris, Beijing also has a ring-like layout. The Forbidden City is *forbidden* because historically, no one could enter or leave without the emperor's permission. What orthogonal base is best for navigating the streets of Beijing?

Theorem 75 shows us that if we have a finite orthogonal set of non-zero vectors in a Hilbert space, then for any element of that Hilbert space, the best approximation of f in terms of those vectors is given by

$$\sum_{n=0}^{N} \frac{\langle f, \phi_n \rangle}{||\phi_n||^2} \phi_n.$$

Here is the setup of questions which can be solved using this theory. Either:

1. You are given functions defined on an interval which are $\mathcal{L}^2$ orthogonal on that interval (possibly with respect to a weight function which is also specified). Either you recognize that they orthogonal because you've seen them before (like sines, cosines, from problems you have solved previously) *or* you *compute* that they are $\mathcal{L}^2$ orthogonal on the interval. Then, you are asked to find the numbers $c_0, c_1, \ldots c_N$ so that the $\mathcal{L}^2$ norm, or the weighted $\mathcal{L}^2$ norm of $f - \sum_{k=0}^{N} c_k \phi_k$ is minimized, where the function f is also specified.

2. You are asked to find the polyonomial of at most degree N such that the $\mathcal{L}^2$ norm (or weighted $\mathcal{L}^2$ norm) of $f - p$ where p is a polynomial is minimized.

In the first case, you compute

$$c_k = \frac{\langle f, \phi_k \rangle}{||\phi_k||^2}.$$

In the second case you need to *build up a set of orthogonal or orthonormal polynomials.* Then, you let ϕ_k be defined to be the polynomial of degree k you have built. Proceed the same as in the first case, and your answer shall be

$$\sum_{k=0}^{N} c_k \phi_k.$$

If you don't like the thought of building up a set of orthogonal polynomials, if you are lucky, then it may be possible to suitably modify some of the French polynomials to be orthogonal on the interval under investigation, with respect to the (possibly weighted) $\mathcal{L}^2$ norm. So, we shall proceed to study the French polynomials.

8.2 The Legendre polynomials and applications

The *Legendre polynomials,* are defined to be

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left((x^2-1)^n \right).$$

Your first reaction might be something like *why on earth are they defined in such a bizarre way?* Admittedly, when I first learned about these polynomials, that is what I thought. What did you expect, they are French polynomials! Of course they are not defined in some simple way, mais non, they must be all fancy and shrouded in mystery and intrigue. Actually though, the reason comes from the PDE in which they arise as solving one part of the separation of variables for the heat and wave equations in three dimensions using spherical coordinates. The best way to familiarize ourselves with these polynomials is to start doing calculations with them:

$$(x^2-1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} (x^2)^k = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} x^{2k}.$$

Therefore, if we differentiate n times, only the terms with $k \geq n/2$ survive. Differentiating a term x^{2k} once we get $2kx^{2k-1}$. Differentiating n times gives

$$\frac{d^n}{dx^n}(x^{2k}) = x^{2k-n} \prod_{j=0}^{n-1} (2k-j).$$

If we want to be really persnickety, we prove this by induction. For $n=1$, we get that

$$(x^{2k})' = 2kx^{2k-1}.$$

Which is correct. If we assume the formula is true for n, then differentiating $n+1$ times using the formula for n we get

$$(2k-n)x^{2k-(n+1)} \prod_{j=0}^{n-1} (2k-j) = x^{2k-(n+1)} \prod_{j=0}^{n} (2k-j).$$

See, it is correct. As a result,

$$P_n(x) = \frac{1}{2^n n!} \sum_{k \geq n/2}^{n} (-1)^{n-k} \binom{n}{k} x^{2k-n} \prod_{j=0}^{n-1} (2k-j).$$

So, we see that this is indeed a polynomial of degree n. The Legendre polynomials are very useful because we can obtain a set of orthogonal polynomials on any arbitrary bounded interval by modifying the Legendre polynomials! As a consequence, we can use them to obtain the *best* approximation by polynomials of any arbitrary function in $\mathcal{L}^2$.

Theorem 76. *The Legendre polynomials are orthogonal in $\mathcal{L}^2(-1,1)$, and*

$$||P_n||^2 = \frac{2}{2n+1}.$$

Proof: We first prove the orthogonality. Assume that $n > m$. Then, since they have this constant stuff out front, we compute

$$2^n n! 2^m m! \langle P_n, P_m \rangle = \int_{-1}^{1} \frac{d^n}{dx^n}(x^2-1)^n \frac{d^m}{dx^m}(x^2-1)^m dx.$$

Let us integrate by parts once:

$$= \frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n \frac{d^m}{dx^m}(x^2-1)^m \bigg|_{-1}^{1} - \int_{-1}^{1} \frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n \frac{d^{m+1}}{dx^{m+1}}(x^2-1)^m.$$

Consider the boundary term:

$$\frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n = \frac{d^{n-1}}{dx^{n-1}}(x-1)^n(x+1)^n.$$

This vanishes at $x = \pm 1$, because the polynomial vanishes to order n whereas we only differentiate $n-1$ times. So, we have shown that

$$2^n n! 2^m m! \langle P_n, P_m \rangle = -\int_{-1}^{1} \frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n \frac{d^{m+1}}{dx^{m+1}}(x^2-1)^m.$$

We repeat this $n-1$ more times. We note that for all $j < n$,

$$\frac{d^j}{dx^j}(x^2-1)^n \text{ vanishes at } x = \pm 1.$$

For this reason, all of the boundary terms from integrating by parts vanish. So, we just get

$$(-1)^n \int_{-1}^{1} (x^2-1)^n \frac{d^{m+n}}{dx^{m+n}}(x^2-1)^m dx = (-1)^n \int_{-1}^{1} (x^2-1)^n \frac{d^n}{dx^n}\frac{d^m}{dx^m}(x^2-1)^m dx$$

Remember that $n > m$. We computed that $\frac{d^m}{dx^m}(x^2-1)^m$ is a polynomial of degree m. So, if we differentiate it more than m times we get zero. So, we're integrating zero! Hence it is zero.

For the second part, we need to compute:

$$(x^2-1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k}(x^2)^k = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} x^{2k}.$$

Therefore, if we differentiate n times, only the terms with $k \geq n/2$ survive. Differentiating a term x^{2k} once we get $2kx^{2k-1}$. Differentiating n times gives

$$\frac{d^n}{dx^n}(x^{2k}) = x^{2k-n} \prod_{j=0}^{n-1} (2k-j).$$

If we want to be really persnickety, we prove this by induction. For $n = 1$, we get that

$$(x^{2k})' = 2kx^{2k-1}.$$

Which is correct. If we assume the formula is true for n, then differentiating $n+1$ times using the formula for n we get

$$(2k-n)x^{2k-(n+1)}\prod_{j=0}^{n-1}(2k-j) = x^{2k-(n+1)}\prod_{j=0}^{n}(2k-j).$$

See, it is correct. As a result,

$$P_n(x) = \frac{1}{2^n n!}\sum_{k\geq n/2}^{n}(-1)^{n-k}\binom{n}{k}x^{2k-n}\prod_{j=0}^{n-1}(2k-j).$$

So, we see that this is indeed a polynomial of degree n. With this formula, we can write

$$P_n(x) = \frac{1}{2^n n!}\sum_{k\geq n/2}^{n}(-1)^{n-k}\binom{n}{k}x^{2k-n}\prod_{j=0}^{n-1}(2k-j).$$

Differentiating n times gives us just the term with the highest power of x, so we have

$$\frac{d^n}{dx^n}P_n(x) = \frac{1}{2^n n!}n!\prod_{j=0}^{n-1}(2n-j) = \frac{(2n)!}{2^n n!}.$$

Consequently,

$$\langle P_n, P_n\rangle = (-1)^n\frac{1}{2^n n!}\frac{(2n)!}{2^n n!}\int_{-1}^{1}(x^2-1)^n dx = (-1)^n\frac{2(2n)!}{2^{2n}(n!)^2}\int_{0}^{1}(x^2-1)^n dx$$

$$= (-1)^n\frac{2(2n)!}{2^{2n}(n!)^2}\int_0^1\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}x^{2k}dx$$

$$= (-1)^n\frac{2(2n)!}{2^{2n}(n!)^2}\sum_{k=0}^{n}(-1)^{n-k}\frac{x^{2k+1}}{2k+1}\binom{n}{k}\Bigg|_0^1$$

$$= (-1)^n\frac{2(2n)!}{2^{2n}(n!)^2}\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\frac{1}{2k+1}$$

$$= \frac{2(2n)!}{2^{2n}(n!)^2}\sum_{k=0}^{n}(-1)^{k}\binom{n}{k}\frac{1}{2k+1}.$$

This looks super complicated. Apparently by some miracle of life

$$\int_0^1(1-x^2)^n dx = \frac{\Gamma(n+1)\Gamma(1/2)}{\Gamma(n+3/2)}.$$

Since

$$\langle P_n, P_n\rangle = (-1)^n\frac{2(2n)!}{2^{2n}(n!)^2}\int_0^1(x^2-1)^n dx = \frac{2(2n)!}{2^{2n}(n!)^2}\int_0^1(1-x^2)^n dx,$$

we get

$$\frac{\Gamma(n+1)\Gamma(1/2)2(2n)!}{2^{2n}(n!)^2\Gamma(n+3/2)}.$$

We use the properties of the Γ function together with the fact that $\Gamma(1/2) = \sqrt{\pi}$ to obtain

$$\frac{\sqrt{\pi}2(2n)!}{2^{2n}n!(n+1/2)\Gamma(n+1/2)}.$$

Let us consider

$$2(n+1/2)\Gamma(n+1/2) = (2n+1)\Gamma(n+1/2).$$

Next consider

$$2(n-1/2)\Gamma(n-1/2) = (2n-1)\Gamma(n-1/2).$$

Proceeding this way, the denominator becomes

$$2^n n!(2n+1)(2n-1)\dots 1\sqrt{\pi}.$$

However, now looking at the first part

$$2^n n! = 2n(2n-2)(2n-4)\dots 2.$$

So together we get

$$(2n+1)!\sqrt{\pi}.$$

Hence putting this in the denominator of the expression we had above, we have

$$\frac{\sqrt{\pi}2(2n)!}{(2n+1)!\sqrt{\pi}} = \frac{2}{2n+1}.$$

The results of this chapter then immediately imply

Corollary 77. *The Legendre polynomials are an orthogonal basis for $\mathcal{L}^2$ on the interval $[-1,1]$.*

We used the orthogonal basis $\{e^{inx}\}_{n\in\mathbb{Z}}$ or equivalently $\{1, \sin(nx), \cos(nx)\}_{n\geq 1}$ on the Hilbert space $\mathcal{L}^2(-\pi,\pi)$. By extending functions evenly or oddly from $(0,\pi)$ to $(-\pi,\pi)$ we proved that we could choose an orthogonal base for $(0,\pi)$ that was comprised only of cosines or sines, respectively. Similarly, we can obtain an orthogonal basis for $\mathcal{L}^2(0,1)$ comprised only of even degree or odd degree Legendre polynomials.

Theorem 78. *The even degree Legendre polynomials $\{P_{2n}\}_{n\in\mathbb{N}}$ are an orthogonal basis for $\mathcal{L}^2(0,1)$. The odd degree Legendre polynomials $\{P_{2n+1}\}_{n\in\mathbb{N}}$ are an orthogonal basis for $\mathcal{L}^2(0,1)$.*

Proof: Let f be defined on $[0,1]$. We can extend f to $[-1,1]$ either evenly or oddly. First, assume we have extended f evenly. Then, since $f \in \mathcal{L}^2$ on $[0,1]$,

$$\int_{-1}^{1} |f_e(x)|^2 dx = 2\int_{0}^{1} |f(x)|^2 dx < \infty.$$

Therefore f_e is in $\mathcal{L}^2$ on the interval $[-1,1]$. We have proven that the Legendre polynomials are an orthogonal basis. So, we can expand f_e in a Legendre polynomial series, as

$$\sum_{n\geq 0} \hat{f}_e(n) P_n,$$

where

$$\hat{f}_e(n) = \frac{\langle f_e, P_n\rangle}{||P_n||^2.}$$

By definition,

$$\langle f_e, P_n \rangle = \int_{-1}^{1} f_e(x) P_n(x) dx.$$

Since f_e is even, the product $f_e(x)P_n(x)$ is an *odd* function whenever n is odd. Hence all of the odd coefficients vanish. Moreover,

$$\langle f_e, P_{2n} \rangle = 2 \int_{0}^{1} f(x) P_{2n}(x)) dx.$$

We also have

$$||P_{2n}||^2 = 2 \int_{0}^{1} |P_{2n}(x)|^2 dx.$$

Consequently

$$f = \sum_{n \in \mathbb{N}} \left(\frac{\int_0^1 f(x) P_{2n}(x) dx}{\int_0^1 |P_{2n}(x)|^2 dx} \right) P_{2n}.$$

We can also extend f oddly. This odd extension satisfies

$$\int_{-1}^{1} |f_o(x)|^2 dx = \int_{-1}^{0} |f_o(x)|^2 dx + \int_{0}^{1} |f_o(x)|^2 dx = 2 \int_{0}^{1} |f_o(x)|^2 dx < \infty.$$

So, the odd extension is also in $\mathcal{L}^2$ on the interval $[-1, 1]$. We can expand f_o in a Legendre polynomial series, as

$$\sum_{n \geq 0} \hat{f}_o(n) P_n,$$

where

$$\hat{f}_o(n) = \frac{\langle f_o, P_n \rangle}{||P_n||^2.}$$

By definition,

$$\langle f_o, P_n \rangle = \int_{-1}^{1} f_o(x) P_n(x) dx.$$

Since f_o is odd, the product $f_o(x)P_n(x)$ is an *odd* function whenever n is *even.* Hence all of the even coefficients vanish. Moreover,

$$\langle f_o, P_{2n+1} \rangle = 2 \int_{0}^{1} f(x) P_{2n+1}(x)) dx,$$

because the product of two odd functions is an even function. We also have

$$||P_{2n+1}||^2 = \int_{-1}^{0} |P_{2n+1}(x)|^2 dx + \int_{0}^{1} |P_{2n+1}(x)|^2 dx = 2 \int_{0}^{1} |P_{2n+1}(x)|^2 dx.$$

Consequently

$$f = \sum_{n \in \mathbb{N}} \left(\frac{\int_0^1 f(x) P_{2n+1}(x) dx}{\int_0^1 |P_{2n+1}(x)|^2 dx} \right) P_{2n+1}.$$

Applications of Legendre polynomials to best approximations on bounded integrals

Exercise 79. *Find the polynomial $q(x)$ of at most degree 10 which minimizes the following integral*

$$\int_{-\pi}^{\pi} |q(x) - \sin(x)|^2 dx.$$

To do this exercise, we need different polynomials... If Legendre polynomials are orthogonal on $(-1,1)$, can we somehow use them to create orthogonal polynomials on $(-\pi,\pi)$? Let's think about changing variables. How about setting

$$t = \frac{x}{\pi}.$$

Then,

$$\int_{-\pi}^{\pi} P_n(x/\pi)\overline{P_m(x/\pi)}dx = \int_{-1}^{1} P_n(t)\overline{P_m(t)}\pi dt = \begin{cases} 0 & n \neq m \\ \frac{2\pi}{2n+1} & n = m \end{cases}.$$

Therefore the polynomials

$$P_n(x/\pi)$$

are orthogonal on $x \in (-\pi,\pi)$, and their norms squared on that interval are

$$\frac{2\pi}{2n+1}.$$

The best approximation is therefore the polynomial

$$q(x) = \sum_{n=0}^{10} a_n P_n(x/\pi), \quad a_n := \frac{\int_{-\pi}^{\pi} \sin(x)\overline{P_n(x/\pi)}dx}{\frac{2\pi}{2n+1}}.$$

Exercise 80. *Find the polynomial $p(x)$ of degree at most 100 which minimizes the following integral*

$$\int_0^{10} |e^{x^2} - p(x)|^2 dx.$$

Yikes! Well, let's not panic just yet. The number 100 is even. Hence, we know that the even degree Legendre polynomials are an orthogonal basis for $\mathcal{L}^2(0,1)$. So, we can use the even degree Legendre polynomials if we can just deal with this interval not being $(0,1)$ but being $(0,10)$. To figure this out, let's think about changing variables... As before, think about changing variables,

$$t = x/10,$$

so that

$$\int_0^{10} P_{2n}(x/10)P_{2m}(x/10)dx = \int_0^1 P_{2n}(t)P_{2m}(t)10dt = \begin{cases} 0 & n \neq m \\ \frac{10}{4n+1} & n = m \end{cases}$$

The last calculation we obtained by recalling our calculation

$$\int_{-1}^{1} |P_n(x)|^2 dx = (-1)^n \frac{(2n)!}{(2^n n!)^2} \int_{-1}^{1} (x^2-1)^n dx = \frac{2}{2n+1} \implies \int_0^1 |P_{2n}(x)|^2 dx = \frac{1}{4n+1}.$$

So, the functions $P_{2n}(x/10)$ are an orthogonal basis for $\mathcal{L}^2(0,10)$. Consequently the Best Approximation Theorem says that the best approximation is given by the polynomial

$$p(x) = \sum_{n=0}^{50} c_n P_{2n}(x/10), \quad c_n = \frac{\int_0^{10} e^{x^2}\overline{P_{2n}(x/10)}dx}{\frac{10}{4n+1}}.$$

Exercise 81. *Find the polynomial $p(x)$ of degree at most 99 which minimizes the following integral*

$$\int_0^{10} |e^{x^2} - p(x)|^2 dx.$$

Here, we can recycle our previous solution since 99 is odd, so we can use the odd degree Legendre polynomials in this case to form an orthogonal basis for $\mathcal{L}^2(0,10)$. Our polynomial shall be

$$p(x) = \sum_{n=0}^{49} c_n P_{2n+1}(x/10), \quad c_n = \frac{\int_0^{10} e^{x^2} \overline{P_{2n+1}(x/10)} dx}{\frac{10}{2(2n+1)+1}}.$$

Legendre polynomials for best approximations on arbitrary intervals

Let's consider a best approximation problem on an interval (a,b). First, we find its midpoint,

$$m = \frac{a+b}{2}.$$

Next, we find its length

$$\ell = \frac{b-a}{2}.$$

Then the interval

$$(a,b) = (m-\ell, m+\ell).$$

Since we know about the Legendre polynomials, P_n, on $(-1,1)$ since $x \mapsto \frac{x-m}{\ell} = t$ sends (a,b) to $(-1,1)$,

$$P_n\left(\frac{x-m}{\ell}\right) \quad \text{are orthogonal on } (a,b).$$

In case this is not super obvious, let us compute using the substitution $t = \frac{x-m}{\ell}$,

$$\int_a^b P_n\left(\frac{x-m}{\ell}\right) P_k\left(\frac{x-m}{\ell}\right) dx = \int_{-1}^1 \ell P_n(t) P_k(t) dt = 0 \text{ if } n \neq k.$$

We have simply used substitution in the integral with $t = \frac{x-m}{\ell}$. So, these modified Legendre polynomials are orthogonal on (a,b). Moreover

$$\int_a^b P_n^2\left(\frac{x-m}{\ell}\right) dx = \int_{-1}^1 \ell P_n^2(t) dt = \ell ||P_n||^2 = \frac{2\ell}{2n+1}.$$

So, we simply expand the function f using this version of the Legendre polynomials. Let

$$c_n = \frac{\int_a^b f(x) P_n\left(\frac{x-m}{\ell}\right) dx}{\int_a^b [P_n((x-m)/\ell)]^2 dx}.$$

The best approximation amongst all polynomials of degree at most N is therefore

$$P(x) = \sum_{n=0}^{N} c_n P_n\left(\frac{x-m}{\ell}\right).$$

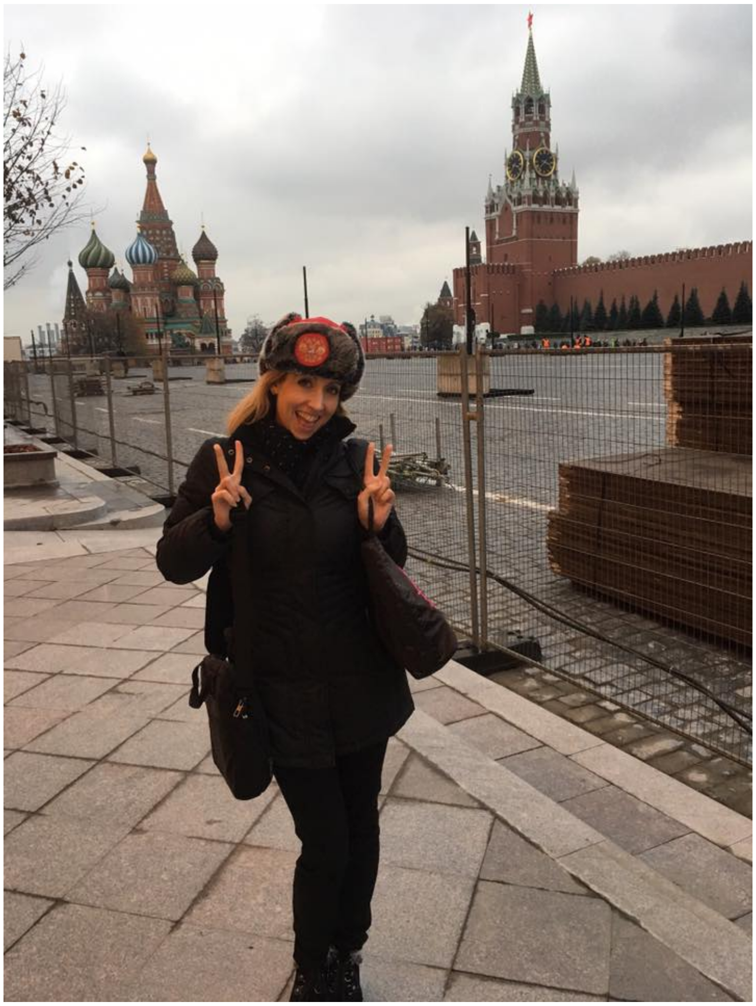


Figure 8.5: Moscow also has a ring-like structure rather than a grid-like structure. I had the pleasure to visit this beautiful city in order to attend a math conference on partial differential equations in honor of the great mathematician B. Y. Sternin. What orthogonal base would you use to navigate Moscow?

8.3 Les polynomes d'hermite

These polynomials shall be a basis for $\mathcal{L}^2(\mathbb{R})$ with respect to the weight function e^{-x^2}. This $\mathcal{L}^2(\mathbb{R})$ is a Hilbert space that is defined analogously to the Hilbert space $\mathcal{L}^2(a,b)$ for an interval (a,b) by letting $a \to -\infty$ and $b \to \infty$.

Definition 82 (The workable definition of $\mathcal{L}^2(\mathbb{R})$)**.** It will suffice for our purposes to treat $\mathcal{L}^2(\mathbb{R})$ as the set of functions on $\mathbb{R}$ which satisfy

$$\int_{\mathbb{R}} |f(x)|^2 dx < \infty.$$

This set of functions, denoted by $\mathcal{L}^2(\mathbb{R})$, is a Hilbert space with the scalar product:

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)\overline{g(x)} d\mu.$$

Hence, by definition, the norm on $\mathcal{L}^2(\mathbb{R})$ is

$$||f||_{\mathcal{L}^2(\mathbb{R})} = \sqrt{\int_{\mathbb{R}} |f(x)|^2 dx}.$$

Definition 83 (The rigorous definition of the Hilbert space $\mathcal{L}^2$)**.** The set

$$\mathcal{L}^2(\mathbb{R}) = \text{ the set of equivalence classes of functions which satisfy:}$$

$$f \text{ is measurable, and } \int_{\mathbb{R}} |f(x)|^2 dx < \infty.$$

The function g belongs to the same equivalence class as f if $g = f$ almost everywhere on $\mathbb{R}$ with respect to the Lebesgue measure on $\mathbb{R}$.

Since everything that would arise in our studies here is measurable, we do not need to worry about the measurability condition nor the equivalence class business, so we can just rely on the workable definition of $\mathcal{L}^2(\mathbb{R})$.

Definition 84. The Hermite polynomials are defined to be

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}.$$

Proposition 85. *The Hermite polynomials are polynomials with the degree of H_n equal to n.*

Proof: The proof is by induction. For $n = 0$, this is certainly true, as $H_0 = 1$. Next, let us assume that

$$\frac{d^n}{dx^n} e^{-x^2} = p_n(x) e^{-x^2},$$

is true for a polynomial, p_n which is of degree n. Then,

$$\frac{d^{n+1}}{dx^{n+1}} e^{-x^2} = \frac{d}{dx}\left(p_n(x)e^{-x^2}\right) = p_n'(x)e^{-x^2} - 2xp_n(x)e^{-x^2} = \left(p_n'(x) - 2xp_n(x)\right)e^{-x^2}.$$

Let

$$p_{n+1} = p_n'(x) - 2xp_n(x).$$

Then we see that since p_n is of degree n, p_{n+1} is of degree $n+1$. Moreover

$$\frac{d^{n+1}}{dx^{n+1}} e^{-x^2} = p_{n+1}(x)e^{-x^2}.$$

So, in fact, the Hermite polynomials satisfy:

$$H_0 = 1, \quad H_{n+1} = -\left(H_n'(x) - 2xH_n(x)\right).$$

Theorem 86. *The Hermite polynomials are orthogonal on* $\mathbb{R}$ *with respect to the weight function* e^{-x^2}. *Moreover, with respect to this weight function* $||H_n||^2 = 2^n n!\sqrt{\pi}$.

Proof: Assume $n > m \geq 0$. We compute

$$\int_{\mathbb{R}} H_n(x) H_m(x) e^{-x^2} dx = \int_{\mathbb{R}} (-1)^n \frac{d^n}{dx^n} e^{-x^2} H_m(x) dx.$$

We use integration by parts n times, noting that the rapid decay of e^{-x^2} kills all boundary terms. We therefore get

$$\int_{\mathbb{R}} e^{-x^2} \frac{d^n}{dx^n} H_m(x) dx = 0.$$

This is because the polyhomial, H_m, is of degree $m < n$. Therefore differentiating it n times results in zero. Finally, for $n = m$, we have by the same integration by parts,

$$\int_{\mathbb{R}} H_n^2(x) e^{-x^2} dx = \int_{\mathbb{R}} e^{-x^2} \frac{d^n}{dx^n} H_n(x) dx.$$

The n^{th} derivative of H_n is just the n^{th} derivative of the highest order term. By our preceding calculation, the highest order term in H_n is

$$(2x)^n.$$

Differentiating n times gives

$$2^n n!.$$

Thus

$$\int_{\mathbb{R}} H_n^2(x) e^{-x^2} dx = 2^n n! \int_{\mathbb{R}} e^{-x^2} dx = 2^n n! \sqrt{\pi}.$$

We may wish to use the following lovely fact, but we shall not prove it.

Theorem 87. *The Hermite polynomials are an orthogonal basis for* $\mathcal{L}^2$ *on* $\mathbb{R}$ *with respect to the weight function* e^{-x^2}.

We shall on the other hand prove the following lovely fact about the Hermite polynomials, known as the generating function for the Hermite polynomials.

Theorem 88 (Generating function for the Hermite polynomials)**.** *For any* $x \in \mathbb{R}$ *and* $z \in \mathbb{C}$, *the Hermite polynomials,*

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2},$$

satisfy

$$\sum_{n=0}^{\infty} H_n(x) \frac{z^n}{n!} = e^{2xz - z^2}.$$

Proof:

Idea!

The key idea with which to begin is to consider instead

$$e^{-(x-z)^2} = e^{-x^2+2xz-z^2}.$$

We consider the Taylor series expansion of this with respect to z, viewing x as a parameter. By definition, the Taylor series expansion for

$$e^{-(x-z)^2} = \sum_{n\geq 0} a_n z^n,$$

where

$$a_n = \frac{1}{n!}\frac{d^n}{dz^n}e^{-(x-z)^2}, \quad \text{evaluated at } z = 0.$$

To compute these coefficients, we use the chain rule, introducing a new variable $u = x - z$. Then,

$$\frac{d}{dz}e^{-(x-z)^2} = -\frac{d}{du}e^{-u^2},$$

and more generally, each time we differentiate, we get a -1 popping out, so

$$\frac{d^n}{dz^n}e^{-(x-z)^2} = (-1)^n\frac{d^n}{du^n}e^{-u^2},$$

Hence, evaluating with $z = 0$, we have

$$a_n = \frac{1}{n!}(-1)^n\frac{d^n}{du^n}e^{-u^2}, \quad \text{evaluated at } u = x.$$

The reason it's evaluated at $u = x$ is because in our original expression we're expanding in a Taylor series around $z = 0$ and $z = 0 \iff u = x$ since $u = x - z$. Now, of course, we have

$$\frac{d^n}{du^n}e^{-u^2}, \quad \text{evaluated at } u = x = \frac{d^n}{dx^n}e^{-x^2}.$$

Hence, we have the Taylor series expansion

$$e^{-(x-z)^2} = e^{-x^2+2xz-z^2} = \sum_{n\geq 0}\frac{z^n}{n!}(-1)^n\frac{d^n}{dx^n}e^{-x^2}.$$

Now, we multiply both sides by e^{x^2} to obtain

$$e^{2xz-z^2} = e^{x^2}\sum_{n\geq 0}\frac{z^n}{n!}(-1)^n\frac{d^n}{dx^n}e^{-x^2}.$$

We can bring e^{x^2} inside because everything converges beautifully. Then, we have

$$e^{2xz-z^2} = \sum_{n\geq 0}\frac{z^n}{n!}e^{x^2}(-1)^n\frac{d^n}{dx^n}e^{-x^2}.$$

Voilà! The definition of the Hermite polynomials is staring us straight in the face! Hence, we have computed

$$e^{2xz-z^2} = \sum_{n\geq 0}\frac{z^n}{n!}H_n(x).$$

8.3.1 Applications to best approximations on $\mathbb{R}$

Exercise 89. *Find the polynomial of at most degree* 40 *which minimizes*

$$\int_{\mathbb{R}} |f(x) - P(x)|^2 e^{-x^2} dx,$$

where f is some function in the weighted $\mathcal{L}^2$ space on $\mathbb{R}$ with weight e^{-x^2}.

We know that the Hermite polynomials are an orthogonal basis for $\mathcal{L}^2$ on $\mathbb{R}$ with the weight function e^{-x^2}. We see that same weight function in the integral. Therefore, we can rely on the theory of the Hermite polynomials! Consequently, we define

$$c_n = \frac{\int_{\mathbb{R}} f(x) H_n(x) e^{-x^2} dx}{||H_n||^2},$$

where

$$||H_n||^2 = \int_{\mathbb{R}} H_n^2(x) e^{-x^2} dx = 2^n n! \sqrt{\pi}.$$

The polynomial we seek is:

$$P(x) = \sum_{n=0}^{40} c_n H_n(x).$$

Some variations on this theme are created by changing the weight function.

Exercise 90. *Find the polynomial of at most degree* 60 *which minimizes*

$$\int_{\mathbb{R}} |f(x) - P(x)|^2 e^{-2x^2} dx.$$

This is not the correct weight function for H_n. However, we can make it so. The correct weight function for $H_n(x)$ is e^{-x^2}. So, if the exponential has $2x^2 = (\sqrt{2}x)^2$, then we should change the variable in H_n as well. We will then have, via the substitution $t = \sqrt{2}x$,

$$\int_{\mathbb{R}} H_n(\sqrt{2}x) H_m(\sqrt{2}x) e^{-2x^2} dx = \int_{\mathbb{R}} H_n(t) H_m(t) e^{-t^2} \frac{dt}{\sqrt{2}} = 0, \quad n \neq m.$$

Moreover, the norm squared is now

$$\int_{\mathbb{R}} H_n^2(t) e^{-t^2} \frac{dt}{\sqrt{2}} = \frac{||H_n||^2}{\sqrt{2}} = \frac{2^n n! \sqrt{\pi}}{\sqrt{2}}.$$

Consequently, the functions $H_n(\sqrt{2}x)$ are an orthogonal basis for $\mathcal{L}^2$ on $\mathbb{R}$ with respect to the weight function e^{-2x^2}. We have computed the norms squared above. The coefficients are therefore

$$c_n = \frac{\int_{\mathbb{R}} f(x) H_n(\sqrt{2}x) e^{-2x^2} dx}{2^n n! \sqrt{\pi}/\sqrt{2}}.$$

The polynomial is

$$P(x) = \sum_{n=0}^{60} c_n H_n(\sqrt{2}x).$$

8.4 The Laguerre polynomials

The Laguerre polynomials come from understanding the quantum mechanics of the hydrogen atom. This will be explored at the end of the chapter.

Definition 91. The Laguerre polynomials,

$$L_n^\alpha(x) = \frac{x^{-\alpha}e^x}{n!}\frac{d^n}{dx^n}(x^{\alpha+n}e^{-x}).$$

We summarize their properties in the following

Theorem 92 (Properties of Laguerre polynomials)**.** *The Laguerre polynomials are an orthogonal basis for $\mathcal{L}^2$ on $(0,\infty)$ with the weight function $x^\alpha e^{-x}$. Their norms squared,*

$$||L_n^\alpha||^2 = \frac{\Gamma(n+\alpha+1)}{n!}.$$

They satisfy the Laguerre equation

$$[x^{\alpha+1}e^{-x}(L_n^\alpha)']' + nx^\alpha e^{-x}L_n^\alpha = 0.$$

For $x > 0$ and $|z| < 1$,

$$\sum_{n=0}^\infty L_n^\alpha(x)z^n = \frac{e^{-xz/(1-z)}}{(1-z)^{\alpha+1}}.$$

The Laguerre polynomials can also be used to obtain best approximations.

Exercise 93. *Find the polynomial of at most degree 7 which minimizes*

$$\int_0^\infty |f(x) - P(x)|^2 x^\alpha e^{-x} dx.$$

Since the Laguerre polynomials are an orthogonal basis for $\mathcal{L}^2(0,\infty)$ with weight function $x^\alpha e^{-x}$, we define

$$c_n = \frac{\int_0^\infty f(x)L_n^\alpha(x)x^\alpha e^{-x}dx}{||L_n^\alpha||^2}.$$

The polynomial we seek is:

$$P(x) = \sum_{n=0}^7 c_n L_n^\alpha(x).$$

8.5 Some functions have Taylor expansions but nearly all functions can be expanded with orthogonal polynomials!

It is natural to wonder what is the difference between a best polynomial approximation and a Taylor expansion? The most important distinction is that we can create a best polynomial approximation for *any function* defined on an interval (a,b) that satisfies

$$\int_a^b |f(x)|^2 dx < \infty.$$

On the other hand, f has a Taylor series expansion, also known as a power series expansion, if and only if f can be differentiated *infinitely many times.* That is a pretty demanding requirement that is not satisfied by a lot of functions. Here is an example

$$f(x) = \sum_{n=0}^{\infty} \alpha^n \cos(\beta^n \pi x), \quad 0 < \alpha < 1, \quad \beta \in \mathbb{N} \text{ is odd, and } \alpha\beta > 1 + \frac{3\pi}{2}.$$

This function is called the Weierstraß function, named after Karl Weierstraß. It is *nowhere differentiable.* It is however continuous, and consequently it is an element of $\mathcal{L}^2$ on any bounded interval. We can therefore expand this function in terms of an orthogonal basis of polynomials. For example, if $\{p_n\}_{n=0}^{\infty}$ are a set of orthogonal polynomials on an interval (a, b), then we have the expansion

$$f(x) = \sum_{n=0}^{\infty} \frac{\langle f, p_n \rangle}{||p_n||^2} p_n(x).$$

On the other hand, this function does not have a Taylor expansion.

For the converse, consider a function $g(x)$ that does have a Taylor expansion. Then it is differentiable infinitely many times, and so as a consequence it is continuous. It is therefore bounded on any bounded interval and hence an element of the Hilbert space $\mathcal{L}^2$ on such an interval. So, we can expand it in terms of the orthogonal polynomials as well:

$$g(x) = \sum_{n=0}^{\infty} \frac{\langle g, p_n \rangle}{||p_n||^2} p_n(x).$$

Thus, we can expand a lot *more* functions using orthogonal polynomials in comparison to using Taylor series. This will be explored further in the exercises!

8.5.1 Best approximations

If we have established that a certain collection of functions

e^{inx}, cos, sin, orthogonal polynomials, Bessel functions, functions obtained by solving SLPs,

are an orthogonal basis on a bounded interval, then we can use these to obtain best approximations. Let us call such functions ϕ_n. Then the best approximation to any f in $\mathcal{L}^2$ of the bounded interval under consideration is its Fourier-ϕ_n expansion, which is

$$\sum \frac{\langle f, \phi_n \rangle}{||\phi_n||^2} \phi_n(x).$$

Recall

$$\langle f, \phi_n \rangle = \int f(y)\overline{\phi_n(y)} w(y) dy, \quad \text{if the weight function is } w(y),$$

and

$$||\phi_n||^2 = \langle \phi_n, \phi_n \rangle.$$

One can also do best approximations using Hermite and Laguerre polynomials on $\mathbb{R}$ and $(0, \infty)$, respectively, with the weight functions e^{-x^2} and $x^\alpha e^{-x}$, respectively. The process is analogous in all cases!

8.6 Orthogonal polynomials in quantum chemistry: the mathematics of the hydrogen atom

In the hydrogen atom, there is an electron and a proton. The proton is about 2,000 times more massive than the electron, so it makes sense to consider the proton as immobile, from the electron's point of view.

The electron is therefore moving in an electrostatic force field with potential $-\epsilon^2/r$, where ϵ is the charge of the proton, and r is the distance from the origin. We assume that the proton is located at the origin.

According to quantum mechanics, when the electron is in a stationary state at energy level E, its wave function u is in $L^2(\mathbb{R}^3)$ and satisfies the Schrödinger equation

$$\frac{\hbar^2}{2m}\Delta u + \frac{\epsilon^2}{r}u + Eu = 0. \tag{8.6.1}$$

Above, $\hbar$ is Planck's constant, and m is the mass of the electron, the Laplace operator Δ is in $\mathbb{R}^3$ equal to

$$\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2,$$

and $r = \sqrt{x^2+y^2+z^2}$ is the distance from the proton located at the origin. Due to the radial symmetry, it is natural to introduce spherical coordinates to try to solve this equation. In this section, we will work through a series of exercises to solve this time-independent Schrödinger equation and understand the energy levels of the hydrogen atom as shown in Figure 8.6.

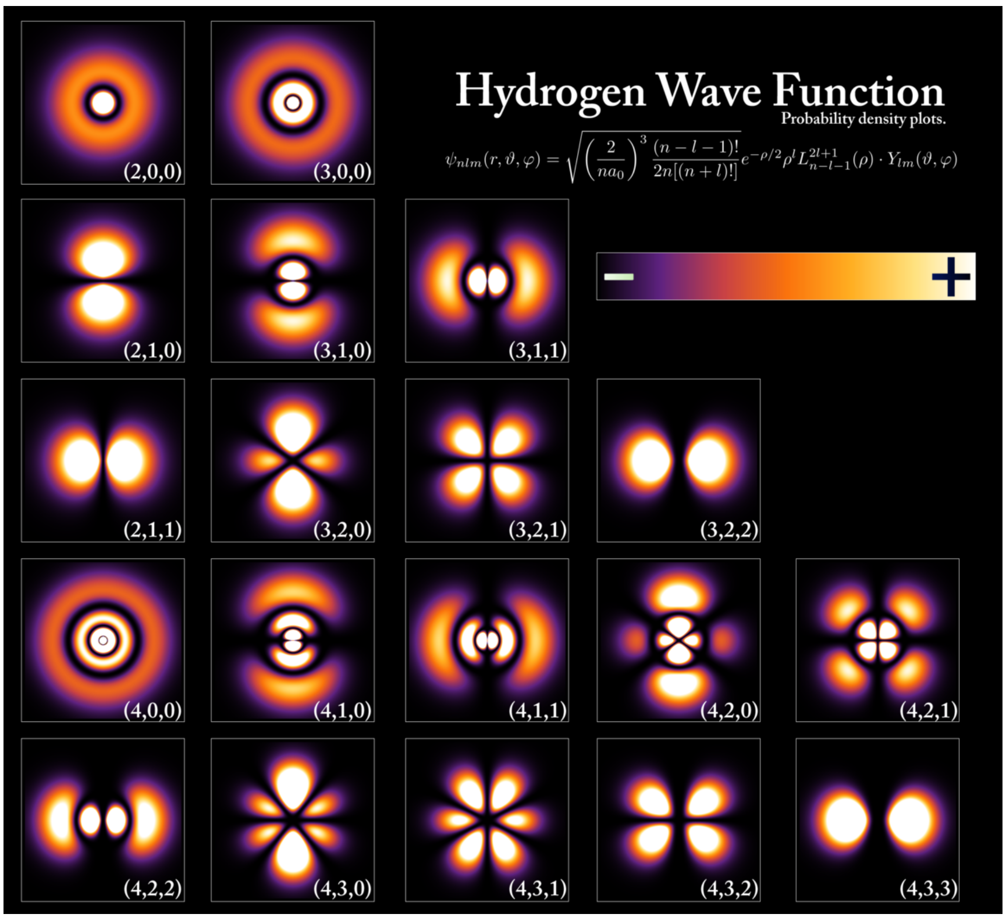


Figure 8.6: The density function u that we seek to find in equation (**??**) describes the probability of finding the electron at the point $(x,y,z) \in \mathbb{R}^3$. This image shows the atomic orbitals at different energy levels. The probability of finding the electron is given by the color. This image is Public Domain, obtained from https://en.wikipedia.org/wiki/Atomic_orbital.

Exercise 94. *Let*

$$x = r\cos\theta\sin\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\phi.$$

Show that

$$\Delta f = f_{rr} + \frac{2}{r} f_r + \frac{1}{r^2 \sin\phi}(f_\phi \sin\phi)_\phi + \frac{1}{r^2 \sin^2\phi} f_{\theta\theta}.$$

Exercise 95. *For a function of the form* $R(r)\Theta(\theta)\Phi(\phi)$*, compute*

$$\Delta(R(r)\Theta(\theta)\Phi(\phi)).$$

Now, the equation (8.6.1) looks a bit more complicated than necessary. We can change the units of mass, so that we can assume $\hbar = m = \epsilon = 1$. Then, our equation becomes

$$\frac{1}{2}\Delta u + \frac{u}{r} + Eu = 0 \iff \Delta u + 2\frac{u}{r} + 2Eu = 0. \tag{8.6.2}$$

Assume our function $u = R(r)\Theta(\theta)\Phi(\phi)$.

Exercise 96. *Using separation of variables, show that (up to a constant multiple) the* Θ *part of the function must be equal to*

$$\Theta(\theta) = e^{im\theta},$$

and that

$$\Phi(\phi) = P_n^{|m|}(\cos\phi),$$

where $n \geq |m|$*. Above,* P_n^m *is the* associated Legendre function,

$$P_n^m(s) = \frac{(1-s^2)^{m/2}}{2^n n!} \frac{d^{n+m}}{ds^{n+m}}(s^2-1)^n.$$

Show that P_n^m *is the solution of the problem for the function* $y = y(s)$ *of one variable,*

$$[(1-s^2)y'] + \frac{m^2 y}{1-s^2} + n(n+1)y = 0, \quad y(-1) = y(1) = 0.$$

Next, we're going to consider the radial part.

Exercise 97. *Show that* R *must satisfy*

$$r^2 R'' + 2rR' + [2Er^2 + 2r - n(n+1)]R = 0.$$

Let's think for a moment about the energy, E. A proton is positively charged. So, if the electron is also positively charged, the two of them repel each other, and the electron runs away. This does not create a hydrogen atom. So, we're interested in negative energy, $E < 0$, because this can create a bond with the proton, so that the electron stays trapped. That's what's happening in a hydrogen atom. So, from now on, we assume

$$E < 0.$$

By introducing a few definitions, we will be able to simplify our problem,

$$\nu = (-2E)^{-1/2}, \quad s = 2\nu^{-1} r, \quad R(r) = S(2\nu^{-1}) = S(s).$$

Exercise 98. *Show that the equation becomes*

$$s^2 S'' + 2sS' + [\nu s - \frac{1}{4}s^2 - n(n+1)]S = 0.$$

Next, let

$$S = s^n e^{-s/2}\Sigma.$$

Show that the equation now becomes

$$s\Sigma'' + (2n+2-s)\Sigma' + (\nu - n - 1)\Sigma = 0. \tag{8.6.3}$$

Verify that this is the Laguerre equation,

$$xy'' + (\alpha + 1 - x)y' + ny,$$

with $\alpha = 2n+1$ *and* n *replaced by* $\nu - n - 1$.

Exercise 99. *Show that the only solutions of* (8.6.3) *that will produce solutions of the form* $u = R\Theta\Phi \in L^2(\mathbb{R}^3)$ *to* (8.6.2) *are the Laguerre polynomials.*

So we now know that $\nu \geq n+1$ and $\nu \in \mathbb{Z}$.

Exercise 100. *Unravel all the substitutions to show that the solution*

$$R_{n\nu}(r) = (2\nu^{-1}r)^n e^{-r/\nu} L^{2n+1}_{\nu-n-1}(2\nu^{-1}r),$$

and

$$u_{mn\nu} = R_{n\nu}(r) e^{im\theta} P_n^{|m|}(\cos(\phi),$$

with

$$E_{mn\nu} = -\frac{1}{2}\nu^{-2}.$$

What is important to notice here is that ν is an *integer*. This means that when ν changes, the energy $E_{mn\nu}$ *jumps*. This is because any two different integers are at least one apart. Therefore, the energy can only come at the levels

$$E_{mn\nu} = -\frac{1}{2}\nu^{-2}, \quad \nu \in \mathbb{Z}, \quad \nu \geq n+1.$$

It is rather fascinating to know that experimental physicists already knew this fact about the energy levels, before the mathematics had been done! Another important observation is that $E_{mn\nu}$ depends only on ν, as long as $\nu \geq n+1$. So, there are a lot of different functions $u_{mn\nu}$ for each $E_{mn\nu}$. What happens to the energy as $\nu \to \infty$?

8.7 Exercises

1. (EO 21) Show that the functions $\varphi_n(x) = \frac{\sin(x/2)}{\pi x} e^{inx}$ are pairwise orthogonal in $\mathcal{L}^2(\mathbb{R})$. Determine coefficients c_n that minimize
$$\int_{-\infty}^{\infty} \left| \frac{1}{1+x^2} - \sum_{n=-N}^{N} c_n \varphi_n(x) \right|^2 dx.$$

2. (EO 36) Determine the polynomial $P(x)$ of at most degree two that minimizes
$$\int_0^\infty |\sqrt{x} - P(x)|^2 e^{-x} dx.$$

3. (EO 37) Determine the polynomial $P(x)$ of at most degree two that minimizes
$$\int_{-\infty}^\infty |x^4 - P(x)|^2 e^{-x^2/2} dx.$$

4. (EO 38) Determine the polynomial $P(x)$ of at most degree two that minimizes
$$\int_0^\infty |e^{x/4} - P(x)|^2 x e^{-x}.$$

5. (EO 39) Determine the polynomial of the form $P(x) = x^3 + ax^2 + bx + c$ that minimizes
$$\int_0^1 |P(x)|^2 dx.$$

6. (EO 41) Compute $H_n'(0)$ for the Hermite polynomials H_n. Hint: use the generating function.

7. [4, 6.1.1] Let $\{p_n\}$ be an orthogonal set of polynomials for $\mathcal{L}^2$ on a bounded interval (a,b), such that p_n is of degree n.

   (a) Fix n. Let $x_1, x_2, \dots, x_k$ be the points in (a,b) where p_n changes sign, and let $q(x) = \prod_{j=1}^{k}(x-x_j)$. Show that $p_n q$ never changes sign on (a,b) and hence $\langle p_n, q\rangle \neq 0$.

   (b) Show that the number k of sign changes in the preceding part is at least n.

   (c) Conclude that p_n has exactly n distinct zeros, all of which lie in (a,b). Geometrically, this shows that p_n becomes more and more oscillatory as $n \to \infty$, similar to $\sin(nx)$.

8. [4, 6.2.1] Show that the n^{th} Legendre polynomial can be written in closed form as

$$P_n(x) = \frac{1}{2^n} \sum_{j \le n/2} \frac{(-1)^j (2n-2j)!}{j!(n-j)!(n-2j)!} x^{n-2j}.$$

   Use this to compute that

$$P_{2k-1}(0) = 0, \quad P_{2k}(0) = \frac{(-1)^k (2k)!}{2^{2k}(k!)^2}.$$

9. [4, 6.2.3] Find the general solution of the Legendre equation

$$\left[(1-x^2)y'\right]' + \lambda y = 0, \quad x \in (-1,1).$$

   It may be helpful to re-write the equation as

$$(1-x^2)y'' - 2xy' + \nu(\nu+1)y = 0,$$

   set $y = \sum_{n\ge 0} a_n x^n$, and solve recursively for the coefficients in terms of a_0 and a_1. Verify that the series converges in the interval $(-1,1)$.

10. [4, 6.2.4] Use the preceding exercise to show that the Legendre equation has a polynomial solution precisely when ν is an integer, and that this solution is a constant multiple of P_ν if $\nu \ge 0$ or $P_{-\nu-1}$ if $\nu < 0$.

11. [4, 6.2.5] Expand x^n for $n = 2, 3, 4$ in series of Legendre polynomials. Hint: no calculus is needed!

12. [4, 6.4.1] Show that the Hermite polynomials can be explicitly computed as

$$H_n(x) = n! \sum_{j \le n/2} \frac{(-1)^j (2x)^{n-2j}}{j!(n-2j)!}.$$

13. [4, 6.4.2] Find the general solution of the Hermite equation

$$y'' - 2xy' + \lambda y = 0,$$

   where $\lambda \in \mathbb{C}$. Do this by assuming $y = \sum_{n\ge 0} a_n x^n$ and solving recursively for the coefficients in terms of a_0 and a_1. Show that the Hermite equation has a polynomial solution of degree n precisely when $\lambda = 2n$, and this solution is a constant multiple of H_n.

14. Compute the best polynomial approximation of degree two in $\mathcal{L}^2(-1,1)$ for e^x. Compare this with the Taylor expansion about zero for e^x.

15. Consider the polynomial $x^3 + 1$. Compute its best polynomial approximations of degree $0, 1, 2$, and 3 in $\mathcal{L}^2(-1,1)$. Compare these with its Taylor expansion.

16. The Laguerre polynomials also have a generating function. It is

$$\sum_{n=0}^{\infty} L_n^{\alpha}(x) z^n = \frac{e^{-xz/(1-z)}}{(1-z)^{\alpha+1}}.$$

If you are really ambitious, you can prove this using the Cauchy integral formula, but this is rather tricky. Instead, use this formula to show that

$$(1-z^2)\frac{\partial}{\partial z} \sum_{n \geq 0} L_n^{\alpha}(x) z^n = [x + (1+\alpha)(z-1)] \sum_{n \geq 0} L_n^{\alpha}(x) z^n.$$

17. Use the preceding exercise to obtain the recursive formula

$$(n+1)L_{n+1}^{\alpha}(x) + (x - \alpha - 2n - 1)L_n^{\alpha}(x) + (n+\alpha)L_{n-1}^{\alpha}(x) = 0.$$

18. Show that

$$(L_n^{\alpha})'(x) - (L_{n-1}^{\alpha})'(x) + L_{n-1}^{\alpha}(x) = 0.$$

19. Show that the Laguerre and Hermite polynomials are related by

$$L_n^{-1/2}(x) = \frac{(-1)^n}{2^{2n} n!} H_{2n}(\sqrt{x}), \quad L_n^{1/2}(x) = \frac{(-1)^n}{2^{2n+1} n!} \frac{H_{2n+1}(\sqrt{x})}{\sqrt{x}}.$$

20. Complete all of the exercises in §8.6.

Chapter 9

The Fourier transform: when a solution needs to be found, transform the problem into sound!

The Fourier transform has become ubiquitous in several different contexts including but not limited to: theoretical mathematics, physics, chemistry, engineering, signal processing, electronics, and medicine. Physically, the Fourier transform takes a function that depends on a variable like $x \in \mathbb{R}$, that we think of as position in one-dimensional space, into a function that depends on $\xi \in \mathbb{R}$, that is the frequency of a wave. We could think of this as changing the problem from something that is happening in physical space to a sound wave. This change in perspective can be useful, for example if we start with an impossible looking equation, or a partial differential equation, or an impossible looking integral. Taking the Fourier transform, we may then obtain a solvable equation, for example by changing a PDE into an ODE, or by obtaining an equivalent but easier looking integral. So, in terms of the (sound) wave frequency variable, the problem becomes solvable, and we solve it. Then, we obtain a solution for the original problem on $x \in \mathbb{R}$ by applying the Fourier inversion theorem (FIT). In §9.9.2, we will see how the Fourier transform can also be used to explain the Heisenberg uncertainty principle, and how it relates to the process of quantizing operators.

Throughout this chapter we will be working in unbounded regions of space. We will first be able to define the Fourier transform for functions such that the integral of the absolute value over the real line is finite. This is a Banach space, sadly it is not a Hilbert space.[1]

Definition 101 (The Banach space $\mathcal{L}^1$)**.** The set

$$\mathcal{L}^1(\mathbb{R}) = \text{ the set of equivalence classes, of functions which satisfy:}$$

$$f \text{ is measurable, and } \int_{\mathbb{R}} |f(x)|dx < \infty.$$

The function g belongs to the same equivalence class as f if $g = f$ almost everywhere on $\mathbb{R}$ with respect to the Lebesgue measure on $\mathbb{R}$.

That is the mathematically rigorous definition, but it suffices for us to work with the 'workable definition' below.

Definition 102 (The workable definition of $\mathcal{L}^1$)**.** The Banach space

$$\mathcal{L}^1(\mathbb{R}) = \text{ the set functions which satisfy: } \int_{\mathbb{R}} |f(x)|dx < \infty.$$

[1] One can prove this statement rigorously, and it is generally done in a functional analysis course. The method is to prove that it is not self-dual, because all Hilbert spaces are self-dual.

Figure 9.1: The Fourier transform does just that: it transforms! In physics, taking the Fourier transform with respect to the spatial variable, we change from a function that depends on position in space into a function that depends on a wave frequency. This can be very useful, because tasks that appear impossible to solve in the space variable can turn into solvable problems in the frequency variable. How does sound propagate? As soundwaves of course! So, we can imagine that the Fourier transforms the problem into sound (waves). Since soundwaves are not visible, to illustrate the concept here is a famous wave, known as the Great Wave off Kanagawa, a woodblock print by the artist Hokusai, published around 1830. Image license and source cc 1.0 `openclipart.org`

Quite a bit changes when we consider the entire real line, or more generally unbounded regions of space, as compared with the previous chapters, that focused on problems and calculations for bounded regions of space. For example, the functions

$$e^{inx}, \sin(x), \cos(x)$$

are all neither in $\mathcal{L}^1(\mathbb{R})$ nor in $\mathcal{L}^2(\mathbb{R})$. Furthermore, there is no relationship between $\mathcal{L}^1(\mathbb{R})$ and $\mathcal{L}^2(\mathbb{R})$. There are functions which are in $\mathcal{L}^1(\mathbb{R})$ but not in $\mathcal{L}^2(\mathbb{R})$. For example, consider

$$f(x) := \begin{cases} 0 & x \leq 0 \\ \frac{1}{\sqrt{x}} & 0 < x < 1 \\ 0 & x \geq 1 \end{cases}.$$

Exercise 103. *Verify that this function f is in $\mathcal{L}^1(\mathbb{R})$ but is not in $\mathcal{L}^2(\mathbb{R})$. Compute its $\mathcal{L}^1(\mathbb{R})$ norm.*

Now consider

$$g(x) = \begin{cases} 0 & x \leq 1 \\ \frac{1}{x} & x > 1 \end{cases}.$$

Exercise 104. *Verify that this function g is in $\mathcal{L}^2(\mathbb{R})$ but is not in $\mathcal{L}^1(\mathbb{R})$. Compute its $\mathcal{L}^2(\mathbb{R})$ norm.*

The function

$$e^{-|x|}$$

is in both $\mathcal{L}^1(\mathbb{R})$ and in $\mathcal{L}^2(\mathbb{R})$.

Exercise 105. *Verify that this function is in both $\mathcal{L}^1(\mathbb{R})$ and $\mathcal{L}^2(\mathbb{R})$. Compute its $\mathcal{L}^1$ and $\mathcal{L}^2$ norms. Come up with your own examples of functions which are*

1. *In $\mathcal{L}^1(\mathbb{R})$ but not in $\mathcal{L}^2(\mathbb{R})$.*
2. *In $\mathcal{L}^2(\mathbb{R})$ but not in $\mathcal{L}^1(\mathbb{R})$.*
3. *In both $\mathcal{L}^1(\mathbb{R})$ and $\mathcal{L}^2(\mathbb{R})$.*

So, all we can say is that

$$\mathcal{L}^1(\mathbb{R}) \not\subset \mathcal{L}^2(\mathbb{R}), \quad \mathcal{L}^2(\mathbb{R}) \not\subset \mathcal{L}^1(\mathbb{R}), \quad \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R}) \neq \emptyset.$$

On the other hand, one can prove that $\mathcal{L}^2(a,b) \subset \mathcal{L}^1(a,b)$ for any bounded interval (a,b).

9.1 A convolution could be the solution!

The convolution is a convoluted way of taking two functions and producing a third function. Let's call the functions f and g (Fred and George). We shift one of the two functions by a real value, x, and then integrate the product, $f(x-y)g(y)$. When we integrate over the whole real line, the result then only depends on the value of x, and so this is a way to obtain a new function. Perhaps we should name it Henry? The convolution has many uses, including allowing us to approximate arbitrary functions with smooth ones. This is a way to cheat out derivatives of functions that are not differentiable. Moreover, the convolution is key to solving the initial value problem for the heat equation in infinite spatial regions.

Definition 106. The convolution of f and g is a function $f * g : \mathbb{R} \to \mathbb{C}$ defined by

$$f * g(x) = \int_{\mathbb{R}} f(x-y)g(y)dy,$$

whenever the integral on the right exists.

Proposition 107. *Assume that f and g are both in $\mathcal{L}^2(\mathbb{R})$. Then*

1. $|f * g(x)| \leq ||f|| ||g||$ *for all* $x \in \mathbb{R}$
2. $f * (ag + bh) = af * g + bf * h$ *for all* $a, b \in \mathbb{C}$
3. $f * g = g * f$
4. $f * (g * h) = (f * g) * h$

Proof: This is useful to do because it helps to familiarize oneself with the convolution. We first estimate

$$|f * g(x)| = \left| \int_{\mathbb{R}} f(x-y)g(y)dy \right| \leq \int_{\mathbb{R}} |f(x-y)||g(y)|dy.$$

The point $x \in \mathbb{R}$ is fixed and arbitrary, so we define a function

$$\phi(y) = f(x-y).$$

Then

$$|f * g(x)| \leq \int_{\mathbb{R}} |\phi(y)||g(y)|dy \leq ||\phi|| ||g||.$$

We compute

$$||\phi||^2 = \int_{\mathbb{R}} |f(x-y)|^2 dy = -\int_{\infty}^{-\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |f(t)|^2 dt = ||f||^2.$$

Above, we used the substitution $t = x - y$ so $dt = -dy$, and the integral got reversed. The $-$ goes away when we re-reverse the integral. So, in the end we see that

$$|f * g(x)| \leq ||f|| ||g||$$

as desired. The second property follows simply by the linearity of the integral itself. For the third property, we will use substitution again:

$$f * g(x) = \int_{\mathbb{R}} f(x-y)g(y)dy.$$

We want to get $g(x-z)$ so we define

$$y = x - z \implies x - y = z, \quad dz = -dy.$$

Hence,

$$f * g(x) = -\int_{\infty}^{-\infty} f(z)g(x-z)dz = \int_{-\infty}^{\infty} g(x-z)f(z)dz = g * f(x).$$

We do something rather similar in the fourth property:

$$f * (g * h)(x) = \int_{\mathbb{R}} f(x-y) \int_{\mathbb{R}} g(y-z)h(z)dzdy.$$

For the other term we have

$$(f * g) * h(x) = \int_{\mathbb{R}} (f * g)(x-y)h(y)dy = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-y-z)g(z)h(y)dzdy.$$

So, we define

$$t = y - z \implies x - y = x - t - z, \quad dt = dy.$$

Then

$$f * (g * h)(x) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x - t - z) g(t) h(z) dz dt.$$

Finally, we call $z = y$ and $t = z$ (sorry if this gives you a headache!) because they are just names, and then we get

$$f * (g * h)(x) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x - y - z) g(y) h(z) dz dy.$$

If you're worried about the order of integration, don't be. Since everything is in $\mathcal{L}^2$, these integrals converge absolutely, so those Italian magicians, Fubini & Tonelli allow us to do the switch-a-roo with the integrals as much as we like.

One of the useful features of convolution is that we can use it to smooth out non-smooth functions. This is known as *mollification,* which comes from the verb, to mollify, which means to make smooth.[2]

Proposition 108 (Mollification)**.** *If $f \in \mathcal{C}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$, $f' \in \mathcal{L}^2(\mathbb{R})$, and $g \in \mathcal{L}^2(\mathbb{R})$, then $f * g \in \mathcal{C}^1(\mathbb{R})$. Moreover $(f * g)' = f' * g$.*

Proof: Everything converges beautifully so just stick that differentiation right under the integral defining

$$f * g(x) = \int_{\mathbb{R}} f(x - y) g(y) dy.$$

Hence

$$(f * g)'(x) = \int_{\mathbb{R}} f'(x - y) g(y) dy = f' * g(x).$$

If you are not satisfied with this explanation, a rigorous proof can be obtained using the Dominated Convergence Theorem, but that is a theorem which we cannot prove in the context of this humble course.

A convoluted example

Let's compute a convolution. Let $f(x) = \frac{1}{1+x^2}$ and

$$g(x) = \begin{cases} 1 & |x| < 3 \\ 0 & |x| > 3 \end{cases}.$$

The function g is not differentiable at the points ± 3. The function f is perfectly smooth on $\mathbb{R}$. Let's convolve them!

$$f * g(x) = \int_{\mathbb{R}} f(x - y) g(y) dy = \int_{\mathbb{R}} \frac{1}{1 + (x - y)^2} g(y) dy = \int_{-3}^{3} \frac{1}{1 + (x - y)^2} dy.$$

[2] One can mollify garlic, tahini, chickpeas, soy sauce, olive oil, oregano, black pepper, lemon juice, in suitable proportions, together with a bit of hot sauce like Cholula, Tabasco, or Sriracha, to make hummus.

If we dig deep into our calculus memory, we vaguely recall that

$$(\arctan(t))' = \frac{1}{1+t^2}.$$

So, this integral becomes:

$$-\arctan(x-y)|_{-3}^{3} = -\arctan(x-3) + \arctan(x+3).$$

This is indeed a smooth function of x.

9.2 The formidable Fourier transform

The Fourier transform is the scalar product of a function $f(x)$ that depends on $x \in \mathbb{R}$ together with a *wave function,* that is a function of the form $e^{ix\xi}$. When the variable $\xi \in \mathbb{R}$, then physically one would refer to the function $e^{-ix\xi}$ as a *wave,* like in Figure 9.1, because

$$e^{ix\xi} = \cos(x\xi) + i\sin(x\xi),$$

so it is the sum of a cosine wave in the real direction and a sine wave in the imaginary direction. The frequency of the wave is ξ. So, when we take the Fourier transform, that is defined as

$$\hat{f}(\xi) = \langle f, e^{ix\xi} \rangle = \int_{\mathbb{R}} f(x)\overline{e^{ix\xi}}dx = \int_{\mathbb{R}} f(x)e^{-ix\xi}dx. \tag{9.2.1}$$

The Fourier transform is now a function of ξ, because the dependence on x has been integrated away. So, the Fourier transform *transforms* the function $f(x)$ that depends on the space variable $x \in \mathbb{R}$ into a function $\hat{f}(\xi)$ that depends on the *frequency* variable ξ. Before we can define the Fourier transform on the Hilbert space $\mathcal{L}^2$, we will define it on the Banach space $\mathcal{L}^1$.

Proposition 109. *Assume that $f \in \mathcal{L}^1(\mathbb{R})$. Then*

$$\hat{f}(\xi) := \int_{\mathbb{R}} f(x)e^{-ix\xi}dx$$

is a well-defined complex number for any $\xi \in \mathbb{R}$.

Proof: Since $f \in \mathcal{L}^1(\mathbb{R})$, by definition we can estimate for $\xi \in \mathbb{R}$

$$\left| \int_{\mathbb{R}} e^{-ix\xi} f(x)dx \right| \le \int_{\mathbb{R}} |f(x)|dx < \infty.$$

This is because for $\xi \in \mathbb{R}$ and $x \in \mathbb{R}$, $|e^{-ix\xi}| = 1$. Consequently, the integral that defines the Fourier transform

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-ix\xi}dx \in \mathbb{C},$$

and so it defines a function $\hat{f} : \mathbb{R} \to \mathbb{C}$. Note that if we do not assume that $\xi \in \mathbb{R}$, then we can say nothing about the integral, because $|e^{-ix\xi}| = e^{x\,\mathrm{Im}(\xi)}$, so if $\mathrm{Im}(\xi) \neq 0$, then this can grow exponentially for large x.

9.2.1 Example of computing a Fourier transform

Let us get a feel for Fourier transforming by doing it! Consider the function $f(x) = e^{-a|x|}$ where $a > 0$. Then it is certainly in $\mathcal{L}^1(\mathbb{R})$ so we ought to be able to compute its Fourier transform. This is by definition

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} e^{-a|x|} dx = \int_{-\infty}^{0} e^{-ix\xi} e^{ax} dx + \int_{0}^{\infty} e^{-ix\xi} e^{-ax} dx.$$

We compute these integrals by finding a primitive for the integrand:

$$\hat{f}(\xi) = \frac{e^{x(a-i\xi)}}{a-i\xi}\Big|_{-\infty}^{0} + \frac{e^{x(-a-i\xi)}}{-a-i\xi}\Big|_{0}^{\infty}$$

$$= \frac{1}{a-i\xi} + \frac{1}{a+i\xi} = \frac{a+i\xi+a-i\xi}{a^2+\xi^2} = \frac{2a}{a^2+\xi^2}.$$

9.3 The formidable Fourier transform's fine features

The following is a useful and fundamental collection of facts about the Fourier transform. It may be useful to introduce the notations

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) = \hat{f}(\xi).$$

Sometimes we feel like a wide hat, sometimes a narrow hat, and sometimes we need that big $\mathcal{F}$. It is useful to be fluent with all three equivalent notations.

Theorem 110 (Properties of the Fourier transform)**.** *Assume that everything below is well defined. Then, the Fourier transform,*

$$\mathcal{F}(f)(\xi) := \hat{f}(\xi) := \int_{\mathbb{R}} f(x) e^{-ix\xi} dx$$

satisfies

1. $\mathcal{F}(f(x-a))(\xi) = e^{-ia\xi}\hat{f}(\xi)$.
2. $\mathcal{F}(f')(\xi) = i\xi\hat{f}(\xi)$
3. $\mathcal{F}(xf(x))(\xi) = i\mathcal{F}(f)'(\xi)$
4. $\mathcal{F}(f * g)(\xi) = \hat{f}(\xi)\hat{g}(\xi)$

Proof: We just compute[3] First

$$\mathcal{F}(f(x-a))(\xi) = \int_{\mathbb{R}} f(x-a) e^{-ix\xi} dx.$$

Change variables. Let $t = x - a$, then $dt = dx$, and $x = t + a$ so

$$\mathcal{F}(f(x-a))(\xi) = \int_{\mathbb{R}} f(t) e^{-i(t+a)\xi} dt = e^{-ia\xi}\hat{f}(\xi).$$

The next one will come from integrating by parts:

$$\int_{\mathbb{R}} f'(x) e^{-ix\xi} dx = f(x) e^{-ix\xi}\big|_{-\infty}^{\infty} - \int_{\mathbb{R}} -i\xi f(x) e^{-ix\xi} dx = i\xi\hat{f}(\xi).$$

[3] We do not need to bother with issues of convergence, because everything is rigorously correct thanks to the dominated convergence theorem. This will be discussed in the appendix eventually.

The boundary terms vanish because of reasons (again it is $\mathcal{L}^1$ and $\mathcal{L}^2$ theory stuff). Similarly we compute

$$\int_{\mathbb{R}} xf(x)e^{-ix\xi}dx = -\frac{1}{i}\int_{\mathbb{R}} f(x)\frac{d}{d\xi}e^{-ix\xi}dx = i\frac{d}{d\xi}\int_{\mathbb{R}} f(x)e^{-ix\xi}dx = i\mathcal{F}(f)'(\xi).$$

Finally,

$$\mathcal{F}(f*g)(\xi) = \int_{\mathbb{R}} f*g(x)e^{-ix\xi}dx = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)g(y)e^{-ix\xi}dydx.$$

We do a little sneaky trick

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)g(y)e^{-ix\xi}e^{-iy\xi}e^{iy\xi}dydx$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)e^{-i(x-y)\xi}g(y)e^{-iy\xi}dydx.$$

Let $z = x - y$. Then $dz = -dy$ so

$$= \int_{\mathbb{R}}\int_{\infty}^{-\infty} f(z)e^{-iz\xi}(-dz)g(y)e^{-iy\xi}dy = \int_{\mathbb{R}}\int_{\mathbb{R}} f(z)e^{-iz\xi}dzg(y)e^{-iy\xi}dy$$

$$= \hat{f}(\xi)\hat{g}(\xi).$$

The following theorem can be proven with sophisticated tools of measure theory and functional analysis that are outside the scope of this text. However, we can understand and apply the statement of the theorem. The first statement is that there is a unique way to make sense of the Fourier transform for functions in $\mathcal{L}^2$, even though this is not very obvious. The proof involves a limiting procedure of estimating $\mathcal{L}^2$ functions by functions that are in $\mathcal{L}^1 \cap \mathcal{L}^2$, because the Fourier transform is well-defined for all of those functions. The second statement in this theorem is the *Fourier inversion theorem,* abbreviated FIT, that shows us how to undo the Fourier transform. Physically, we transform the function of space into a function of frequency, and then the FIT takes us back to space again.

Theorem 111 (Extension of Fourier transform to $\mathcal{L}^2$ and getting FIT)**.** *There is a well defined unique extension of the Fourier transform to $\mathcal{L}^2(\mathbb{R})$. The Fourier transform of an element of $\mathcal{L}^2(\mathbb{R})$ is again an element of $\mathcal{L}^2(\mathbb{R})$. Moreover, for any $f \in \mathcal{L}^2(\mathbb{R})$ we have the FIT (Fourier Inversion Theorem):*

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\xi)e^{ix\xi}d\xi. \tag{9.3.1}$$

If two functions in $\mathcal{L}^2(\mathbb{R})$ have the same Fourier transform, then they are equal.

The following theorem is extremely useful for computing impossible looking integrals. The theorem shows us how the scalar product of two functions in $\mathcal{L}^2$ is related to the scalar product of their Fourier transforms.

Theorem 112 (Plancharel)**.** *For any $f \in \mathcal{L}^2(\mathbb{R})$, $\hat{f} \in \mathcal{L}^2(\mathbb{R})$. Moreover,*

$$\langle \hat{f}, \hat{g}\rangle = 2\pi\langle f, g\rangle,$$

and thus

$$||\hat{f}||^2_{\mathcal{L}^2} = 2\pi||f||^2,$$

for all f and g in $\mathcal{L}^2(\mathbb{R})$.

Figure 9.2: If we take a lot of waves with random frequencies and start adding them up, and letting the frequencies tend to infinity, they will start to cancel each other out. This is the physical meaning of the Riemann & Lebesgue lemma.

Proof: Start with the right side and use the FIT on f, to write

$$2\pi\langle f, g\rangle = 2\pi \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{2\pi} e^{ix\xi} \hat{f}(\xi)\overline{g(x)} d\xi dx = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{ix\xi} \hat{f}(\xi)\overline{g(x)} d\xi dx.$$

Move the complex conjugate to engulf the $e^{ix\xi}$,

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(\xi)\overline{g(x)e^{-ix\xi}} d\xi dx.$$

Swap the order of integration and integrate x first:

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(\xi)\overline{g(x)e^{-ix\xi}} dx d\xi = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi)} d\xi = \langle \hat{f}, \hat{g}\rangle.$$

We may from time to time use the following fact that we also shall not prove. It is named after *two* mathematicians, Riemann and Lebesgue. Roughly speaking, it shows that if we take an $\mathcal{L}^1$ function and compute its Fourier transform, the transform tends to zero as the frequency of the wave tends to infinity. This is related to a physical concept known as the *stationary phase method.* The idea is that if we add up a bunch of waves with random frequencies, and we let these random frequencies tend to infinity, we will get a lot of cancellation. Integration is a mathematical way of summing over a continuous range of values, and so we can view the Riemann & Lebesgue lemma as a consequence of the stationary phase method. An illustration of this method is shown in Figure 9.2.

Lemma 113 (Riemann & Lebesgue)**.** *Assume* $f \in \mathcal{L}^1(\mathbb{R})$*. Then,*

$$\lim_{\xi \to \pm\infty} \hat{f}(\xi) = 0.$$

9.4 The big bad(*ss) convolution approximation theorem: the big bad CAT

This theorem and its proof are both rather long. The proof relies very heavily on knowing the definition of limits and how to work with those definitions, so if you're not comfortable with ϵ and δ style arguments, it would be advisable to review these. The efforts required to prove this theorem are worthwhile, because we will apply it to solving partial differential equations including the heat equation.

To help get you through this, is a picture of my old cat, Romeo in Figure 9.3. He was a big Bengal boy, and if I didn't play with him enough, he destroyed my stuff. It is a bit of an analogy with the big bad CAT; if you don't pay enough attention to it, it could screw up your day.

Figure 9.3: The big bad CAT is your best ally, if you just pay enough attention to it. Similar to an energetic cat or dog. They are by nature good, and if they're misbehaving, it is always due to human error.

Theorem 114 (The big bad CAT)**.** *Assume that $g \in L^1(\mathbb{R})$. Define*

$$\alpha = \int_{-\infty}^{0} g(x)dx, \quad \beta = \int_{0}^{\infty} g(x)dx.$$

Assume that f is piecewise continuous on $\mathbb{R}$ and its left and right sided limits exist for all points of $\mathbb{R}$. Assume that at least one of the following two options is true:

1. *g is zero outside a bounded interval;*
2. *f is bounded on $\mathbb{R}$.*

For $\varepsilon > 0$,

$$g_\epsilon(x) := \frac{g(x/\epsilon)}{\epsilon}.$$

Then

$$\lim_{\epsilon \searrow 0} f * g_\epsilon(x) = \alpha f(x+) + \beta f(x-) \quad \forall x \in \mathbb{R}.$$

Proof. **Idea!**

Do manipulations to get a "left side" statement and a "right side" statement.

We would like to show that

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}} f(x-y) g_\varepsilon(y) dy = \alpha f(x+) + \beta f(x-)$$

which is equivalent to showing that

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}} f(x-y) g_\varepsilon(y) dy - \alpha f(x+) - \beta f(x-) = 0.$$

We now insert the definitions of α and β, so we want to show that

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}} f(x-y) g_\varepsilon(y) dy - \int_{-\infty}^{0} f(x+) g(y) dy - \int_{0}^{\infty} f(x-) g(y) dy = 0.$$

We can prove this if we show that

$$\heartsuit : \lim_{\varepsilon \to 0} \int_{-\infty} f(x-y) g_\varepsilon(y) dy - \int_{-\infty}^{0} f(x+) g(y) dy = 0$$

and also

$$\star : \lim_{\varepsilon \to 0} \int_{0}^{\infty} f(x-y) g_\varepsilon(y) dy - \int_{0}^{\infty} f(x-) g(y) dy = 0.$$

We will prove that $\heartsuit$ holds. The argument is the same for both, so proving one of them is sufficient.

Exercise 115. *To practice this proof on your own, work out the proof for $\star$.*

Hence, we would like to show that by choosing ε sufficiently small, we can make

$$\int_{-\infty}^{0} f(x-y) g_\varepsilon(y) dy - \int_{-\infty}^{0} f(x+) g(y) dy$$

as small as we like. To make this precise, let us assume that "as small as we like" is quantified by a very small $\delta > 0$. Then we show that for sufficiently small ε we obtain

$$\left| \int_{-\infty}^{0} f(x-y) g_\varepsilon(y) dy - \int_{-\infty}^{0} f(x+) g(y) dy \right| < \delta.$$

Idea!

Smash the two integrals together:

$$\int_{-\infty}^{0} \left(f(x-y)g_\varepsilon(y) - f(x+)g(y)\right) dy.$$

Well, this is a bit inconvenient, because in the first part we have g_ε, but in the second part it's just g.

Idea!

Sneak g_ε into the second term. We make a small observation,

$$\int_{-\infty}^{0} g(y)dy = \int_{-\infty}^{0} g(z/\varepsilon)\frac{dz}{\varepsilon} = \int_{-\infty}^{0} g_\varepsilon(z)dz$$

Above, we have made the substitution $z = \varepsilon y$, so $y = z/\varepsilon$, and $dz/\varepsilon = dy$. The limits of integration don't change. By this calculation,

$$\int_{-\infty}^{0} f(x+)g(y)dy = \int_{-\infty}^{0} f(x+)g_\varepsilon(y)dy.$$

(Above the integration variable was called z, but what's in a name? The name of the integration variable doesn't matter!). Moreover, note that $f(x+)$ is a constant, so it's just sitting there doing nothing. Hence, we have computed that

$$\int_{-\infty}^{0} \left(f(x-y)g_\varepsilon(y) - f(x+)g(y)\right) dy = \int_{-\infty}^{0} g_\varepsilon(y)\left(f(x-y) - f(x+)\right) dy.$$

Remember that $y \leq 0$ where we're integrating. Therefore, $x - y \geq x$.

Idea!

Use the definition of right hand limit:

$$\lim_{y\uparrow 0} f(x-y) = f(x+) \implies \lim_{y\uparrow 0} f(x-y) - f(x+) = 0.$$

By the definition of limit there exists $y_0 < 0$ such that for all $y \in (y_0, 0)$

$$|f(x-y) - f(x+)| < \widetilde{\delta}.$$

We are using $\widetilde{\delta}$ for now, to indicate that $\widetilde{\delta}$ is going to be something in terms of δ, engineered in such a way that at the end of our argument we get that for ε sufficiently small,

$$\left| \int_{-\infty}^{0} g_\varepsilon(y) \left(f(x-y) - f(x+) \right) dy \right| < \delta.$$

To figure out this $\widetilde{\delta}$, we use our estimate on the part of the integral from y_0 to 0,

$$\left| \int_{y_0}^{0} (f(x-y) - f(x+)) g_\varepsilon(y) dy \right| \leq \int_{y_0}^{0} |f(x-y) - f(x+)||g_\varepsilon(y)| dy$$

$$\leq \widetilde{\delta} \int_{y_0}^{0} |g_\varepsilon(y)| dy \leq \widetilde{\delta} \int_{\mathbb{R}} |g_\varepsilon(y)| dy = \widetilde{\delta} ||g||.$$

Above, we have used the same substitution trick to see that

$$\int_{\mathbb{R}} |g_\varepsilon(y)| dy = \int_{\mathbb{R}} |g(z)| dz = ||g||,$$

where $||g||$ is the $L^1(\mathbb{R})$ norm of g. By assumption, $g \in L^1(\mathbb{R})$, so this L^1 norm is finite. So, let

$$\widetilde{\delta} = \frac{\delta}{2||g|| + 1}.$$

Note that we're not dividing by zero. So, this is a perfectly decent number. Then, we have the estimate (repeating the above estimate)

$$\left| \int_{y_0}^{0} (f(x-y) - f(x+)) g_\varepsilon(y) dy \right| \leq \int_{y_0}^{0} |f(x-y) - f(x+)||g_\varepsilon(y)| dy$$

$$\leq \widetilde{\delta} \int_{y_0}^{0} |g_\varepsilon(y)| dy \leq \widetilde{\delta} \int_{\mathbb{R}} |g_\varepsilon(y)| dy \leq \frac{\delta}{2}.$$

Idea!

To deal with the other part of the integral, from $-\infty$ to y_0, consider the two cases given in the statement of the theorem separately. It is important to remember that

$$y_0 < 0.$$

So, we wish to estimate

$$\left| \int_{-\infty}^{y_0} (f(x-y) - f(x+)) g_\varepsilon(y) dy \right|.$$

Assume that g vanishes outside a bounded interval. We retain the first part of our estimate, that is

$$\int_{y_0}^{0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy < \frac{\delta}{2}.$$

Next, we again observe that

$$\lim_{\varepsilon \downarrow 0} \frac{y_0}{\varepsilon} = -\infty.$$

By assumption that g vanishes outside a bounded interval, here exists some $R > 0$ such that

$$g(x) = 0 \forall x \in \mathbb{R} \text{ with } |x| > R.$$

Hence, we may choose ε sufficient small so that

$$\frac{y_0}{\varepsilon} < -R.$$

Specifically, let

$$\varepsilon_0 = \frac{1}{-Ry_0} > 0.$$

Then for all $\varepsilon \in (0, \varepsilon_0)$ we compute that

$$\frac{y_0}{\varepsilon} < -R.$$

Hence for all $y \in (-\infty, y_0/\varepsilon)$ we have $g(y) = 0$. Thus, we compute as before using the substitution $z = y/\varepsilon$,

$$\int_{-\infty}^{y_0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy = \int_{-\infty}^{y_0/\varepsilon} |f(x-\varepsilon z) - f(x+)||g(z)|dz = 0,$$

because $g(z) = 0 \forall z \in (-\infty, y_0/\varepsilon)$. Thus, we have the total estimate that for all $\varepsilon \in (0, \varepsilon_0)$,

$$\left| \int_{-\infty}^{0} g_\varepsilon(y) \left(f(x-y) - f(x+) \right) dy \right|$$

$$\leq \int_{-\infty}^{0} |g_\varepsilon(y)||f(x-y) - f(x+)|dy \leq \int_{-\infty}^{y_0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy + \int_{y_0}^{0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy$$

$$< 0 + \frac{\delta}{2} \leq \delta.$$

Finally, we assume that f is bounded, which means that there exists $M > 0$ such that $|f(x)| \leq M$ holds for all $x \in \mathbb{R}$. Hence

$$|f(x-y) - f(x+)| \leq |f(x-y)| + |f(x+)| \leq 2M.$$

So, we have the estimate

$$\left| \int_{-\infty}^{y_0} (f(x-y) - f(x+)) g_\varepsilon(y) dy \right| \leq \int_{-\infty}^{y_0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy \leq 2M \int_{-\infty}^{y_0} |g_\varepsilon(y)|dy.$$

We shall do a substitution now, letting $z = y/\varepsilon$. Then, as we have computed before,

$$\int_{-\infty}^{y_0} |g_\varepsilon(y)|dy = \int_{-\infty}^{y_0/\varepsilon} |g(z)|dz.$$

Here the limits of integration do change, because $y_0 < 0$. Specifically $y_0 \neq 0$, which is why the top limit changes. We're integrating between $-\infty$ and y_0/ε. We know that $y_0 < 0$. So, when we divide it by a really small, but still positive number, like ε, then $y_0/\varepsilon \to -\infty$ as $\varepsilon \to 0$. Moreover, we know that

$$\int_{-\infty}^{0} |g(y)|dy < \infty.$$

What this really means is that

$$\lim_{R\to-\infty} \int_{R}^{0} |g(y)|dy = \int_{-\infty}^{0} |g(y)|dy < \infty.$$

Hence,

$$\lim_{R\to-\infty} \int_{-\infty}^{0} |g(y)|dy - \int_{R}^{0} |g(y)|dy = 0.$$

Of course, we know what happens when we subtract the integral, which shows that

$$\lim_{R\to-\infty} \int_{-\infty}^{R} |g(y)|dy = 0.$$

Since

$$\lim_{\varepsilon\to 0} y_0/\varepsilon = -\infty,$$

this shows that

$$\lim_{\varepsilon\to 0} \int_{-\infty}^{y_0/\varepsilon} |g(y)|dy = 0.$$

Hence, by definition of limit there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$,

$$\int_{-\infty}^{y_0/\varepsilon} |g(y)|dy < \frac{\delta}{4(M+1)}.$$

Then, combining this with our estimates, above, which we repeat here,

$$\left| \int_{-\infty}^{y_0} (f(x-y) - f(x+))g_\varepsilon(y)dy \right| \leq \int_{-\infty}^{y_0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy \leq 2M \int_{-\infty}^{y_0} |g_\varepsilon(y)|dy$$

$$< 2M \frac{\delta}{4(M+1)} < \frac{\delta}{2}.$$

Therefore, we have the estimate that for all $\varepsilon \in (0, \varepsilon_0)$,

$$\left| \int_{-\infty}^{0} g_\varepsilon(y) \left(f(x-y) - f(x+)\right) dy \right|$$

$$\leq \int_{-\infty}^{0} |g_\varepsilon(y)||f(x-y) - f(x+)|dy \leq \int_{-\infty}^{y_0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy + \int_{y_0}^{0} |f(x-y) - f(x+)||g_\varepsilon(y)|dy$$

$$< \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

$\square$

Corollary 116. *Assume the same hypotheses as for the Big Bad CAT, and in addition assume that f is continuous everywhere, and that $\alpha = \beta = \frac{1}{2}$. Then we have for all $x \in \mathbb{R}$,*

$$\lim_{\varepsilon \searrow 0} f * g_\varepsilon(x) = f(x).$$

Proof: We apply the theorem, noting that since f is assumed to be continuous, for all x the left and right limits are equal to the value of the function, specifically

$$f(x_+) = f(x_-) = f(x) \implies \alpha f(x_+) + \beta f(x_-) = f(x), \quad \forall x \in \mathbb{R}.$$

As another corollary we solve the initial value problem for the homogeneous heat equation.

Corollary 117. *Assume that f is a bounded, continuous function defined on $\mathbb{R}$. Then, let*

$$f(x,t) := \int_{\mathbb{R}} f(x-y) \frac{e^{-y^2/(4t)}}{\sqrt{4\pi t}} dy.$$

This function satisfies

$$\partial_t f(x,t) - \partial_{xx} f(x,t) = 0 \quad \forall t > 0, \quad x \in \mathbb{R},$$

and

$$\lim_{t \searrow 0} f(x,t) = f(x).$$

Figure 9.4: Having proven the theorem and its corollary, it is perfectly reasonable to rest and ponder the meaning of these powerful mathematical results.

Proof: First, we note that for $t > 0$ the integral converges beautifully, so we can differentiate under the integral. Since we don't know anything about f being differentiable, we use the fact that convolution is commutative so we can move the x into the exponential, specifically

$$f(x,t) = \int_{\mathbb{R}} f(y) \frac{e^{-(x-y)^2/(4t)}}{\sqrt{4\pi t}} dy.$$

Now we differentiate with respect to t:

$$\partial_t f(x,t) = \int_{\mathbb{R}} f(y) \left[\frac{(x-y)^2}{4t^2} e^{-(x-y)^2/(4t)} (4\pi t)^{-1/2} - \frac{1}{2} \frac{1}{\sqrt{4\pi}} t^{-3/2} e^{-(x-y)^2/(4t)} \right] dy$$

$$= \int_{\mathbb{R}} f(y) \left[\frac{(x-y)^2}{t^{3/2} 8\sqrt{\pi}} - \frac{1}{4\sqrt{\pi} t^{3/2}} \right] e^{-(x-y)^2/(4t)} dy.$$

Next we differentiate with respect to x:

$$\partial_x f(x,t) = \int_{\mathbb{R}} f(y) \frac{-2(x-y)}{4t} e^{-(x-y)^2/(4t)} (4\pi t)^{-1/2} dy$$

$$= \int_{\mathbb{R}} f(y) e^{-(x-y)^2/(4t)} \frac{-(x-y)}{4t^{3/2}\sqrt{\pi}} dy$$

and

$$\partial_{xx} f(x,t) = \int_{\mathbb{R}} f(y) e^{-(x-y)^2/(4t)} \left(-\frac{1}{4\sqrt{\pi} t^{3/2}} + \frac{-2(x-y)}{4t} \left(\frac{-(x-y)}{4t^{3/2}\sqrt{\pi}} \right) \right) dy$$

$$= \int_{\mathbb{R}} f(y) \left(\frac{-1}{4t^{3/2}\sqrt{\pi}} + \frac{(x-y)^2}{8t^{3/2}\sqrt{\pi}} \right) e^{-(x-y)^2/(4t)} dy.$$

Does this look familiar? It's the same as our calculation of $\partial_t f(x,t)$. So, indeed this function satisfies the heat equation. To see that it also satisfies the initial condition, in the sense that $\lim_{t \searrow 0} f(x,t) = f(x)$, let's express this with help from

$$g(y) := \frac{e^{-y^2}}{\sqrt{\pi}}.$$

This function is certainly in $\mathcal{L}^1(\mathbb{R})$. Moreover

$$\int_{-\infty}^{0} g(y) dy = \frac{1}{2} = \int_{0}^{\infty} g(y) dy.$$

So, the convolution approximation theorem tells us that

$$\lim_{\varepsilon \searrow 0} f * g_\varepsilon(x) = f(x).$$

On the left side this is

$$\int_{\mathbb{R}} f(x-y) g(y/\varepsilon) \frac{dy}{\varepsilon} = \int_{\mathbb{R}} f(x-y) e^{-y^2/\varepsilon^2} \frac{dy}{\varepsilon\sqrt{\pi}}.$$

All that matters here is that $\varepsilon > 0$ and it tends to zero. Well, what if we just call $\sqrt{4t} = \varepsilon$. Then since $\varepsilon \to 0$ is equivalent to sending $t \to 0$ we get

$$\lim_{t \searrow 0} \int_{\mathbb{R}} f(x-y) e^{-y^2/(4t)} \frac{dy}{\sqrt{4\pi t}} = f(x).$$

9.5 Applications of the Fourier transform: the initial value problem for the heat equation

The order in which we are doing things may seem a little strange, because we rather amazingly pulled the solution to the heat equation out of thin air. How did we obtain it?

9.5.1 Homogeneous heat equation solved with help from the Fourier transform

We wish to solve:

$$\begin{cases} u_t(x,t) - u_{xx}(x,t) = 0, \quad x \in \mathbb{R}, \quad t > 0 \\ u(x,0) = f(x), \end{cases}$$

where our initial data f is assumed to be bounded and continuous.

Idea!

Fourier transform the PDE with respect to the x variable, because $x \in \mathbb{R}$, whereas $t > 0$, but the Fourier transform integrates over all of $\mathbb{R}$, thus x is the wise choice. We just drive over the whole PDE with the Fourier transform!

So, doing that we have

$$\hat{u}_t(\xi, t) - \hat{u}_{xx}(\xi, t) = 0.$$

Now, we use the theorem which gave us the properties of the Fourier transform. It says that if we take the Fourier transform of a derivative, $\hat{f'}(\xi) = i\xi \hat{f}(\xi)$. Using this twice,

$$\hat{u}_{xx}(\xi, t) = -\xi^2 \hat{u}(\xi, t).$$

Now, those of you who are picky about switching limits may not like this, but it is in fact rigorously valid:

$$\partial_t \hat{u}(\xi, t) + \xi^2 \hat{u}(\xi, t) = 0.$$

Hence

$$\partial_t \hat{u}(\xi, t) = -\xi^2 \hat{u}(\xi, t).$$

This is a first order homogeneous ODE for u in the t variable. We can solve it!!! We do that and get

$$\hat{u}(\xi, t) = e^{-\xi^2 t} c(\xi).$$

The constant can depend on ξ but not on t. To figure out what the constant should be, we use the IC:

$$\hat{u}(\xi, 0) = \hat{f}(\xi) \implies c(\xi) = \hat{f}(\xi).$$

Thus, we have found

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

Now, we use another property of the Fourier transform which says

$$\widehat{f * g}(\xi) = \hat{f}(\xi) \hat{g}(\xi).$$

So, if we can find a function whose Fourier transform is $e^{-\xi^2 t}$, then we can express u as a convolution of that function and f. So, we are looking to find

$$g(x,t) \text{ such that } \hat{g}(x,t) = e^{-\xi^2 t}.$$

We know how to go backwards from the Fourier transform, and that is done by using the FIT! The FIT guarantees that

$$g(x,t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} e^{-\xi^2 t} d\xi.$$

We can use some complex analysis to compute this integral. To do this, we shall complete the square in the exponent:

$$-\xi^2 t + ix\xi = -\left(\xi\sqrt{t} - \frac{ix}{2\sqrt{t}}\right)^2 - \frac{x^2}{4t}.$$

Therefore we are computing

$$\int_{\mathbb{R}} \exp\left(-\left(\xi\sqrt{t} - \frac{ix}{2\sqrt{t}}\right)^2 - \frac{x^2}{4t}\right) d\xi.$$

Using a contour integral, we can in fact ignore the imaginary part. To see this, first note that we are integrating with respect to ξ, so we can for the moment just consider:

$$\int_{-\infty}^{\infty} \exp\left(-\left(\xi t - \frac{ix}{2\sqrt{t}}\right)^2\right) d\xi.$$

We draw a box. The box has vertices in the complex plane at the points $\pm R$ and $\pm R + \frac{ix}{2\sqrt{t}}$. The integrand above is holomorphic for all ξ inside this box. Therefore the integral around the boundary of the box is zero. When $\xi = \pm R$, the integrand is very small, thus the integrals on the vertical sides of the box tend to zero. Hence the integrals along the two horizontal sides of the box are also adding up to zero, which shows that

$$\int_{-\infty}^{\infty} \exp\left(-\left(\xi t - \frac{ix}{2\sqrt{t}}\right)^2\right) d\xi = \int_{-\infty}^{\infty} \exp(-\xi^2 t^2) d\xi.$$

So, we compute (using a change of variables to $y = \xi\sqrt{t}$ so $t^{-1/2} dy = d\xi$)

$$\int_{\mathbb{R}} e^{-\xi^2 t} d\xi = \frac{1}{\sqrt{t}} \int_{\mathbb{R}} e^{-y^2} dy = \frac{\sqrt{\pi}}{\sqrt{t}}.$$

Hence,

$$\int_{\mathbb{R}} \exp\left(-\left(\xi\sqrt{t} - \frac{ix}{2\sqrt{t}}\right)^2 - \frac{x^2}{4t}\right) d\xi = \frac{\sqrt{\pi}}{\sqrt{t}} e^{-\frac{x^2}{4t}}.$$

Recalling the factor of $1/(2\pi)$ we have

$$g(x,t) = \frac{1}{2\pi} \frac{\sqrt{\pi}}{\sqrt{t}} e^{-\frac{x^2}{4t}} = \frac{1}{2\sqrt{\pi t}} e^{-\frac{x^2}{4t}}.$$

Hence the solution is

$$u(x,t) = f * g(x) = \int_{\mathbb{R}} \frac{1}{2\sqrt{\pi t}} f(x-y) e^{-y^2/(4t)} dy.$$

If we naively set $t = 0$, we obtain an expression that does not make sense. So, how do we know that this expression indeed gives us our initial data at $t = 0$? We use the BB CAT and its corollary!

The corollary shows that

$$\lim_{t \searrow 0} \int_{\mathbb{R}} f(x-y) \frac{e^{-\frac{y^2}{4t}}}{\sqrt{4\pi t}} dy = f(x), \quad \forall x \in \mathbb{R}.$$

Moreover, we also obtain that $u(x,t)$ is smooth in both x and t for all $t > 0$, because the function we are convolving with has these properties. The solution is also the unique one if we require the solution of the homogeneous heat equation is a non-increasing function of $t > 0$. That makes sense physically, because there is no source in our heat equation, so the heat should not be increasing as time passes.

9.5.2 Inhomogeneous heat equation

If you have an inhomogeneous IVP for the heat equation, here are two ways to deal with that:

1. If the inhomogeneity is *time independent,* look for a steady state solution to solve the inhomogeneous equation. Then, solve the homogeneous equation, but change your initial data. If f is your steady state solution and v was your initial data (before f came along), solve the IVP for the homogeneous heat equation with IC $v - f$ rather than just v.

2. If the inhomogeneity is *time dependent,* you can solve by Fourier transforming the whole PDE!

The second method subsumes the first, in fact, because it works for both time dependent as well as time independent inhomogeneities, so we focus on that more powerful method.

Consider an inhomogeneous heat equation on $\mathbb{R}$:

$$u_t - u_{xx} = G(x,t), \quad u(x,0) = f(x) \text{ is continuous and bounded.}$$

We begin by Fourier transforming the whole equation

$$\partial_t \hat{u}(\xi,t) + \xi^2 \hat{u}(\xi,t) = \hat{G}(\xi,t).$$

This is a first order ODE, and there is a method for solving it. The mμthod, the method of integrating factor that is often called μ begins by computing

$$e^{\int \xi^2 dt} = e^{\xi^2 t}.$$

At this stage one can ignore the constant of integration. The solution to the inhomogeneous ODE is then obtained as

$$\frac{\int_0^t e^{\xi^2 s} \hat{G}(\xi,s) ds + C(\xi)}{e^{\xi^2 t}} = e^{-\xi^2 t} \int_0^t e^{\xi^2 s} \hat{G}(\xi,s) ds + C(\xi) e^{-\xi^2 t}.$$

We would like the initial condition to be satisfied, so when $t = 0$ we should obtain that this is equal to the Fourier transform of the initial data,

$$\hat{f}(\xi).$$

Since the integral from 0 to 0 is just zero, this shows us that the constant (at least from the perspective of time, this is a constant) should be

$$C(\xi) = \hat{u}(\xi,0) = \hat{f}(\xi).$$

The last term is therefore

$$\hat{f}(\xi) e^{-\xi^2 t}.$$

This is the same as what we obtained in solving the homogeneous heat equation. The properties of the Fourier transform indicate that this is the Fourier transform of the convolution of f and the function whose

Fourier transform is $e^{-\xi^2 t}$. We can either look back at how we solved the homogeneous heat equation or look it up on a table, and either way we obtain that

$$\hat{f}(\xi)e^{-\xi^2 t} \text{ is the Fourier transform in } x \text{ at } \xi \text{ of } \int_{\mathbb{R}} f(x-y)e^{-y^2/(4t)}(4\pi t)^{-1/2}dy.$$

We therefore just need to understand the first part of the solution,

$$e^{-\xi^2 t}\int_0^t e^{\xi^2 s}\hat{G}(\xi,s)ds.$$

We can bring in that factor to the integral:

$$= \int_0^t e^{\xi^2(s-t)}\hat{G}(\xi,s)ds.$$

So, what we would really like to do right now is invert the Fourier transform, by computing

$$\frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}\int_0^t e^{\xi^2(s-t)}\hat{G}(\xi,s)dsd\xi.$$

The convergence is absolute, so we can switch the orders of the integrals

$$\int_0^t \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}e^{\xi^2(s-t)}\hat{G}(\xi,s)d\xi ds.$$

Consequently, we just need to inverse Fourier transform $e^{\xi^2(s-t)}\hat{G}(\xi,s)$. This is a product, so by the properties of the Fourier transform, if we can identify a function whose Fourier transform is $e^{\xi^2(s-t)}$, then the inverse Fourier transform of $e^{\xi^2(s-t)}\hat{G}(\xi,s)$ is the convolution of that function with G. Since $0 \le s \le t$, we write $e^{\xi^2(s-t)} = e^{-\xi^2(t-s)}$, because this makes it clear that the exponent is not positive. So, we know how to inverse Fourier transform this! The function whose Fourier transform is this is our heat kernel at time $t-s$,

$$H(y,t-s) = \frac{1}{\sqrt{4\pi(t-s)}}e^{-y^2/(4(t-s)}.$$

Consequently the function whose Fourier transform is $e^{-\xi^2(t-s)}\hat{G}(\xi,s)$ is the convolution

$$\int_{\mathbb{R}} G(x-y,s)H(y,t-s)dy = \int_{\mathbb{R}} G(x-y,s)\frac{1}{\sqrt{4\pi(t-s)}}e^{-y^2/(4(t-s)}dy.$$

So, this part of the solution is therefore

$$\int_0^t\int_{\mathbb{R}} G(x-y,s)\frac{1}{\sqrt{4\pi(t-s)}}e^{-y^2/(4(t-s)}dyds,$$

and the full solution is

$$u(x,t) = \int_0^t\int_{\mathbb{R}} G(x-y,s)\frac{1}{\sqrt{4\pi(t-s)}}e^{-y^2/(4(t-s)}dyds + \int_{\mathbb{R}} f(x-y)e^{-y^2/(4t)}(4\pi t)^{-1/2}dy.$$

This solution satisfies our initial data because

$$\lim_{t\downarrow 0}\int_0^t \frac{1}{2\sqrt{\pi(t-s)}}\int_{\mathbb{R}} e^{-\frac{(y)^2}{4(t-s)}}G(x-y,s)dyds = 0,$$

and the BB CAT helps us with its corollary that guarantees

$$\lim_{t\downarrow 0}\int_{\mathbb{R}} \frac{1}{2\sqrt{\pi t}}e^{-\frac{(y)^2}{4t}}f(x-y)dy = f(x) \quad \forall x\in\mathbb{R}.$$

9.5.3 Computing tricky integrals: sometimes π falls out of the sky!

In some of my research I wish to compute complicated integrals. Like in this pre-print, for example https://arxiv.org/abs/2010.02776. I wanted very much to cheat and make Mathematica compute some integrals there, but it could not do it. So, we had to do them by hand. In that work we used the residue theorem from complex analysis, and that is a wonderful technique for computing difficult integrals. Now, you can add a further technique to your mathematical toolbox: the Fourier transform.

Idea!

The following is a very useful observation:

$$\hat{f}(0) = \int_{\mathbb{R}} f(x)dx.$$

So, if you have the integral of a function, this is equal to the value of its Fourier transform at $\xi = 0$. So, if you can look up the Fourier transform of the function in a table or list of Fourier transforms, like the Mathematical Handbook Beta, then to compute the integral, no need for fancy contour integrals, simply pop $\xi = 0$ into the Fourier transform. Here we have collected some of the most common functions and their Fourier transforms in Table 9.1. You are welcome to use these!

Here is an example:

$$\text{compute: } \int_{\mathbb{R}} \frac{1}{x^2+9}dx. \tag{9.5.1}$$

We see the integrand is an item in Table 9.1; it is number 10 with $a = 3$. On the right side of the table, we get the Fourier transform (with $a = 3$) is given by

$$\frac{\pi}{3}e^{-3|\xi|}.$$

So, the integral in (9.5.1) is equal to the Fourier transform of $\frac{1}{x^2+9}$ evaluated when $\xi = 0$. So, to compute the value of the integral, we just take the Fourier transform on the right side of the table and plug in $\xi = 0$. When we do this we obtain $\frac{\pi}{3}$. Hence the value of the integral is

$$\int_{\mathbb{R}} \frac{1}{x^2+9}dx = \frac{\pi}{3}.$$

This could look surprising because the integral on the left is a rational function, and yet what comes out is a transcendental number. (π is transcendental, and therefore $\frac{\pi}{3}$ is also transcendental.) If we think about

1.	$f(x)$	$\hat{f}(\xi)$
2.	$f(x-c)$	$e^{-ic\xi}\hat{f}(\xi)$
3.	$e^{icx}f(x)$	$\hat{f}(\xi-c)$
4.	$f(ax)$	$a^{-1}\hat{f}(a^{-1}\xi)$
5.	$f'(x)$	$i\xi\hat{f}(\xi)$
6.	$xf(x)$	$i(\hat{f})'(\xi)$
7.	$(f*g)(x)$	$\hat{f}(\xi)\hat{g}(\xi)$
8.	$f(x)g(x)$	$(2\pi)^{-1}(\hat{f}*\hat{g})(\xi)$
9.	$e^{-ax^2/2}$	$\sqrt{2\pi/a}e^{-\xi^2/(2a)}$
10.	$(x^2+a^2)^{-1}$	$(\pi/a)e^{-a\|\xi\|}$
11.	$e^{-a\|x\|}$	$2a(\xi^2+a^2)^{-1}$
12.	$\chi_a(x) := \begin{cases} 1 & \|x\| < a \\ 0 & \|x\| > a \end{cases}$	$2\xi^{-1}\sin(a\xi)$
13.	$x^{-1}\sin(ax)$	$\pi\chi_a(\xi) = \begin{cases} \pi & \|\xi\| < a \\ 0 & \|\xi\| > a \end{cases}$

Table 9.1: On the left we have functions and on the right are their Fourier transforms. Here $a > 0$ and $c \in \mathbb{R}$ are constants.

Figure 9.5: It might seem weird that when we compute the integral of a rational function, like $(x^2 + a^2)$, over the real line, the result involves π. This is somewhat similar to how computing seemingly impossible sums using Fourier series, we has had π falling out of the sky. With the Fourier transform, to make a distinction, it's savory π, like a pizza pie, that falls out of the sky.

complex analysis, however, it makes sense, because there is a 2π in the residue theorem. So, this pizza π falling out of the sky is not such a surprise, but still we enjoy Figure 9.5.

Here is another example, we wish to compute

$$\int_{\mathbb{R}} f(x)g(x)dx,$$

with some complicated functions f and g (see extra övning # 9). Now, you can use that the Fourier transform of a product is

$$(2\pi)^{-1}(\hat{f} * \hat{g})(\xi).$$

Hence, what you have above is

$$\int_{\mathbb{R}} f(x)g(x)dx = \int_{\mathbb{R}} e^{-i(0)x} f(x)g(x)dx = (2\pi)^{-1}(\hat{f} * \hat{g})(0).$$

So, if the Fourier transforms of these functions are somewhat better than the functions f and g, then the stuff on the right could be nicely computable and give you the integral on the left. We will see examples of this in the exercises at the end of the chapter.

As another example, let's say that somehow you know the Fourier transform of $f(t)$ is $\frac{1}{|w|^3+1}$. We then would like to compute

$$\int_{\mathbb{R}} |f * f'|^2 dt.$$

This may seem impossible at first, because we only know the Fourier transform of $f(t)$. We could start by trying to use the FIT, that says

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{iwt} \frac{1}{|w|^3 + 1} dw.$$

In fact, if you are in love with complex analysis, you could compute this integral explicitly using the residue theorem. Then, you would need to compute the integral of $|f * f'|^2$, and I reasonably expect things could become extremely complicated. There is an easier way to the solution.

By the Plancharel theorem,

$$\int_{\mathbb{R}} |f * f'|^2 dt = \frac{1}{2\pi} \int_{\mathbb{R}} |\widehat{f * f'}|^2 dt. \tag{9.5.2}$$

Now we use the theorem on the properties of the Fourier transform which says

$$\widehat{f * f'}(\xi) = \hat{f}(\xi)\widehat{f'}(\xi).$$

Now we use that same theorem to say that

$$\widehat{f'}(\xi) = i\xi \hat{f}(\xi).$$

So, the stuff on the right in (9.5.2) is

$$\frac{1}{2\pi} \int_{\mathbb{R}} |\hat{f}(\xi) i\xi \hat{f}(\xi)|^2 d\xi.$$

We are given what the Fourier transform is, so we put it in there:

$$\frac{1}{2\pi} \int_{\mathbb{R}} \frac{\xi^2}{(|\xi|^3 + 1)^4} d\xi.$$

Now this isn't so terrible. It's an even function so this is

$$\frac{1}{\pi} \int_0^\infty \frac{\xi^2}{(\xi^3 + 1)^4} d\xi.$$

It just so happens that the derivative of

$$\frac{1}{(\xi^3+1)^3} \text{ is } \frac{-9\xi^2}{(\xi^3+1)^4},$$

so

$$\frac{1}{\pi} \int_{\mathbb{R}} \frac{\xi^2}{(\xi^3+1)^4} d\xi = \frac{-1}{9\pi} \left. \frac{1}{(\xi^3+1)^3} \right|_0^\infty = \frac{1}{9\pi}.$$

Voilá, more (pizza) π falls out of the sky!

9.6 Fourier sine and cosine transforms and applications to PDEs on half-spaces

Today we shall investigate some transforms related to the Fourier transform. The first two can be used to solve PDEs on half lines, *if the boundary condition* is suitable.

9.6.1 Fourier sine and cosine transforms and their inverse formulas

Similar to the Fourier sine and cosine series, the Fourier sine and cosine transforms are obtained by taking a function that is in $\mathcal{L}^2(0,\infty)$, and extending it to $\mathcal{L}^2(\mathbb{R})$ to be either odd (sine transform) or even (cosine transform).

Definition 118. Let f be in $\mathcal{L}^1$ or $\mathcal{L}^2$ on $(0,\infty)$. The Fourier cosine transform,

$$\mathcal{F}_c(f)(\xi) := \int_0^\infty f(x)\cos(\xi x)dx.$$

The Fourier sine transform,

$$\mathcal{F}_s(f)(\xi) := \int_0^\infty f(x)\sin(\xi x)dx.$$

As with the Fourier transform, the Fourier sine and cosine transforms also have inversion formula.

Theorem 119. *Assume that $f \in \mathcal{L}^2[0,\infty)$. Then we have the Fourier cosine inversion formula*

$$f(x) = \frac{2}{\pi}\int_0^\infty \mathcal{F}_c(f)(\xi)\cos(x\xi)d\xi.$$

We also have the Fourier sine inversion formula

$$f(x) = \frac{2}{\pi}\int_0^\infty \sin(x\xi)\mathcal{F}_s(f)(\xi)d\xi.$$

Proof: First, let us extend f evenly to $\mathbb{R}$, denoting this extension by f_e, so that $f_e(-x) = f_e(x)$. We compute the standard Fourier transform:

$$\hat{f}_e(\xi) = \int_{\mathbb{R}} f_e(x)e^{-ix\xi}dx = \int_{\mathbb{R}} f_e(x)(\cos(x\xi) - i\sin(x\xi))dx = 2\int_0^\infty f(x)\cos(x\xi)dx.$$

The term with the sine has dropped out because $f_e(x)\sin(x\xi)$ is an odd function of x. The term with the cosine gets doubled because $f_e(x)\cos(x\xi)$ is an even function. So, all together we have computed:

$$\hat{f}_e(\xi) = 2\int_0^\infty f(x)\cos(x\xi)dx = 2\mathcal{F}_c(f)(\xi).$$

Since the cosine is an even function,

$$\hat{f}_e(\xi) = \hat{f}_e(-\xi).$$

So, we also have that $\mathcal{F}_c(f)$ is an even function. The inversion formula (FIT) says that

$$f_e(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}\hat{f}_e(\xi)d\xi = \frac{1}{\pi}\int_{\mathbb{R}} e^{ix\xi}\mathcal{F}_c(f)(\xi)d\xi$$

$$= \frac{1}{\pi}\int_{\mathbb{R}} (\cos(x\xi) + i\sin(x\xi))\mathcal{F}_c(f)(\xi)d\xi = \frac{2}{\pi}\int_0^\infty e^{ix\xi}\mathcal{F}_c(f)(\xi)d\xi.$$

This is the cosine-FIT! Above we have used the fact that $\mathcal{F}_c(f)$ is an even function. Hence its product with the cosine is also an even function, so that part of the integral gets a factor of two when we integrate only over the positive real line. The product of an even function like $\mathcal{F}_c(f)$ with an odd function, like the sine, is odd, so that integral vanishes.

On the other hand, we may also define the odd extension, f_o which satisfies $f_o(-x) = -f_o(x)$ (for $x \neq 0$). The value of f at zero is not really important at this moment.[4] We compute the standard Fourier transform of the odd extension:

$$\hat{f}_o(\xi) = \int\limits_{\mathbb{R}} f_o(x)e^{-ix\xi}dx = \int\limits_{\mathbb{R}} f_o(x)(\cos(x\xi) - i\sin(x\xi))dx = -2i\int\limits_0^\infty f(x)\sin(x\xi)dx$$

$$= -2i\mathcal{F}_s(f)(\xi).$$

Above, we have used the fact that f_o is odd, and therefore so is its product with the cosine. On the other hand, the product with the sine is an even function, which explains the factor of 2. Since the sine itself is odd, we have that $\hat{f}_o$ is an odd function and similarly $\mathcal{F}_s(f)(\xi)$ is also an odd function. We apply the FIT:

$$f_o(x) = \frac{1}{2\pi}\int\limits_{\mathbb{R}} e^{ix\xi}\hat{f}_o(\xi)d\xi = -\frac{i}{\pi}\int\limits_{\mathbb{R}} (\cos(x\xi) + i\sin(x\xi))\mathcal{F}_s(f)(\xi)d\xi$$

$$= \frac{1}{\pi}\int\limits_{\mathbb{R}} \sin(x\xi)\mathcal{F}_s(f)(\xi)d\xi = \frac{2}{\pi}\int\limits_0^\infty \sin(x\xi)\mathcal{F}_s(f)(\xi)d\xi).$$

This is the sine-FIT! Above we have used the fact that $\mathcal{F}_s(f)$ is an odd function, and therefore so is its product with the cosine. On the other hand the product of two odd functions is an even function, so that is the reason for the factor of 2.

9.6.2 Solving the heat equation on a semi-infinite rod with insulated end

We wish to solve the problem:

$$u_t - u_{xx} = 0, \quad u_x(0,t) = 0, \quad u(x,0) = f(x), \quad x \in [0,\infty).$$

Assume that by some method, we have obtained a solution $u(x,t)$ defined on $[0,\infty)_x \times [0,\infty)_t$. To see if we may use a Fourier sine or cosine transform method, let us see what happens when we extend our solution evenly or oddly. The even extension would satisfy, by the cosine-FIT:

$$u_e(x,t) = \frac{2}{\pi}\int\limits_0^\infty \mathcal{F}_c(u)(\xi)\cos(x\xi)d\xi.$$

The odd extension would satisfy, by the sine-FIT

$$u_o(x,t) = \frac{2}{\pi}\int\limits_0^\infty \sin(x\xi)\mathcal{F}_s(f)(\xi)d\xi.$$

[4] This is because we are working in $\mathcal{L}^2$ which ignores sets of measure zero, and a single point is a set of measure zero.

OBS! The extension matches up with our original function on the positive real line (that is how an extension works!) We need the derivative with respect to x to vanish at $x = 0$. Let's just differentiate these expressions. Note that the x dependence is only in the sine or cosine term so we have:

$$\partial_x u_e(x,t) = -\frac{2}{\pi}\int_0^\infty \mathcal{F}_c(u)(\xi)\xi\sin(x\xi)d\xi \implies \partial_x u_e(0,t) = 0.$$

On the other hand

$$\partial_x u_o(x,t) = \frac{2}{\pi}\int_0^\infty \xi\cos(x\xi)\mathcal{F}_s(u)(\xi)d\xi \implies \partial_x u_o(0,t) = \frac{2}{\pi}\int_0^\infty \xi\mathcal{F}_s(u)(\xi)d\xi = ???$$

The even extension automatically gives us the desired boundary condition whereas the odd extension leads to something complicated looking, which we have no reason to know is zero.

Although we could try to work with the Fourier cosine transform to solve this problem, we would then need to prove an analogue of Theorem 110 for the Fourier cosine transform, and that is rather a lot of work. It is a bit easier to do this procedure instead:

1. Extending the initial data *evenly* to the real line.
2. Solving the problem using the Fourier transform on the real line.
3. Verifying that the solution satisfies all the conditions: the PDE, the IC, and the BC.

We do this. Extend f evenly, and write the extension as f_e. Then the solution to the homogeneous heat equation on the real line with initial data f_e is

$$u_e(x,t) = \frac{1}{2\sqrt{\pi t}}\int_{\mathbb{R}} f_e(y)e^{-\frac{(x-y)^2}{4t}}dy.$$

We split up the integral:

$$\int_{-\infty}^0 f_e(y)e^{-(x-y)^2/(4t)}dy + \int_0^\infty f_e(y)e^{-(x-y)^2/(4t)}dy$$

$$= -\int_\infty^0 f_e(z)e^{-(z+x)^2/(4t)}dz + \int_0^\infty f_e(y)e^{-(x-y)^2/(4t)}dy.$$

Above we made the substitution that $z = -y$ in the first integral. Due to the evenness of f_e, nothing happens when we change $y = -z$. Reversing the limits of integration we obtain

$$-\int_\infty^0 f_e(z)e^{-(z+x)^2/(4t)}dz = \int_0^\infty f_e(z)e^{-(z+x)^2/(4t)}dz = \int_0^\infty f_e(y)e^{-(x+y)^2/(4t)}dy.$$

So, all together we have

$$u_e(x,t) = \frac{1}{2\sqrt{\pi t}}\int_0^\infty f(y)\left(e^{-\frac{(x-y)^2}{4t}} + e^{-\frac{(x+y)^2}{4t}}\right)dy.$$

Is this an even function? Let us verify:

$$e^{-\frac{(x-y)^2}{4t}} + e^{-\frac{(x+y)^2}{4t}} = e^{-\frac{(-x-y)^2}{4t}} + e^{-\frac{(-x+y)^2}{4t}}.$$

AWESOME! Our solution to the heat equation in this way is EVEN. Therefore, it is the same on the left and right sides. So, we can simply let

$$u(x,t) = u_e(x,t) = \frac{1}{2\sqrt{\pi t}} \int_0^\infty f(y) \left(e^{-\frac{(x-y)^2}{4t}} + e^{-\frac{(x+y)^2}{4t}} \right) dy.$$

The way we have built it, it satisfies the IC, BC, and the PDE!

Exercise 120. *Solve:*

$$u_t - u_{xx} = 0, \quad u(0,t) = 0, \quad u(x,0) = \phi(x), \quad x \in [0, \infty).$$

Above, we assume that ϕ is bounded, continuous, and in $\mathcal{L}^2$. Hint: extend ϕ oddly *this time, and follow the analogous procedure.*

9.7 Dirichlet problem in a quadrant

Consider the problem

$$u_{xx} + u_{yy} = 0, \quad x, y > 0, \quad u(x,0) = f(x), \quad u(0,y) = g(y).$$

To deal with these inhomogeneities let us instead solve two nicer problems:

1. $w_{xx} + w_{yy} = 0, \quad x, y > 0, \quad w(x,0) = f(x), \quad w(0,y) = 0.$
2. $v_{xx} + v_{yy} = 0, \quad x, y > 0, \quad v(x,0) = 0, \quad v(0,y) = g(y).$

The full solution will then be obtained by setting

$$u(x,y) = w(x,y) + v(x,y).$$

Exercise 121. *Verify that if w and v solve the problems above, then indeed u solves the original problem.*

We would like to use Fourier methods, but the problems we have above $x, y > 0$. The Fourier transform is defined on the whole plane. So, we may wish to use an even or odd extension.

Idea!

To solve a problem like $w_{xx} + w_{yy} = 0, \quad x, y > 0, \quad w(x,0) = f(x), \quad w(0,y) = 0$, look at the boundary condition. The solution should vanish at $x = 0$. Now think about sine and cosine. Which of these vanishes at $x = 0$? The sine. That is an odd function. So this gives us the clue to extend oddly.

We define therefore

$$f_o(x) := \begin{cases} f(x) & x > 0 \\ -f(-x) & x < 0 \end{cases}.$$

Now, we take the Fourier transform of the PDE in the x variable. We obtain:

$$-\xi^2 \hat{w}(\xi, y) + \partial_{yy} \hat{w}(\xi, y) = 0 \implies \hat{w}(\xi, y) = A(\xi) e^{-\xi y} + B(\xi) e^{\xi y}.$$

The boundary condition we have says that

$$\hat{w}(\xi, 0) = \hat{f}(\xi) = A(\xi) + B(\xi).$$

Now, in our problem, we have $y > 0$. So, if we look at the term $e^{\xi y}$ this grows exponentially as $y \to \infty$. Thus, it did not come from a function in $\mathcal{L}^2$ or in $\mathcal{L}^1$. It did not come from something Fourier-transformable. So, we shall try to solve the problem using only the part which decays as $y \to infty$. The boundary condition at $y = 0$ acts like an initial condition, at least from x's perspective:

$$\hat{w}(\xi, 0) = A(\xi) = \hat{f}_o(\xi) \implies \hat{w}(\xi, y) = \hat{f}_o(\xi) e^{-\xi y}.$$

We look at the table to find a function whose Fourier transform is $e^{-\xi y}$. OBS! The transform is occurring in the x variable, from whose perspective y is a constant. Thus, the item on the table is a slight modification of 10, in particular the function

$$\frac{y}{\pi}(x^2 + y^2)^{-1} \text{ has Fourier transform in the } x \text{ variable } e^{-y|\xi|}.$$

Thus, we have found

$$\hat{w}(\xi, y) = \hat{f}_o(\xi) \frac{y}{\pi} \widehat{(x^2 + y^2)^{-1}}(\xi).$$

The Fourier transform sends convolutions to products, which tells us that

$$w(x, y) = \int_{-\infty}^{\infty} f_o(z) \frac{y}{\pi((x-z)^2 + y^2)} dz = \int_{-\infty}^{0} f_o(z) \frac{y}{\pi((x-z)^2 + y^2)} dz + \int_{0}^{\infty} f(z) \frac{y}{\pi((x-z)^2 + y^2)} dz$$

We do a substitution in the first integral, with $t = -z$

$$= \int_{-\infty}^{0} f_o(z) \frac{y}{\pi((x-z)^2 + y^2)} dz = -\int_{\infty}^{0} f_o(-t) \frac{y}{\pi((x+t)^2 + y^2)} dt$$

$$= \int_{\infty}^{0} f(t) \frac{y}{\pi((x+t)^2 + y^2)} dt = -\int_{0}^{\infty} f(t) \frac{y}{\pi((x+t)^2 + y^2} dt.$$

Re-naming the variable of integration z, we get

$$w(x, y) = \int_{0}^{\infty} f(z) \left[\frac{y}{\pi((x-z)^2 + y^2)} - \frac{y}{\pi((x+z)^2 + y^2)} \right] dz.$$

The other problem is basically identical, we simply Fourier transform in the y variable. Thus the solution to the second problem is

$$v(x, y) = \int_{0}^{\infty} g(z) \left[\frac{x}{\pi((y-z)^2 + x^2)} - \frac{x}{\pi((y+z)^2 + x^2)} \right] dz.$$

We obtain the full solution by adding:

$$u(x, y) = w(x, y) + v(x, y).$$

9.8 The Sampling Theorem and the discrete and fast Fourier transforms

The Sampling Theorem is an amazing fact. It states that if the Fourier transform of a function lives in a bounded interval, then the function is completely specified by its values at a discrete set of points. A few explanations are in order. We say that a function *lives* in a bounded interval if it is never seen outside that interval, in other words, if it *vanishes* outside that interval. A function vanishes when it is equal to zero. So, this means that such a function only assumes nonzero values within some interval (a, b) for some $a < b$ which are real numbers. By defining

$$L = \max\{|a|, |b|\},$$

then we can also say that the function only assumes nonzero values on the (probably larger) interval $[-L, L]$. The discrete set of values mentioned is then the set of points

$$\left\{\frac{n\pi}{L}\right\}_{n\in\mathbb{Z}}.$$

Sure, there are infinitely many such points, but they are discrete in the sense that they are evenly spaced, they never cluster up. The theorem says that we can specify the value of f at *all* $t \in \mathbb{R}$, including those values that fall between these points. Now, the *vast majority* of points of $\mathbb{R}$ are *not* contained in this discrete set, so this is a pretty amazing fact! The reason it is called the *sampling* theorem is that we can think of measuring the value of f at these points as *sampling* f. We just need to sample f at these points, and then we recover f completely on all of $\mathbb{R}$. This is pretty amazing, and it is also related to the Heisenberg uncertainty principle. The theorem is proven by expanding $\hat{f}$ in a Fourier series, which can be done because $\hat{f}$ lives on a bounded interval. Then we do some clever manipulations and invoke the FIT to prove the theorem. It is a beautiful interplay between Fourier series and the Fourier transform!

9.8.1 The sampling theorem

Theorem 122 (Sampling Theorem)**.** *Let $f \in L^2(\mathbb{R})$. We take the definition of the Fourier transform of f to be*

$$\int_{\mathbb{R}} e^{-ix\xi} f(x)dx,$$

and we then assume that there is $L > 0$ so that $\hat{f}(\xi) = 0$ $\forall \xi \in \mathbb{R}$ with $|\xi| > L$. Then:

$$f(t) = \sum_{n\in\mathbb{Z}} f\left(\frac{n\pi}{L}\right) \frac{\sin(n\pi - tL)}{n\pi - tL}.$$

Proof:

Idea!

Since the Fourier transform $\hat{f}$ has compact support, (meaning it lives inside a bounded interval and is zero everywhere else) we can expand it as a Fourier series.

We therefore have

$$\hat{f}(x) = \sum_{-\infty}^{\infty} c_n e^{in\pi x/L}, \quad c_n = \frac{1}{2L} \int_{-L}^{L} e^{-in\pi x/L} \hat{f}(x)dx.$$

Idea!

Use the FIT to express f in terms of its Fourier transform.

We therefore have

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ixt} \hat{f}(x) dx = \frac{1}{2\pi} \int_{-L}^{L} e^{ixt} \hat{f}(x) dx.$$

On the left we have used the fact that $\hat{f}$ is supported in the interval $[-L, L]$, thus the integrand is zero outside of this interval, so we can throw that part of the integral away.

Idea!

Substitute the Fourier expansion of $\hat{f}$ into the integral.

So, we have

$$f(t) = \frac{1}{2\pi} \int_{-L}^{L} e^{ixt} \sum_{-\infty}^{\infty} c_n e^{in\pi x/L} dx.$$

From here until the end of the proof, we will essentially just be computing. The coefficients

$$c_n = \frac{1}{2L} \int_{-L}^{L} e^{-in\pi x/L} \hat{f}(x) dx = \frac{1}{2L} \int_{\mathbb{R}} e^{ix(-n\pi/L)} \hat{f}(x) dx = \frac{2\pi}{2L} f\left(\frac{-n\pi}{L}\right).$$

In the second equality we have used the fact that $\hat{f}(x) = 0$ for $|x| > L$, so by including that part we don't change the integral. In the third equality we have used the FIT!!! So, we now substitute this into our formula above for

$$f(t) = \frac{1}{2\pi} \int_{-L}^{L} e^{ixt} \sum_{-\infty}^{\infty} \frac{\pi}{L} f\left(\frac{-n\pi}{L}\right) e^{in\pi x/L} dx$$

This is approaching the form we wish to have in the theorem, but the argument of the function f has a pesky negative sign. That can be remedied by switching the order of summation, which does not change the sum, so

$$f(t) = \frac{1}{2L} \int_{-L}^{L} e^{ixt} \sum_{-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) e^{-in\pi x/L} dx.$$

We may also interchange the summation with the integral[5]

$$f(t) = \frac{1}{2L} \sum_{-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) \int_{-L}^{L} e^{x(it - in\pi/L)} dx.$$

We then compute

$$\int_{-L}^{L} e^{x(it-in\pi/L)} dx = \frac{e^{L(it-in\pi/L)}}{i(t - n\pi/L)} - \frac{e^{-L(it-in\pi/L)}}{i(t - n\pi/L)} = \frac{2i}{i(t - n\pi/L)} \sin(Lt - n\pi).$$

[5] None of this makes sense pointwise; we are working over L^2. The key property which allows interchange of limits, integrals, sums, derivatives, etc is *absolute convergence*. This is the case here because elements of L^2 have $\int |f|^2 < \infty$. That is precisely the type of absolute convergence required.

Substituting,

$$f(t) = \sum_{-\infty}^{\infty} f\left(\frac{n\pi}{L}\right) \frac{\sin(Lt - n\pi)}{Lt - n\pi}.$$

Of course my dyslexia has ended up with things being backwards, but it is not a problem because sine is odd so

$$\sin(Lt - n\pi) = -\sin(n\pi - Lt),$$

so

$$\frac{\sin(Lt - n\pi)}{Lt - n\pi} = \frac{-\sin(n\pi - Lt)}{Lt - n\pi} = \frac{\sin(n\pi - Lt}{n\pi - Lt}.$$

9.8.2 Discrete and fast Fourier transforms

We have seen that computing the Fourier transform is not the easiest thing in the world. The example with the Gaussian involving all those tricks: completing the square, complex analysis and contour integral is a nice and easy case. However, in the *real world* you may come across functions and not know how to compute the Fourier transform by hand, nor be able to find it in BETA. Or it could simply never have been computed analytically. In this case you may compute something called the *discrete Fourier transform.*

We start with a function, $f(t)$, and think of analyzing $f(t)$ as *time analysis,* whereas analyzing $\hat{f}(\xi)$ as *frequency analysis.* We shall consider a finite dimensional Hilbert space:

$$\mathbb{C}^N = \left\{ (s_n)_{n=0}^{N-1}, \quad s_n \in \mathbb{C}, \quad \langle (s_n), (t_n) \rangle := \sum_{n=0}^{N-1} s_n \overline{t_n} \right\}.$$

Now let

$$e_k(n) := \frac{e^{2\pi i k n/N}}{\sqrt{N}}.$$

Proposition 123. *Let*

$$e_k := (e_k(n))_{n=0}^{N-1}.$$

Then

$$\{e_k\}_{k=0}^{N-1}$$

are an ONB of $\mathbb{C}^N$.

Proof: We simply compute. It is so cute and discrete!

$$\langle e_k, e_j \rangle = \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i k n/N} e^{-2\pi i j n/N} = \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i (k-j) n/N}.$$

If $j = k$ the terms are all 1, and so the total is N which divided by N gives 1. Otherwise, we may without loss of generality assume that $k > j$ (swap names if not the case). Then we are staring at a geometric series! We know how to sum it

$$\sum_{n=0}^{N-1} e^{2\pi i (k-j) n/N} = \frac{1 - e^{2\pi i (k-j) N/N}}{1 - e^{2\pi i (k-j)/N}} = 0.$$

Here it is super important that $k - j$ is a number between 1 and $N - 1$. We know this because $0 \le j < k \le N - 1$. Hence when we subtract j from k, we get something between 1 and $N - 1$. So we are not dividing by zero!

Now we shall fix T small and N large and look at $f(t)$ just on the interval $[0,(N-1)T]$. Let

$$f(t_n) := f(nT), \quad t_n = nT.$$

Basically, we're going to identify f with an element of $\mathbb{C}^N$, namely

$$(f(t_n))_{n=0}^{N-1}.$$

Definition 124. Let

$$w_k := \frac{2\pi k}{NT}.$$

The discrete Fourier transform of f at w_k is defined to be

$$F(w_k) := \langle (f(t_n)), e_k \rangle = \sum_{n=0}^{N-1} \frac{f(t_n)e^{-2\pi i kn/N}}{\sqrt{N}}.$$

This can also be written as

$$\sum_{n=0}^{N-1} \frac{f(t_n)e^{-iw_k t_n}}{\sqrt{N}}.$$

Example 125. *One of the fun facts about the discrete Fourier transform is that we can Fourier transform functions which are neither in $\mathcal{L}^2$ nor in $\mathcal{L}^1$. For example, let's compute the discrete Fourier transform of*

$$f(x) = x, \quad T = \frac{1}{10}, \quad N = 5.$$

So, we identify f with the vector

$$(0, 0.1, 0.2, 0.3, 0.4).$$

Then,

$$F(w_k) := \sum_{n=0}^{4} \frac{ne^{-2\pi i kn/5}}{10\sqrt{5}}.$$

So, we identify the Fourier transform of f with the vector

$$\left(\sum_{n=0}^{4} \frac{n}{10\sqrt{5}}, \sum_{n=0}^{4} \frac{ne^{-2\pi i n/5}}{10\sqrt{5}}, \sum_{n=0}^{4} \frac{ne^{-4\pi i n/5}}{10\sqrt{5}}, \sum_{n=0}^{4} \frac{ne^{-6\pi i n/5}}{10\sqrt{5}}, \sum_{n=0}^{4} \frac{ne^{-8\pi i n/5}}{10\sqrt{5}} \right).$$

Proposition 126. *We have the inversion formula*

$$f(t_n) = \sum_{k=0}^{N-1} F(w_k)e_n(k) = \langle (F(w_k)), \overline{e}_n \rangle.$$

Proof: We simply compute. By definition

$$\langle (F(w_k)), \overline{e}_n \rangle = \sum_{k=0}^{N-1} F(w_k)e_n(k).$$

Now, we insert the definition of $F(w_k)$ which gives us another sum, so we use a different index there. Hence we have

$$\sum_{k=0}^{N-1} \sum_{m=0}^{N-1} \frac{f(t_m)e^{-iw_k t_m}}{\sqrt{N}} \frac{e^{2\pi i kn/N}}{\sqrt{N}} = \frac{1}{N} \sum \sum f(t_m)e^{-2\pi i km/N} e^{2\pi i kn/N}$$

$$= \frac{1}{N} \sum \sum f(t_m) e^{2\pi i k(n-m)/N} = \frac{1}{N} \sum_{m=0}^{N-1} f(t_m) \sum_{k=0}^{N-1} e^{2\pi i k(n-m)/N}$$

$$= \sum_{m=0}^{N-1} f(t_m) \sum_{k=0}^{N-1} \frac{e^{-2\pi i k m/N}}{\sqrt{N}} \frac{\overline{e^{-2\pi i k n/N}}}{\sqrt{N}} = \sum_{m=0}^{N-1} f(t_m) \langle e_m, e_n \rangle.$$

By the proposition we just proved before,

$$\langle e_m, e_n \rangle = \delta_{n,m} = \begin{cases} 0 & n \neq m \\ 1 & n = m. \end{cases}$$

So, the only term which survives is when $m = n$, and so we get

$$f(t_n).$$

Example 127. *Now, let's see if the inversion formula actually works for our example... First, we should have*

$$\sum_{k=0}^{4} F(w_k) e_0(k) = \sum_{k=0}^{4} \sum_{n=0}^{4} \frac{n e^{-2\pi i k n/5}}{10\sqrt{5}} \frac{1}{\sqrt{5}}$$

$$= \frac{1}{50} \sum_{n=0}^{4} n \sum_{k=0}^{4} e^{-2\pi i k n/5} = \frac{1}{50} \sum_{n=1}^{4} \frac{1 - e^{-2\pi i n}}{1 - e^{-2\pi i n/5}} = 0 = f(t_0).$$

Let's try another value:

$$\sum_{k=0}^{4} F(w_k) e_1(k) = \sum_{k=0}^{4} \sum_{m=0}^{4} \frac{m e^{-2\pi i k m/5}}{10\sqrt{5}} \frac{e^{2\pi i k/5}}{\sqrt{5}}$$

$$= \frac{1}{50} \sum_{n=1}^{4} n \sum_{k=0}^{4} e^{-2\pi i k(n-1)/5}.$$

For $n = 2, 3, 4$, the sum over k gives

$$\frac{1 - e^{-2\pi i (n-1)}}{1 - e^{-2\pi i (n-1)/5}} = 0.$$

For $n = 1$, the sum over k gives 5. Thus, the only term that survives is the term with $n = 1$, for which we obtain

$$\frac{1}{50}(1)(5) = \frac{1}{10} = f(t_1).$$

So, it is indeed working as it should. This is rather tedious, however.

Now, we can see this as matrix multiplication. In the discrete Fourier transform, we sampled f at the finitely many points $t_0, \ldots, t_{N-1}$. We therefore identify f with a vector

$$\begin{bmatrix} f(t_0) \\ f(t_1) \\ \vdots \\ f(t_{N-1}) \end{bmatrix}.$$

Similarly, the Fourier transform can be identified with the vector:

$$\begin{bmatrix} F(w_0) \\ F(w_1) \\ \dots \\ F(w_{N-1}) \end{bmatrix}.$$

This vector is the product of the matrix

$$\begin{bmatrix} \bar{e}_0 & \bar{e}_1 & \dots \bar{e}_{N-1} \end{bmatrix}$$

whose columns are

$$\bar{e}_n = \frac{1}{\sqrt{N}} \begin{bmatrix} e^0 \\ e^{-2\pi i n/N} \\ e^{-2\pi i(2)n/N} \\ \dots e^{-2\pi i k n/N} \\ \dots \\ e^{-2\pi i n(N-1)/N} \end{bmatrix}$$

together with the vector

$$\begin{bmatrix} f(t_0) \\ f(t_1) \\ \dots \\ f(t_{N-1}) \end{bmatrix}$$

That is

$$\begin{bmatrix} F(w_0) \\ F(w_1) \\ \dots \\ F(w_{N-1}) \end{bmatrix} = \begin{bmatrix} \bar{e}_0 & \bar{e}_1 & \dots \bar{e}_{N-1} \end{bmatrix} \begin{bmatrix} f(t_0) \\ f(t_1) \\ \dots \\ f(t_{N-1}) \end{bmatrix}$$

This entails a LOT of calculations. We can speed it up by being clever. Many calculations are repeated in fact. Assume that $N = 2^X$ for some giant power X. The idea is to split up into even and odd terms. We do this:

$$F(w_k) = \frac{1}{\sqrt{N}} \left[\sum_{j=0}^{\frac{N}{2}-1} f(t_{2j}) e^{-2\pi i k(2j)/N} + \sum_{j=0}^{\frac{N}{2}-1} f(t_{2j+1}) e^{-2\pi i k(2j+1)/N} \right].$$

We introduce the slightly cumbersome notation:

$$e_N^k(n) = e^{-2\pi i k n/N}.$$

Then,

$$e_N^k(2j) = e^{-2\pi i k(2j)/N} = e^{-2\pi i k j/(N/2)} = e_{N/2}^k(j).$$

Now we only need an $\frac{N}{2} \times \frac{N}{2}$ matrix! You see, writing this way,

$$F(w_k) = \frac{1}{\sqrt{N}} \left[\sum_{j=0}^{\frac{N}{2}-1} f(t_{2j}) e_{N/2}^k(j) + e_N^k(1) \sum_{j=0}^{\frac{N}{2}-1} f(t_{2j+1}) e_{N/2}^k(j) \right].$$

We can repeat this many times because N is a power of 2. We just keep chopping in half. If we do this as many times as possible, we will need to do on the order of

$$\frac{N}{2} \log_2(N)$$

computations. This is in comparison to the original method which had an $N \times N$ matrix and was thus on the order of N^2 computations. For example, if $N = 2^{10}$, then comparing $N^2 = 2^{20}$ to $\frac{N}{2} \log_2 N = 2^9 * 10$, we see that

$$\frac{2^{10} * 5}{2^{20}} = \frac{x}{100} \implies 100 * 2^{10} * 5 = 2^{20} x \implies 2^2 * 5^3 * 2^{10} 2^{-20} = x,$$

so

$$5^3 2^{-8} = x \approx 0.488.$$

This means the amount of work we are doing by using the FFT is less than 0.5% of the work done using the standard DFT. In other words, we save over 99.5% by doing the FFT. That's why it's called FAST.

9.9 Fourier transforms in mathematical physics

In signal processing, it is impossible for a signal to be both band-limited and time-limited. Mathematically, this means that it is impossible for both a function and its Fourier transform to vanish outside a finite interval, unless the function is identically zero, meaning there is no signal. Here we prove this mathematical fact, and therefore prove the real life consequence to signal processing.

9.9.1 A signal cannot be both band-limited and time-limited

It is convenient to introduce a common mathematical definition here.

Definition 128. A function is said to have *compact support* if it vanishes everywhere outside a compact set. The aforementioned compact set is *the support* of the function, and we say that the function is *supported on that set.* For functions from $\mathbb{R}^n$ to $\mathbb{C}$, compact sets of $\mathbb{R}^n$ are those sets which are closed and bounded. A function from $\mathbb{R}$ to $\mathbb{C}$ has compact support if and only if there is $L > 0$ such that $f(x) = 0$ for all x with $|x| > L$.

The following proposition is of its own independent interest.

Proposition 129. *Assume that $f \in \mathcal{L}^2(\mathbb{R})$, and that f has compact support. Then $f \in \mathcal{L}^1(\mathbb{R})$.*

Proof: We will use the Cauchy & Schwarz inequality. Assume that $f(x) = 0$ for all $x \in \mathbb{R}$ with $|x| > L$ for some L. Then,

$$\int_{\mathbb{R}} |f(x)| dx = \int_{\mathbb{R}} \chi_{[-L,L]}(x) |f(x)| dx \leq \sqrt{\int_{\mathbb{R}} \chi_{[-L,L]}(x)^2 dx} \sqrt{\int_{\mathbb{R}} |f(x)|^2 dx}$$

$$= \sqrt{2L} ||f||_{\mathcal{L}^2(\mathbb{R})} < \infty.$$

Above we have used the characteristic function that is 1 inside the interval $(-L, L)$ and zero outside of it

$$\chi_{[-L,L]}(x) = \begin{cases} 1 & |x| < L \\ 0 & |x| > L. \end{cases}.$$

We have also used the assumption that $f \in \mathcal{L}^2(\mathbb{R})$, so its $\mathcal{L}^2$ norm is finite.

We consider the Fourier transform but now we wish to evaluate the transform at complex values $z \in \mathbb{C}$, so we begin by investigating

$$F(z) := \int_{\mathbb{R}} e^{-ixz} f(x) dx = \int_{\mathbb{R}} e^{-ix \operatorname{Re}(z)} e^{x \operatorname{Im}(z)} f(x) dx.$$

Since f vanishes for $|x| > L$, we have the estimate

$$\int_{\mathbb{R}} |e^{x \operatorname{Im}(z)} f(x)| dx \leq e^{L|\operatorname{Im}(z)|} \int_{\mathbb{R}} |f(x)| dx < \infty,$$

because we have proven in the proposition that $f \in \mathcal{L}^1(\mathbb{R})$.

Consequently, the function $e^{x \operatorname{Im}(z)} f(x) \in \mathcal{L}^1(\mathbb{R})$ for any $z \in \mathbb{C}$ and therefore has a well defined Fourier transform, and so

$$F(z) = \int_{\mathbb{R}} e^{-ixz} f(x) dx = \int_{\mathbb{R}} e^{-ix \operatorname{Re}(z)} e^{x \operatorname{Im}(z)} f(x) dx$$

is a well defined function on $\mathbb{C}$. We would like to prove that it is entire by applying the dominated convergence theorem. Since

$$\frac{d}{dz} e^{-ixz} f(x) = (-ix) e^{-ixz} f(x),$$

we similarly estimate

$$|(-ix) e^{-ixz} f(x)| \leq L e^{L|\operatorname{Im}(z)|} |f(x)| \in \mathcal{L}^1(\mathbb{R}),$$

because

$$\int_{\mathbb{R}} L e^{L|\operatorname{Im}(z)|} |f(x)| dx = L e^{L|\operatorname{Im}(z)|} \int_{\mathbb{R}} |f(x)| dx = L e^{L|\operatorname{Im}(z)|} ||f||_{\mathcal{L}^1(\mathbb{R})}.$$

The dominated convergence theorem therefore allows us to differentiate under the integral and proves that $F(z)$ is an entire function of z. If $F(z)$ restricted to the real axis is zero outside of a bounded interval, then it is zero everywhere in $\mathbb{C}$ by the identity theorem. If this is the case, then the FIT immediately implies that the original function f was zero everywhere. Consequently, presuming f is non-zero, its Fourier transform (on the real axis) cannot be confined to a bounded interval. This always makes me think of peanut butter. If we start with a function that lives in a bounded interval, nice and neat and tidy, and then take its Fourier transform, the values of the function that were constrained within the bounded interval get smeared out over the whole real line. If you don't like peanut butter, imagine some other tasty spread. If you really like it, you want to spread it out over every corner of your cracker or bread, not missing a spot. This is what the Fourier transform does; it takes the function that is constrained to its little container (interval) and smears it out over the whole real line (bread).

Conversely, if we assume the Fourier transform of a function, say $\hat{g}$ lives in a bounded interval, we apply the same argument to the function

$$G(z) := \frac{1}{2\pi} \int_{\mathbb{R}} e^{izx} \hat{g}(x) dx.$$

This function is equal to $g(z)$ when $z \in \mathbb{R}$, but it is also a holomorphic function on all of $\mathbb{C}$ by the same arguments as above.

Exercise 130. *Work out all the details of the proof in this case, mimicking the arguments for the case when f lives in a bounded interval. That is, assuming that $\hat{g}$ lives in a bounded interval, prove that g lives in a bounded interval if and only if both g and $\hat{g}$ are identically equal to zero.*

9.9.2 The Heisenberg uncertainty principle

The fact that a signal cannot be both time-limited and band-limited is related to the uncertainty principle. In quantum mechanics, a particle like an electron moving in one dimension is described by a wave function $f : \mathbb{R} \to \mathbb{C}$, such that the $\mathcal{L}^2$ norm is equal to one. The value of $|f(x)|^2$ at $x \in \mathbb{R}$ is interpreted as the probability density that the particle is found at the point x. For this reason the $\mathcal{L}^2$ norm is equal to one, because the particle is *somewhere.* A slight modification of the Fourier transform of the wave function

$$\frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f(x) e^{-ixp/\hbar} dx =: \breve{f}(p) \tag{9.9.1}$$

Figure 9.6: The Fourier transform takes a function that is confined to a bounded interval, like a jar, and then smears it out over the entire real like, like peanut butter. Or, if you prefer a more exotic spread for your bread, try kaya! Kaya is a coconut jam made from a base of coconut milk, eggs and sugar. It is popular with Malaysians, Indonesians and Singaporeans. There are two varieties: nyonya kaya, which is is a lighter-green colour as shown here, and Hainanese kaya, which is a darker brown kind that uses caramelized sugar, changing the colour. Image license and source: creative commons 1.0 https://openclipart.org

is the probability density of the momentum of the particle. The position-momentum uncertainty principle states that we cannot know both the position and the momentum of the particle. We make a precise formulation of this that is known as Heisenberg's inequality and prove it here!

Definition 131. The dispersion of a function f about a point a is defined to be

$$\mathcal{D}_a f = \frac{\int_{\mathbb{R}} (x-a)^2 |f(x)|^2 dx}{\int_{\mathbb{R}} |f(x)|^2 dx},$$

whenever the integrals on the right side are defined.

The reason we call this the dispersion of f near a is because it measures how much f is concentrated near the point a. Let's take a quantum wave function, for example, so the denominator is equal to one. If the particle is very likely to be found near the point a, for example if f is zero everywhere outside of a small interval around a, $[a-\varepsilon, a+\varepsilon]$, then for $\alpha \in (a-\varepsilon, a+\varepsilon)$

$$\mathcal{D}_\alpha f = \int_{a-\varepsilon}^{a+\varepsilon} (x-\alpha)^2 |f(x)|^2 dx \leq \varepsilon^2,$$

because the interval over $\mathbb{R}$ of $|f|^2$ is equal to one. On the other hand, for α very far from a, let's say

$$|\alpha - a| > R + \varepsilon \implies \mathcal{D}_\alpha f = \int_{a-\varepsilon}^{a+\varepsilon} (x-\alpha)^2 |f(x)|^2 dx \geq R^2.$$

So the farther α is from where f lives, the larger $\mathcal{D}_\alpha f$ becomes. So, $\mathcal{D}_a$ is small if the support of f is concentrated around a, and $\mathcal{D}_a$ is large if the support of f is far away from a. Heisenberg's inequality states that f and $\hat{f}$ cannot both be concentrated near single points.

Theorem 132 (Heisenberg's inequality)**.** *For any $f \in \mathcal{L}^2(\mathbb{R})$ that is not identically zero, and for all $a, \alpha \in \mathbb{R}$,*

$$(\mathcal{D}_a f)(\mathcal{D}_\alpha \hat{f}) \geq \frac{1}{4}. \tag{9.9.2}$$

Proof: First, note that if $xf(x)$ is not in $\mathcal{L}^2(\mathbb{R})$, this means that

$$\int_{\mathbb{R}} x^2 |f(x)|^2 dx = \infty \implies \mathcal{D}_a(f) = \infty.$$

Note that since the integrand in the definition of $\mathcal{D}_a$ is non-negative, $\mathcal{D}_a$ is zero if and only if $f = 0$, and so if $xf(x)$ is not in $\mathcal{L}^2(\mathbb{R})$ then $f \neq 0$ implying that $\hat{f} \neq 0$, and therefore in this case the left side of (9.9.2) is infinite. So, henceforth we assume that $xf(x) \in \mathcal{L}^2(\mathbb{R})$.

Since smooth functions are dense in $\mathcal{L}^2$, by a limiting argument using the dominated convergence theorem, we assume that f is differentiable. We then have by the Plancherel theorem

$$\int_{\mathbb{R}} |f'(x)|^2 dx = \frac{1}{2\pi} \int_{\mathbb{R}} (\widehat{(f')}(\xi))^2 d\xi = \frac{1}{2\pi} \int_{\mathbb{R}} \xi^2 |\hat{f}(\xi)|^2 d\xi.$$

If $\xi \hat{f}(\xi)$ is not in $\mathcal{L}^2(\mathbb{R})$, then the right side of this equation is infinite, and by the same argument as above, the left side of (9.9.2) is also infinite. So, we may henceforth assume that we also have $f' \in \mathcal{L}^2(\mathbb{R})$.

For the sake of simplicity, let us now address the case in which $a = \alpha = 0$. Then we compute using integration by parts that for real numbers A and B

$$\int_A^B x\overline{f(x)} f'(x) dx = x|f(x)|^2 \Big|_A^B - \int_A^B (|f(x)|^2 + xf(x)\overline{f'(x)}) dx.$$

Re-arranging

$$\int_A^B |f(x)|^2 dx = x|f(x)|^2\Big|_A^B - \int_A^B x\overline{f(x)}f'(x)dx - \int_A^B xf(x)\overline{f'(x)}dx.$$

Since x is real-valued we note that

$$\int_A^B x\overline{f(x)}f'(x)dx = \int_A^B \overline{xf(x)\overline{f'(x)}}dx.$$

Adding a number to its complex conjugate we obtain twice the real part, and so

$$\int_A^B (x\overline{f(x)}f'(x) + xf(x)\overline{f'(x)})dx = \int_A^B 2\operatorname{Re}(x\overline{f(x)}f'(x))dx = 2\operatorname{Re}\int_A^B x\overline{f(x)}f'(x)dx.$$

Consequently,

$$\int_A^B |f(x)|^2 dx = x|f(x)|^2\Big|_A^B - 2\operatorname{Re}\int_A^B x\overline{f(x)}f'(x)dx.$$

Since $xf(x)$, f, and f' are all in $\mathcal{L}^2(\mathbb{R})$, the limits exist

$$\lim_{A\to-\infty, B\to\infty}\int_A^B |f(x)|^2 dx, \qquad \lim_{A\to-\infty, B\to\infty}\operatorname{Re}\int_A^B x\overline{f(x)}f'(x)dx.$$

In the last integral, this is the $\mathcal{L}^2$ scalar product of $xf(x)$ and $f'(x)$ which is finite since each of these functions is in $\mathcal{L}^2$. Consequently, re-arranging, the limits

$$\lim_{A\to-\infty} A|f(A)|^2, \qquad \lim_{B\to\infty} B|f(B)|^2$$

both exist. These limits must both be zero, because otherwise, if $A|f(A)|^2 \to c$ then for all large x, $|f(x)|^2 \approx \frac{c}{x}$ but the integral of $\frac{c}{x}$ does not converge as $x\to\infty$. The same argument shows that $B|f(B)|^2 \to 0$ as well. So, we obtain

$$\int_{\mathbb{R}} |f(x)|^2 dx = -2\operatorname{Re}\int_{\mathbb{R}} x\overline{f(x)}f'(x)dx.$$

By the Cauchy & Schwarz inequality,

$$\left(\int_{\mathbb{R}} |f(x)|^2 dx\right)^2 \le 4\left(\int_{\mathbb{R}} x^2|f(x)|^2 dx\right)\left(\int_{\mathbb{R}} |f'(x)|^2 dx\right) = 4\left(\int_{\mathbb{R}} x^2|f(x)|^2 dx\right)\left(\int_{\mathbb{R}} |f'(x)|^2 dx\right)$$

As we have seen with the Plancharel theorem calculation above,

$$\int_{\mathbb{R}} |f'(x)|^2 dx = \frac{1}{2\pi}\int_{\mathbb{R}} ((\widehat{f'})(\xi))^2 d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} \xi^2|\hat{f}(\xi)|^2 d\xi.$$

The Plancharel theorem also tells us that

$$\frac{1}{2\pi}\int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi = \int_{\mathbb{R}} |f(x)|^2 dx.$$

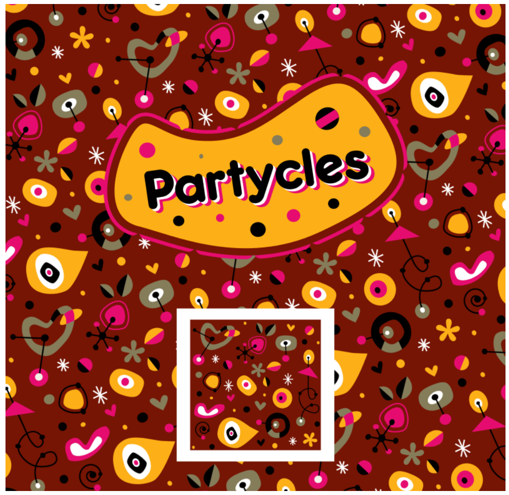


Figure 9.7: I could not resist these partycles. Quantum particles are rather like crazy partying partycles. They dance around so much that if you see a partycle in one location, you have no idea where it is headed to next. If you don't know its location but know its momentum, like the beat to which it dances, you have no idea where it's at. Image license and source: creative commons 1.0 https://openclipart.org

Consequently

$$\left(\int_{\mathbb{R}} |f(x)|^2 dx\right)^2 = \int_{\mathbb{R}} |f(x)|^2 dx \frac{1}{2\pi} \int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi \leq 4 \left(\int_{\mathbb{R}} x^2 |f(x)|^2 dx\right) \frac{1}{2\pi} \int_{\mathbb{R}} \xi^2 |\hat{f}(\xi)|^2 d\xi.$$

Since we have assumed that f is not identically zero, neither is its Fourier transform, and so we may divide obtaining

$$\frac{1}{4} \leq \frac{\int_{\mathbb{R}} x^2 |f(x)|^2 dx}{|f(x)|^2 dx} \frac{\int_{\mathbb{R}} \xi^2 |\hat{f}(\xi)|^2 d\xi}{\int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi} = \mathcal{D}_0(f)\mathcal{D}_0(\hat{f}).$$

To complete the proof, let $F(x) := e^{-i\alpha x} f(x+a)$. Then if f satisfies the hypotheses of the theorem, so does F. Moreover $\mathcal{D}_a f = \mathcal{D}_0 F$, and $\mathcal{D}_\alpha \hat{f} = \mathcal{D}_0 \hat{F}$. We therefore apply the same arguments above to obtain that

$$(\mathcal{D}_a f)(\mathcal{D}_\alpha \hat{f}) = (\mathcal{D}_0 F)(\mathcal{D}_0 \hat{F}) \geq \frac{1}{4}.$$

As a corollary we obtain the precise formulation of the position-momentum uncertainty principle for quantum particles.

Corollary 133 (Position-momentum uncertainty principle)**.** *Let $f(x)$ be a wave function associated to a quantum particle, that is a function in $\mathcal{L}^2(\mathbb{R})$ such that*

$$\int_{\mathbb{R}} |f(x)|^2 dx = 1,$$

so that $f(x)$ is the probability density of the quantum particle. Then

$$(\mathcal{D}_a f)(\mathcal{D}_\alpha \breve{f}) \geq \frac{\hbar^2}{4},$$

with $\check{f}$ the probability distribution of the momentum of the quantum particle as defined in (9.9.1).

Proof: Recalling the definition of $\check{f}$ in (9.9.1) we see that

$$\check{f}(p) = \frac{1}{\sqrt{2\pi\hbar}} \hat{f}(p/\hbar).$$

we compute that

$$\int_{\mathbb{R}} (x-a)^2 |\check{f}(x)|^2 dx = \frac{1}{2\pi\hbar} \int_{\mathbb{R}} (x-a)^2 |\hat{f}(x/\hbar)|^2 dx = \frac{\hbar}{2\pi} \int_{\mathbb{R}} (x/\hbar - a/\hbar)^2 |\hat{f}(x/\hbar)|^2 dx$$

$$= \frac{\hbar^2}{2\pi} \int (y - a/\hbar)^2 |\hat{f}(y)^2| dy.$$

In the last step we made the change of variables $y = \frac{x}{\hbar}$. Similarly we compute that

$$\int_{\mathbb{R}} |\check{f}(x)|^2 dx = \frac{1}{2\pi\hbar} \int_{\mathbb{R}} |\hat{f}(x/\hbar)|^2 dx = \frac{1}{2\pi} \int_{\mathbb{R}} |\hat{f}(y)|^2 dy.$$

Consequently

$$\mathcal{D}_a(\check{f}) = \frac{\hbar^2}{2\pi} \frac{\int (y-a/\hbar)^2 |\hat{f}(y)^2| dy}{\frac{1}{2\pi} \int_{\mathbb{R}} |\hat{f}(y)|^2 dy} = \hbar^2 \mathcal{D}_{a/\hbar}(\hat{f}).$$

By the Heisenberg inequality with $\alpha = a/\hbar$ we have

$$\mathcal{D}_a(f)\mathcal{D}_a(\check{f}) = \hbar^2 \mathcal{D}_a(f)\mathcal{D}_{a/\hbar}(\hat{f}) \geq \frac{\hbar^2}{4}.$$

The uncertainty principle is often considered one of the mysteries of quantum mechanics, but the principle itself is not so strange from the perspective of a wave. The strange part is the fact that quantum particles behave like waves! From the perspective of a wave, the inverse relationship between its temporal or spatial localization, that is the more localized it is in time the more spread out it is in space and vice versa, is quite natural. We see this with the peanut butter smearing effect of the Fourier transform: start with a function with compact support, then this support gets smeared out over the whole real line creating a Fourier transform that does *not* have compact support but rather is smeared out over the whole real line. Similarly, if a quantum particle is localized in space, then its momentum distribution is impossible to predict, and vice versa. You can know where the particle is, but you can't know where it's going. Or you could know where the particle is going, but then you don't know its starting point! This seems mysterious, but the real mystery is: why do quantum particles display these wave-like features?

9.9.3 Quantization of pseudodifferential operators

With all this investigation of quantum mechanics and quantum particles, I'd like to connect with the same 'quantum' terminology that is used in my field of research. Researchers in my field talk about 'quantizing' operators. When I first heard this, I was very confused. How on earth do you quantize, for example, a differential operator? And *why* on earth is it called *quantizing?* Well, it actually makes quite a lot of sense and connects with the rather strange fact that quantum particles can act like waves. The process of quantization makes sense for a very general class of operators known as pseudodifferential operators; all partial differential operators and ordinary differential operators are pseudodifferential operators but not the other way around. The process of quantization allows us to represent the operator with an integral that involves a wave function! So, we take the operator, and we make it wave-like, similar to the way quantum particles behave wave-like. I am not sure if the people who invented the terminology 'to quantize a pseudodifferential operator' justified their terminology with this explanation, but I think it makes sense. The quantization of a pseudodifferential operator is, for an operator Ψ,

$$\Psi u(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{i(x-y)\xi} \sigma(x, y, \xi) u(y) dy d\xi.$$

As an operator, Ψ acts on functions, and so for such a function Ψu is another function. The value of this function at the point x is given by the integral expression on the right. The function $\sigma(x, y, \xi)$ is a function that depends on the variables x, y, and ξ. We can interpret the product $e^{i(x-y)\xi}\sigma(xy, \xi)$ as a wave with the amplitude of the wave given by the function σ. This is intimately related to the Fourier transform, as we will illustrate with an example. Assume that f and its derivative are Fourier transformable. Then by the FIT,

$$f'(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \widehat{(f')}(\xi) d\xi = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{i(x-y)\xi} f'(y) dy d\xi = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} (i\xi) e^{i(x-y)\xi} f(y) dy d\xi,$$

having used integration by parts in the last step. So, with this example the quantization of the differential operator $\frac{d}{dx}$ is obtained by taking $\sigma = i\xi$, a function that depends only on ξ. Representing the differential operator in this way we have replaced differentiation with integration of the product of a wave function, $e^{-(x-y)\xi}$, a symbol function (in this case $\sigma = i\xi$), and the function that we are operating on (in this case f). This is pretty awesome. Essentially we can transform differential operators into integral operators that act by integrating the product of the function we're operating on together with a wave, viewing σ as the amplitude of the wave function $e^{i(x-y)\xi}$. Quantum particles have wave-like characteristics, and (pseudo)-differential operators are *quantized* by representing them as integration with the product of an wave function and its amplitude function.

9.10 Exercises

1. (EO 6a, b) Compute the Fourier transforms of:

$$\frac{t}{(t^2 + a^2)^2}, \quad \frac{1}{(t^2 + a^2)^2}.$$

2. [4, 7.2.8] Given $a > 0$ let $f(x) = e^{-x} x^{a-1}$ for $x > 0$, $f(x) = 0$ for $x \leq 0$. Show that $\hat{f}(\xi) = \Gamma(a)(1+i\xi)^{-a}$ where Γ is the Gamma function.

3. (EO 6d,e) Compute the Fourier transform of:

$$e^{-a|t|} \sin(bt), \quad (a, b > 0), \quad \frac{t}{t^2 + 2t + 5}.$$

4. [4, 7.2.12] For $a > 0$ let

$$f_a(x) = \frac{a}{\pi(x^2 + a^2)}, \quad g_a(x) = \frac{\sin(ax)}{\pi x}.$$

Use the Fourier transform to show that: $f_a * f_b = f_{a+b}$ and $g_a * g_b = g_{\min(a,b)}$.

5. (EO 12) Let

$$f(t) = \int_0^1 \sqrt{w} e^{w^2} \cos(wt) dw.$$

Compute

$$\int_{\mathbb{R}} |f'(t)|^2 dt.$$

6. [4, 7.2.13.b] Use Plancharel's theorem to compute:

$$\int_{\mathbb{R}} \frac{t^2}{(t^2+a^2)(t^2+b^2)} dt = \frac{\pi}{a+b}.$$

7. (EO 7) A function has Fourier transform

$$\hat{f}(\xi) = \frac{\xi}{1+\xi^4}.$$

   Compute

$$\int_{\mathbb{R}} t f(t) dt, \quad f'(0).$$

8. (EO 9) Compute (with help of Fourier transform)

$$\int_{\mathbb{R}} \frac{\sin(x)}{x(x^2+1)} dx.$$

9. (EO 15) Find a solution to the equation

$$u(t) + \int_{-\infty}^{t} e^{\tau - t} u(\tau) d\tau = e^{-2|t|}.$$

10. (EO 55) The function $f(x)$ is continuous and has Fourier transform equal to

$$\hat{f}(\xi) = \frac{\ln(1+\xi^2)}{\xi^2}.$$

    Determine $f(0)$ and $\int_{\mathbb{R}} f(x)dx$.

11. [4, 7.3.1] Use the Fourier transform to find a solution of the ordinary differential equation $u'' - u + 2g(x) = 0$ where $g \in \mathcal{L}^1(\mathbb{R})$. The solution obtained in this way is the one that vanishes at $\pm\infty$.

12. [4, 7.4.1.a.b] Compute the Fourier sine and cosine transforms of e^{-kx}. These are defined, respectively, to be

$$\mathcal{F}_s[f](\xi) = \int_{0}^{\infty} f(x)\sin(\xi x)dx, \quad \mathcal{F}_c[f](\xi) = \int_{0}^{\infty} f(x)\cos(\xi x)dx.$$

13. [4, 7.4.4] Solve the heat equation $u_t = k u_{xx}$ on the half line $x > 0$ with boundary conditions $u(x,0) = f(x)$ and initial condition $u(0,t) = 0$. Do the same for the inhomogeneous heat equation $u_t = k u_{xx} + G(x,t)$ with the same initial and boundary conditions.

14. [4, 7.4.6] Solve Laplace's equation $u_{xx} + u_{yy} = 0$ in the semi-infinite strip $x > 0$, $0 < y < 1$ with the boundary conditions $u_x(0,y) = 0$, $u_y(x,0) = 0$, $u(x,1) = e^{-x}$. Express the answer as a Fourier integral.

15. (EO 11) For the function

$$f(t) = \int_{0}^{2} \frac{\sqrt{w}}{1+w} e^{iwt} dw,$$

    compute

$$\int_{\mathbb{R}} f(t)\cos(t)dt, \quad \int_{\mathbb{R}} |f(t)|^2 dt.$$

16. (EO 14) Solve for u:

$$\int_0^\infty e^{-\tau} u(t-\tau)d\tau - \int_{-\infty}^0 e^{\tau} u(t-\tau)d\tau = \sqrt{3}u(t) - e^{-|t|}.$$

17. (EO 45) Find a bounded solution to

$$\begin{cases} u_t = ku_{xx}, & x \in \mathbb{R},\ t > 0, \\ u(x,0) = (1-2x^2)e^{-x^2}, & x \in \mathbb{R}. \end{cases}$$

18. [4, 7.4.7] Solve Laplace's equation $u_{xx} + u_{yy} = 0$ in the semi-infinite strip $x > 0$, $0 < y < 1$ with the boundary conditions $u(0,y) = 0$, $u(x,0) = 0$, $u(x,1) = e^{-x}$. Express the answer as a Fourier integral.

19. (EO 47) Assume that $f \in \mathcal{L}^2(\mathbb{R})$. Find a solution to

$$\begin{cases} u_{xx} + u_{yy} = 0 & x \in \mathbb{R},\ 0 < y < a, \\ u(x,0) = 0, & u(x,a) = f(x). \end{cases}$$

Show that

$$\int_{\mathbb{R}} |u(x)|^2 dx \le \int_{\mathbb{R}} |f(x)|^2 dx.$$

20. [4, 7.1.2] Let $f(x) = |x|^{-p}$ where $\frac{1}{2} < p < 1$. Show that f is in neither $\mathcal{L}^1(\mathbb{R})$ nor $\mathcal{L}^2(\mathbb{R})$ but that f can be expressed as the sum of an $\mathcal{L}^1(\mathbb{R})$ function and an $\mathcal{L}^2(\mathbb{R})$ function.

21. (EO 67) Compute the Fourier transform of the characteristic function for the interval (a,b) both directly and by using the known case for the interval $(-a,a)$.

22. [4, 7.3.2] Use the Fourier transform to derive the solution of the inhomogeneous heat equation $u_t = ku_{xx} + G(x,t)$ with initial condition $u(x,0) = f(x)$ (assume $f \in \mathcal{L}^2(\mathbb{R})$:

$$u(x,t) = f * H(x) + \int_{\mathbb{R}} \int_0^t G(y,s) K_{t-s}(x-y) ds dy.$$

Here the heat kernel (with the heat conductivity parameter k) is

$$H(x) = \frac{1}{\sqrt{4\pi kt}} e^{-x^2/4kt}.$$

23. [4, 7.1.1] Which of the following functions are in $\mathcal{L}^1(\mathbb{R})$? In $\mathcal{L}^2(\mathbb{R})$?
    (a) $\frac{\sin(x)}{|x|^{3/2}}$
    (b) $(1+x^2)^{-1/2}$
    (c) $\frac{1}{x^2-1}$
    (d) $\frac{1-\cos x}{x^2}$

24. [4, 7.1.3] Let

$$f(x) = \begin{cases} 1 & |x| < 1 \\ 0 & |x| > 1. \end{cases}$$

(a) Compute $f * f$ and $f * f * f$.

(b) Let $f_\varepsilon(x) = \varepsilon^{-1} f(\varepsilon^{-1}x)$ as in the BBC, and let $g(x) = x^3 - x$. Compute $f_\varepsilon * g$ and check that $f_\varepsilon * g \to 2g$ as $\varepsilon \to 0$. Note that $\int_f (x)dx = 2$.

25. [4, 7.1.4] Let $f(x) = e^{-x^2}$ and $g(x) = e^{-2x^2}$. Compute $f * g$. (Hint: complete the square in the exponent as done here in the text.)

26. [4, 7.1.5] For $t > 0$ let $f_t(x) = (4\pi t)^{-1/2} e^{-x^2/(4t)}$. Show that $f_t * f_s = f_{t+s}$. (Hint: the preceding exercise is helpful.)

27. [4, 7.1.6] For $t > 0$ let $f_t(x) = x^{1-t}/\Gamma(t)$ for $x > 0$ and $f_t(x) = 0$ for $x \leq 0$. Show that $f_t * f_s = f_{t+s}$. (Hint: you may need to call upon the beta function).

28. [4, 7.1.7] Show that for any $\delta > 0$ there is a function ϕ on $\mathbb{R}$ with the following properties: $\phi \in \mathcal{C}^\infty$, $0 \leq \phi \leq 1$, $\phi(x) = 1$ for all $x \in [0, 1]$, $\phi(x) = 0$ for all $x < -\delta$ and $x > 1 + \delta$. (Hint: start with an $f(x) = 1$ for $-\frac{1}{2}\delta \leq x \leq 1 + \frac{\delta}{2}$ and 0 otherwise. Show that $f * K_\varepsilon$ does the job for

$$K(y) = \begin{cases} C^{-1} e^{-1/(1-y^2)} & |y| < 1 \\ 0 & |y| \geq 1, \end{cases} \qquad C = \int_{-1}^{1} e^{-1/(1-y^2)} dy,$$

and for $\varepsilon < \frac{\delta}{2}$. This may seem weird but this type of function is very useful in quite a lot of analysis and geometry. One can do the same to create so-called *bump* functions that are 1 on an interval of our choice, quickly vanish to zero outside the interval, and are smooth. In particular these are often used to create *partitions of unity* in analysis and geometry.

29. [4, 7.1.8] Show that for any $f \in \mathcal{L}^2(\mathbb{R})$ and any $\delta > 0$ there is a function g such that $g \in \mathcal{C}^\infty$, g vanishes outside a finite interval, and $||f - g|| < \delta$. This is the $\mathcal{L}^2$ norm. (Hint: start with $F(x) = f(x)$ for $|x| < N$ and 0 otherwise. Show that $||F - f|| < \frac{\delta}{2}$ if N is sufficiently large. Then show that $g = F * K_\varepsilon$ does the job if K is as in the preceding exercise and ε is sufficiently small.) This might also seem weird, but it shows us that we can approximate $\mathcal{L}^2$ functions with smooth ones. That is extremely useful.

30. [4, 7.2.3] Assume that $g \in \mathcal{L}^1(\mathbb{R})$ has $\int_{\mathbb{R}} g(x)dx = 1$, and that $\hat{g} \in \mathcal{L}^1(\mathbb{R})$. Show that $\hat{g}(\delta\xi) \to 1$ as $\delta \to 0$ for all $\xi \in \mathbb{R}$. Show that for any continuous $f \in \mathcal{L}^1(\mathbb{R})$

$$\lim_{\delta \to 0} \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \hat{g}(\delta\xi) \hat{f}(\xi) d\xi = f(x).$$

What happens if f is only piecewise continuous?

31. [4, 7.2.6] Show that $\frac{\sin(x)}{x}$ is not in $\mathcal{L}^1(\mathbb{R})$. (Hint: Show that $\int_{(n-1)\pi}^{n\pi} x^{-1} |\sin(x)| dx > \frac{2}{n}$ for $n \in \mathbb{N}$.)

32. [4, 7.2.7] Assume that f is continuous and piecewise $\mathcal{C}^1$, with f and f' both in $\mathcal{L}^2(\mathbb{R})$. Show that $\hat{f} \in \mathcal{L}^1(\mathbb{R})$. (Hint: first show that $\int_{\mathbb{R}} (1+\xi^2) |\hat{f}(\xi)|^2 d\xi < \infty$ and then use the Cauchy-Schwarz inequality.)

33. [4, 7.2.9] Use the Residue Theorem to compute the Fourier transform of $(x^4 + 1)$. You should obtain

$$\frac{\pi}{\sqrt{2}} e^{-|\xi|^2/\sqrt{2}} \left(\cos \frac{\xi}{\sqrt{2}} + \sin \frac{|\xi|}{2} \right).$$

34. [4, 7.2.10] Let $f(x) = (\sinh(ax))/(\sinh(\pi x))$ for some $0 < a < \pi$. Use the Residue Theorem to show that

$$\hat{f}(\xi) = 2i \sum_{n \geq 1} (-1)^n e^{-n|\xi|} \sinh(ina).$$

Use the fact that $2\sinh(ina) = e^{ina} - e^{-ina}$ and sum the geometric series to show that

$$\hat{f}(\xi) = \frac{\sin(a)}{\cosh\xi + \cos a}.$$

35. [4, 7.2.11] Given $\nu > -\frac{1}{2}$ let $f(x) = (1-x^2)^{\nu-1/2}$ for $|x| < 1$ and 0 otherwise. Show that

$$\hat{f}(\xi) = 2^\nu\sqrt{\pi}\Gamma(\nu+1/2)\xi^{-\nu}J_\nu(\xi).$$

36. [4, 7.2.14] Let h_n be the n^{th} Hermite function, that is defined to be $e^{-x^2/2}H_n(x)$, where H_n is the n^{th} Hermite polynomial. Show that

$$\hat{h_n}(\xi) = \sqrt{2\pi}(-i)^n h_n(\xi).$$

(Hint: use induction on n.) This shows that the Hermite functions are eigenfunctions for the Fourier transform. In fact, they are a basis of eigenfunctions for $\mathcal{L}^2(\mathbb{R})$.

37. [4, 7.2.15] Let $\ell_n(x) = e^{-x/2}L_n^0(x)$ for $x > 0$, and 0 otherwise, where L_n^0 denotes the Laguerre polynomial, and let $\phi_n(\xi) = (2\pi)^{-1/2}\hat{\ell}_n(\xi)$. Show that

$$\phi_n(\xi) = \sqrt{\frac{2}{\pi}}\frac{(2i\xi-1)^n}{(2i\xi+1)^{n+1}}.$$

(Hint: plug the definition of L_n^0 into the formula defining $\hat{\ell}_n$ and integrate by parts n times).

38. [4, 7.3.3] Consider the wave equation $u_{tt} = c^2u_{xx}$ with initial conditions $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$. Assuming that all Fourier transforms in question exist, show that

$$\hat{u}(\xi,t) = \hat{f}(\xi)\cos(ct\xi) + \hat{g}(\xi)(c\xi)^{-1}\sin(ct\xi).$$

Invert the Fourier transform to obtain d'Alembert's formula for u.

39. [4, 7.3.4] Solve the Dirichlet problem in an infinite strip: $u_{xx} + u_{yy} = 0$ for $x \in \mathbb{R}$ and $0 < y < b$ with $u(x,0) = f(x)$ and $u(x,b) = g(x)$. (Hint: first do the case $f = 0$. The case $g = 0$ reduces to this one by the substitution $y \to b - y$, and the general case is obtained by superposition.)

40. [4, 7.3.5] Let S be the infinite cylinder of radius a, given in cylindrical coordinates by the equation $r = a$. Solve the problem

$$u_{rr} + r^{-1}u_r + u_{zz} = 0 \text{ inside } S, \quad u(a,z) = 1 \text{ if } |z| < \ell, \quad u(a,z) = 0 \text{ otherwise.}$$

Physically, the function u is the electrostatic potential inside S if the portion of S with $|z| < \ell$ is held at potential 1, and the rest of S is held at potential 0. Use the Fourier transform in z, and express the answer as an integral.

41. [4, 7.3.6] Suppose $f \in \mathcal{L}^2(\mathbb{R})$ represents a signal. Show that the best approximation to f in the $\mathcal{L}^2(\mathbb{R})$ norm among all signals that are band-limited in the interval $[-\Omega,\Omega]$ is

$$g_0(t) = (2\pi)^{-1}\int_{-\Omega}^{\Omega}\hat{f}(\omega)e^{i\omega t}d\omega.$$

That is show that $||g_0 - f|| \leq ||g - f||$ for all g such that $\hat{g}(\omega) = 0$ for $|\omega| > \Omega$. (Hint: use the Plancharel theorem.)

42. [4, 7.3.7] State and prove a version of the sampling theorem for signals whose Fourier transforms vanish outside an interval $[a,b]$. (Hint: consider $g(t) = e^{-i(b-a)t/2}f(t)$.)

43. [4, 7.4.1] Compute the Fourier cosine transform of $(1+x)e^{-x}$. Compute the Fourier sine transform of xe^{-x}.

44. [4, 7.4.2] Let f and g be in $\mathcal{L}^1(0,\infty)$. Show that the product of the Fourier sine transform of f and the Fourier cosine transform of g is the Fourier sine transform of h where

$$h(x) = \int_0^\infty f(y)\frac{g(|x-y|)-g(x+y)}{2}dy,$$

and that the product of the Fourier sine transform of f and the Fourier sine transform of g is

$$H(x) = \int_0^\infty f(y)\frac{\operatorname{sgn}(x-y)g(|x-y|)-g(x+y)}{2}dy,$$

where $\operatorname{sgn}(t)=1$ if $t>0$, and $\operatorname{sgn}(t)=-1$ if $t<0$.

45. [4, 7.4.3] Suppose f is continuous and piecewise smooth, and that both f and f' are in $\mathcal{L}^1(0,\infty)$. Show that

$$\mathcal{F}_c[f'](\xi) = \xi\mathcal{F}_s[f](\xi) - f(0), \quad \mathcal{F}_s[f'](\xi) = -\xi\mathcal{F}_c[f](\xi),$$

where $\mathcal{F}_c$ denotes the Fourier cosine transform, and $\mathcal{F}_s$ denotes the Fourier sine transform.

46. [4, 7.4.5] Solve the Dirichlet problem in the first quadrant: $u_{xx}+u_{yy}=0$ for $x,y>0$, with $u(x,0)=f(x)$ and $u(0,y)=g(y)$. (Hint: start with the special cases $f=0$ and $g=0$. Then use superposition.)

47. [4, 7.4.8] Find the steady-state temperature in a plate occupying the semi-infinite strip $x>0$, $0<y<1$ if the edges $y=0$ and $x=0$ are insulated, the edge at $y=1$ is maintained at temperature for 1 for $x<c$ and at temperature 0 for $x>c$, and the faces of the plate lose heat to the surroundings according to Newton's law with proportionality constant h. That is solve

$$u_{xx}+u_{yy}-hu=0, \quad u_x(0,y)=u_y(x,0)=0,$$

$$u(x,1)=1 \text{ if } x<c, \quad u(x,1)=0 \text{ if } x\geq c.$$

The answer may be expressed as an integral.

Chapter 10

The Laplace transform can be applied as long as we stay on the heavyside!

Another useful transform, known as the Laplace transform, is a very close relative of the Fourier transform. There is an important distinction, and that is that we can only apply the Laplace transform to functions that *stay on the heavyside.* The heavyside function

$$\Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0. \end{cases}$$

Therefore the heavyside function *stays on the heavyside,* which is the positive real axis, because it vanishes everywhere else. The Laplace transform is defined for functions that also *stay on the heavyside.*

Definition 134. Assume that

$$f(t) = 0 \quad \forall t < 0, \tag{10.0.1}$$

and that there exists $a, C > 0$ such that

$$|f(t)| \leq Ce^{at} \quad \forall t \geq 0. \tag{10.0.2}$$

Then for we define for $z \in \mathbb{C}$ with $\mathrm{Re}(z) > a$ the Laplace transform of f at the point z to be

$$\mathfrak{L}f(z) = \hat{f}(-iz) = \int_0^\infty f(t)e^{-zt}dt.$$

We may also use the notation

$$\widetilde{f}(z) = \mathfrak{L}f(z).$$

It is in general a good idea to check that this transform is well-defined, and for this we estimate

$$|\mathfrak{L}f(z)| \leq \int_0^\infty |f(t)e^{-zt}|dt \leq \int_0^\infty Ce^{at}|e^{-zt}|dt = \int_0^\infty e^{at}e^{-\mathrm{Re}(z)t}dt$$

$$= \frac{e^{t(a-\mathrm{Re}(z))}}{a - \mathrm{Re}(z)}\Bigg|_0^\infty = \frac{1}{\mathrm{Re}(z) - a}.$$

Above we have used the fact that

$$\left|e^{\text{complex number}}\right| = e^{\text{real part}}.$$

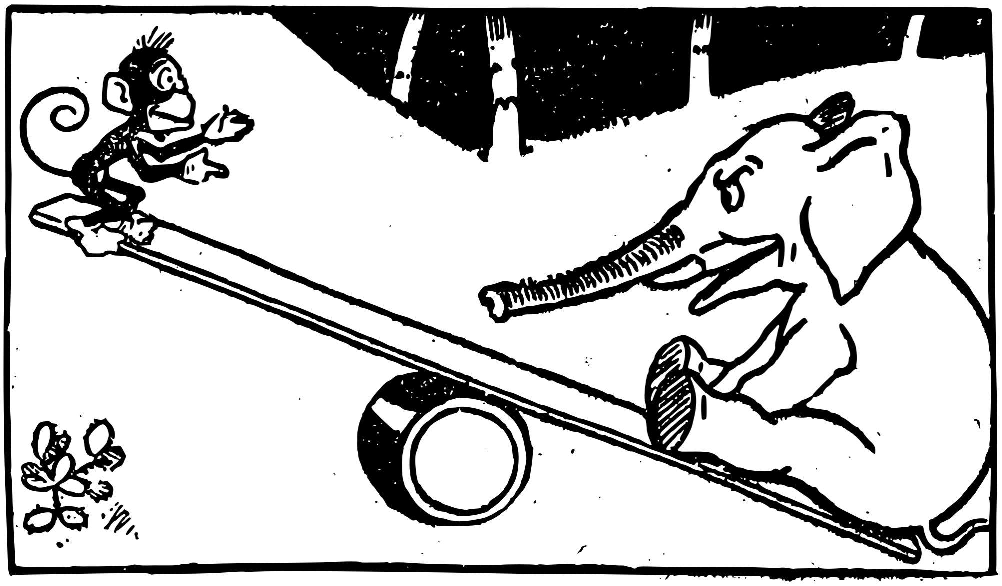


Figure 10.1: The Laplace transform can be applied, as long as we stay on the heavyside. Here, the elephant is on the right side, similar to the heavyside function that lives on the right side of the real line, in the sense that it vanishes everywhere else. So, it is only ever seen on the right side of the real line, the positive real axis. It lives on the heavyside! Image license and source: creative commons 1.0 https://openclipart.org.

Due to this beautiful convergence, $\mathfrak{L}f(z)$ is holomorphic in the half plane $\mathrm{Re}(z) > a$. This is because we may differentiate under the integral sign due to the absolute convergence of the integral. The assumption that $f(t) = 0$ for all negative t is not actually necessary, we could just make it so. For this purpose we define the *heavyside* function, commonly denoted by

$$\Theta(t) := \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}.$$

If we have some f defined on $\mathbb{R}$ which satisfies (10.0.2) but is not (10.0.1), we can apply the Laplace transform to Θf. Another thing which can happen is that we have a function which is only defined on $[0, \infty)$. In that case, we can just extend it to be identically zero on $(-\infty, 0)$.

10.1 Properties of the Laplace transform

We will become familiar with the Laplace transform by demonstrating some of its fundamental properties.

Proposition 135 (Properties of $\mathfrak{L}$)**.** *Assume f and g satisfy* (10.0.2) *and* (10.0.1)*, then*

1. $\mathfrak{L}f(x + iy) \to 0$ *as* $|y| \to \infty$ *for all* $x > a$.
2. $\mathfrak{L}f(x + iy) \to 0$ *as* $x \to \infty$ *for all* y.
3. $\mathfrak{L}(\Theta(t - a)f(t - a))(z) = e^{-az}\mathfrak{L}f(z)$.
4. $\mathfrak{L}(e^{ct}f(t))(z) = \mathfrak{L}f(z - c)$.
5. $\mathfrak{L}(f(at)) = a^{-1}\mathfrak{L}f(a^{-1}z)$.
6. **** If f is continuous and piecewise $\mathcal{C}^1$ on* $[0, \infty)$*, and f' satisfies* (10.0.2) *and* (10.0.1)*, then*

$$\mathfrak{L}(f')(z) = z\mathfrak{L}f(z) - f(0).$$

7. $\mathfrak{L}(\int_0^t f(s)ds)(z) = z^{-1}\mathfrak{L}f(z)$.
8. $\mathfrak{L}(tf(t))(z) = -(\mathfrak{L}f)'(z)$.
9. $\mathfrak{L}(f * g)(z) = \mathfrak{L}f(z)\mathfrak{L}g(z)$.
10. *If* $t^{-1}f(t)$ *satisfies* (10.0.1) *and* (10.0.2)*, then*

$$\mathfrak{L}(t^{-1}f(t))(z) = \int_z^\infty \mathfrak{L}f(w)dw.$$

The integral is any contour in the w*-plane which starts at* z *along which* $\operatorname{Im} w$ *stays bounded and* $\operatorname{Re} w \to \infty$.

Proof: There's a bunch of stars next to #6 because it's the reason the Laplace transform is useful for solving PDEs and ODEs. It's quite similar to how the Fourier transform takes in derivatives and spits out multiplication. Intuitively, this fact about $\mathfrak{L}$ should jive with the similar fact about $\mathcal{F}$ because well, the Laplace transform is just the Fourier transform evaluated at a complex number.

(1) The first statement

$$\mathfrak{L}f(z) = \int_0^\infty e^{-(x+iy)t}f(t)dt = \int_0^\infty e^{-xt}f(t)e^{-iyt}dt = \hat{g}(y),$$

for the function

$$g(t) = e^{-xt}f(t).$$

The Riemann-Lebesgue Lemma says that $\hat{g}(y) \to 0$ when $|y| \to \infty$.

(2) The second statement is more satisfying because we just compute and estimate directly. We did this estimate above already, where we got

$$|\mathfrak{L}f(z)| \leq \frac{1}{\operatorname{Re}(z) - a} \to 0 \text{ when } \operatorname{Re}(z) = x \to \infty.$$

(3) The third statement is also a direct computation:

$$\mathfrak{L}(\Theta(t-a)f(t-a))(z) = \int_0^\infty \Theta(t-a)f(t-a)e^{-zt}dt = \int_{-a}^\infty \Theta(s)f(s)e^{-z(s+a)}ds.$$

Above we did the substitution $s = t - a$ so $ds = dt$. Since f and the Heavyside function are zero for negative s, and the Heavyside function is 1 for positive s, this is

$$e^{-za}\int_0^\infty f(s)e^{-zs}ds = e^{-za}\mathfrak{L}f(z).$$

(4) Similarly, we directly compute

$$\mathfrak{L}(e^{ct}f)(z) = \int_0^\infty e^{ct}e^{-zt}f(t)dt = \int_0^\infty e^{-(z-c)t}f(t)dt = \mathfrak{L}f(z-c).$$

(5) Again no surprise, we compute

$$\mathfrak{L}(f(at))(z) = \int_0^\infty e^{-zt}f(at)dt = \int_0^\infty e^{-zs/a}f(s)\frac{ds}{a} = a^{-1}\mathfrak{L}f(z/a).$$

Here we used the substitution $s = at$ so $a^{-1}ds = dt$.

(6) Now we are finally getting to the important one:

$$\mathfrak{L}(f')(z) = \int_0^\infty e^{-zt} f'(t)dt = e^{-zt}f(t)\big|_0^\infty + \int_0^\infty ze^{-zt}f(t)dt.$$

We have used integration by parts above. By (10.0.2) and since $\mathrm{Re}(z) > a$, the limit as $t \to \infty$ is zero, and so we get

$$\mathfrak{L}(f')(z) = -f(0) + z\mathfrak{L}f(z).$$

Awesome.

(7) Next we define

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

Then, we use the preceding fact:

$$\mathfrak{L}(F')(z) = z\mathfrak{L}F(z) - F(0) = z\mathfrak{L}F(z).$$

Since $F' = f$ we get

$$z^{-1}\mathfrak{L}(f)(z) = \mathfrak{L}(\int_0^t f(s)ds)(z).$$

(8) Next, we compute:

$$\mathfrak{L}(tf(t))(z) = \int_0^\infty te^{-zt}f(t)dt = \int_0^\infty \frac{d}{dz}\left(-e^{-zt}\right) f(t)dt$$

$$= \frac{d}{dz}\left(-\int_0^\infty e^{-zt}f(t)dt\right) = -(\mathfrak{L}f)'(z).$$

Yes, we have used the absolute convergence of the integral to swap limits. It's legit yo.

(9) Nearing the finish line, we compute

$$\mathfrak{L}(f * g)(z) = \mathcal{F}(f * g)(-iz) = \hat{f}(-iz)\hat{g}(-iz) = \mathfrak{L}f(z)\mathfrak{L}g(z).$$

(10) Finally, note that by (10.0.2), if $t^{-1}f(t)$ satisfies this, then at the point $t = 0$ apparently the function f vanishes, so that the function $t^{-1}f(t)$ is well defined. So, don't panic about this point!!! We next define the holomorphic function

$$F(z) = \int_z^\infty \widetilde{f}(w)dw.$$

Since $\widetilde{f}(w) \to 0$ when $\mathrm{Re}(w) \to \infty$ and $\mathrm{Im}(w)$ stays bounded, the fundamental theorem of calculus says that

$$F'(z) = -\widetilde{f}(z).$$

On the other hand,

$$\frac{d}{dz}\int_0^\infty t^{-1}f(t)e^{-zt}dt = \int_0^\infty -f(t)e^{-zt}dt = -\widetilde{f}(z).$$

Figure 10.2: There is a saying that an elephant never forgets. This elephant is to remind us that the Laplace transform can be applied as long as we stay on the heavyside. If we have any doubt about this, we should include heavyside functions together with every function we Laplace transform, and just transform $f(t)\Theta(t)$ where $\Theta(t) = 0$ for $t < 0$ and $\Theta(t) = 1$ for $t > 0$, is the heavyside function.

Hence,

$$F(z) = \int_0^\infty t^{-1} f(t) e^{-zt} dt + c,$$

for some constant c. Since

$$\lim_{\operatorname{Re} z \to \infty} F(z) = 0 = \lim_{\operatorname{Re}(z) \to \infty} \int_0^\infty t^{-1} f(t) e^{-zt} dt \implies c = 0.$$

10.1.1 Laplace transform tables

Here we present some Laplace transforms; the function is on the left, its transform on the right.

Definition 136. There are two important functions known as the error function, denoted

$$\operatorname{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt,$$

and the complementary error function, denoted

$$\operatorname{erfc}(z) := 1 - \operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt.$$

1.	$f(t)$	$\mathfrak{L}f(z) = \widetilde{f(z)}$
2.	$\Theta(t-a)f(t-a)$	$e^{-az}\widetilde{f(z)})$
3.	$e^{ct}f(t)$	$\widetilde{f(z-c)}$
4.	$f(at)$	$a^{-1}\widetilde{f(a^{-1}z)}$
5.	$f'(t)$	$z\widetilde{f(z)} - f(0)$
6.	$f^{(k)}(t)$	$z^k\widetilde{f(z)} - \sum_0^{k-1} z^{k-1-j}f^{(j)}(0)$
7.	$\int_0^t f(s)ds$	$z^{-1}\widetilde{f(z)}$
8.	$tf(t)$	$-\widetilde{f}'(z)$
9.	$t^{-1}f(t)$	$\int_z^\infty \widetilde{f(w)}dw$
10.	$f * g(t)$	$\widetilde{f}(z)\widetilde{g}(z)$
11.	$t^\nu e^{ct}$	$\Gamma(\nu+1)(z-c)^{-\nu-1}$
12.	$(t+a)^{-1}$	$e^{az}\int_{az}^\infty \frac{e^{-u}}{u}du$
13.	$\sin(ct)$	$\frac{c}{z^2+c^2}$
14.	$\cos(ct)$	$\frac{z}{z^2+c^2}$

Table 10.1: Here $a > 0$ is constant and $c \in \mathbb{C}$.

10.2 Application to solving linear constant coefficient ordinary differential equations

By the properties of the Laplace transform we have demonstrated

$$\mathfrak{L}(f')(z) = z\mathfrak{L}f(z) - f(0).$$

If we differentiate again, we obtain

$$\mathfrak{L}(f'')(z) = z\mathfrak{L}(f')(z) - f'(0) = z\left(z\mathfrak{L}f(z) - f(0)\right) - f'(0) = z^2\mathfrak{L}f(z) - zf(0) - f'(0).$$

A pattern is emerging.

Proposition 137. *Assume that f and all derivatives of f up to the k^{th} satisfy the assumptions of Definition 134, so that we may apply the Laplace transform. Then*

$$\mathfrak{L}(f^{(k)})(z) = z^k\mathfrak{L}f(z) - \sum_{j=1}^{k} f^{(k-j)}(0)z^{j-1}.$$

15.	$\sinh(ct)$	$\frac{c}{z^2-c^2}$
16.	$\cosh(ct)$	$\frac{z}{z^2-c^2}$
17.	$\sin(\sqrt{at})$	$\sqrt{\pi a}(4z^3)^{-1/2}e^{-a/(4z)}$
18.	$t^{-1}\sin(\sqrt{at})$	$\pi \operatorname{erf}(\sqrt{a/(4z)}$
19.	$e^{-a^2t^2}$	$(\sqrt{\pi}/(2a))e^{z^2/(4a^2)}\operatorname{erfc}(z/2a)$
20.	$\operatorname{erf}(at)$	$z^{-1}e^{z^2/(4a^2)}\operatorname{erfc}(z/(2a))$
21.	$\operatorname{erf}(\sqrt{t})$	$(z\sqrt{z+1})^{-1}$
22.	$e^t\operatorname{erf}(\sqrt{t})$	$((z-1)\sqrt{z})^{-1}$
23.	$\operatorname{erfc}(a/(2\sqrt{t}))$	$z^{-1}e^{-a\sqrt{z}}$
24.	$t^{-1/2}e^{-\sqrt{at}}$	$\sqrt{\pi/z}e^{a/(4z)}\operatorname{erfc}(\sqrt{a/(4z)})$
25.	$t^{-1/2}e^{-a^2/(4t)}$	$\sqrt{\pi/z}e^{-a\sqrt{z}}$
26.	$t^{-3/2}e^{-a^2/(4t)}$	$2a^{-1}\sqrt{\pi}e^{-a\sqrt{z}}$
27.	$t^\nu J_\nu(t)$	$2^\nu\pi^{-1/2}\Gamma(\nu+1/2)(z^2+1)^{-\nu-1/2}$
28.	$J_0(\sqrt{t})$	$z^{-1}e^{-1/(4z)}$

Table 10.2: Here $a > 0$ is constant and $c \in \mathbb{C}$.

Proof: It is reasonable to prove this using induction, because we have already established the base case

$$\mathfrak{L}(f')(z) = z\mathfrak{L}f(z) - f(0).$$

Here $k = 1$ and the sum has only one term with $j = k = 1$. It works. Now we assume the above formula holds for k, and we show it then also holds for $k+1$. We compute

$$\mathfrak{L}(f^{(k+1)})(z) = \mathfrak{L}((f^{(k)})')(z) = z\mathfrak{L}(f^{(k)})(z) - f^{(k)}(0).$$

By the assumption that the formula holds for k, this is

$$z\left(z^k\mathfrak{L}f(z) - \sum_{j=1}^{k} f^{(k-j)}(0)z^{j-1}\right) - f^{(k)}(0).$$

This is

$$z^{k+1}\mathfrak{L}f(z) - \sum_{j=1}^{k} f^{(k-j)}(0)z^{j} - f^{(k)}(0).$$

Let us change our sum: let $j+1=l$. Then our sum is

$$\sum_{l=2}^{k+1} f^{k-(l-1)}(0)z^{l-1} = \sum_{l=2}^{k+1} f^{(k+1-l)}(0)z^{l-1}.$$

Observe that

$$f^{(k)}(0) = f^{k+1-1}(0)z^{1-1}.$$

Hence

$$-\sum_{j=1}^{k} f^{(k-j)}(0)z^{j} - f^{(k)}(0) = -\sum_{l=1}^{k+1} f^{(k+1-l)}(0)z^{l-1}.$$

So, we have computed

$$\mathfrak{L}(f^{(k+1)})(z) = z^{k+1}\mathfrak{L}f(z) - \sum_{l=1}^{k+1} f^{(k+1-l)}(0)z^{l-1}.$$

That is the formula for $k+1$, which is what we needed to obtain.

For this reason one can use $\mathfrak{L}$ to solve *all* linear constant coefficient ODEs which can be *non-homogeneous!* Any linear, constant coefficient ODE of order n can be re-arranged to take the form

$$\sum_{k=0}^{n} c_k u^{(k)}(t) = f(t).$$

In order for the solution to be unique, there must be specified initial conditions on u, that is the values of

$$u(0), u'(0), \ldots u^{(n-1)}(0)$$

should all be specified. The right side of the equation is a function $f(t)$ that is given. If it is the zero function, then the equation is called homogeneous. If that is not the case, then the equation is inhomogeneous. General inhomogeneous ODEs of this type are notoriously difficult to solve! Well, not anymore! With the power of the Laplace transform, here we give a recipe to solve any and every linear constant coefficient ODE, homogeneous or not!

The first step is to hit the entire equation with the Laplace transform. With our motto reminding us to stay on the heavyside like in Figure 10.1, we can imagine the Laplace transform is like a great big elephant that tramples across the whole equation:

$$\sum_{k=0}^{n} c_k \mathfrak{L}(u^{(k)})(z) = \widetilde{f}(z).$$

Let's write out the left side using our proposition.

First we have

$$c_0\widetilde{u}(z).$$

Then we have

$$c_1\left(z\widetilde{u}(z) - u(0)\right).$$

By our proposition, we computed that for $k \geq 1$,

$$\mathfrak{L}(c_k u^{(k)})(z) = c_k\left(z^k\widetilde{u}(z) - \sum_{j=1}^{k} u^{(k-j)}(0)z^{j-1}\right).$$

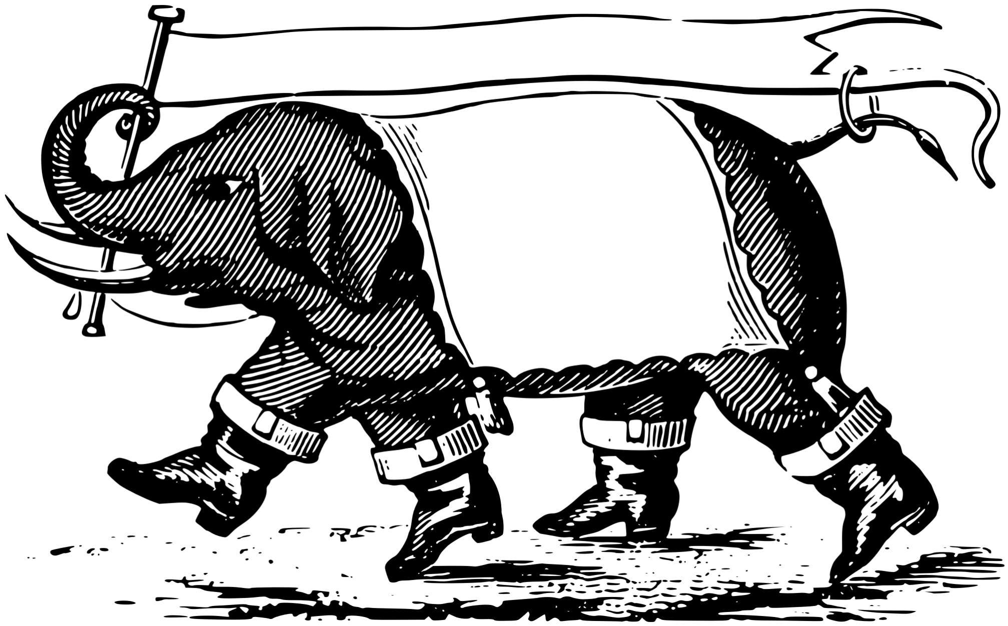


Figure 10.3: To solve a linear, constant coefficient ordinary differential equation, we start by trampling over the whole equation with the Laplace transform. It is like a great big elephant that runs across the whole equation, and it is not bothered in case the right side of the equation is non-zero. The elephant can trample an inhomogeneous ODE, that is one whose right side is non-zero, just as well as it can trample a homogeneous ODE, that is one whose right side is zero!

Therefore the left side of the ODE becomes

$$\begin{aligned} c_0\widetilde{u}(z) + \sum_{k=1}^{n} c_k \left(z^k \widetilde{u}(z) - \sum_{j=1}^{k} u^{(k-j)}(0) z^{j-1} \right) \\ = \sum_{k=0}^{n} c_k z^k \widetilde{u}(z) - \sum_{k=1}^{n} c_k \sum_{j=1}^{k} u^{(k-j)}(0) z^{j-1}. \end{aligned}$$

We therefore define two polynomials

$$P(z) := \sum_{k=0}^{n} c_k z^k,$$

$$Q(z) := -\sum_{k=1}^{n} c_k \sum_{j=1}^{k} u^{(k-j)}(0) z^{j-1}.$$

Our ODE has been LAPLACE-TRANSFORMED, or Laplace-trampled by the elephant in Figure 10.3 into

$$P(z)\widetilde{u}(z) + Q(z) = \widetilde{f}(z).$$

We can solve this for $\widetilde{u}(z)$:

$$\widetilde{u}(z) = \frac{\widetilde{f}(z) - Q(z)}{P(z)}.$$

Hence to get our solution $u(t)$ we just need to invert the Laplace transform of the right side, that is our solution will be

$$u(t) = \mathfrak{L}^{-1}\left(\frac{\widetilde{f}(z) - Q(z)}{P(z)}\right).$$

This is one of the reason that literal entire books have been written dedicated solely to computing Laplace transforms of functions; see for example [17] noting that it is over 400 pages.

10.3 Application of the Laplace transform to solving PDEs

It is a bit more work, but we can also use the Laplace transform to solve partial differential equations. The *telegraph equation,* generalizes both the homogeneous heat equation as well as the homogeneous wave equation, by taking certain choices of the constants in the equation below

$$u_{xx} = \alpha u_{tt} + \beta u_t + \gamma u.$$

The equation reduces to the heat equation if we take $\alpha = \gamma = 0$, and $\beta = 1$. It becomes the wave equation if $\beta = \gamma = 0$, and $\alpha = 1$. In its full generality, with the coefficients α, β and γ all non-zero, this equation describes an electromagnetic signal on a cable.

We wish to solve the problem on a half line with the following boundary and initial conditions:

$$u(0,t) = f(t), \quad u(x,0) = u_t(x,0) = 0.$$

Idea!

If we have a half-line problem with boundary condition at $x = 0$ that is *a function of* t try using the Laplace transform in the t variable.

We follow the hint and hit the whole PDE with the Laplace transform in the t variable; here comes the elephant! This gives

$$\widetilde{u}_{xx}(x,z) = \alpha\mathfrak{L}(u_{tt})(x,z) + \beta\mathfrak{L}(u_t)(x,z) + \gamma\widetilde{u}(x,z).$$

We use the properties of the Laplace transform and the initial conditions which say

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

so

$$\widetilde{u}_{xx}(x,z) = \alpha z^2\widetilde{u}(x,z) + \beta z\widetilde{u}(x,z) + \gamma\widetilde{u}(x,z).$$

This is simply

$$\widetilde{u}_{xx}(x,z) = \left(\alpha z^2 + \beta z + \gamma\right)\widetilde{u}(x,z).$$

It's a second order, linear, constant coefficient, homogeneous ODE for the x variable. Let

$$q = \sqrt{\alpha z^2 + \beta z + \gamma}.$$

Our solution to the ODE is of the form

$$\widetilde{u}(x,z) = a(z)e^{qx} + b(z)e^{-qx}.$$

We have that lovely BC at $x = 0$: $u(0,t) = f(t)$. Hence,

$$\widetilde{u}(0,z) = \widetilde{f}(z) \implies a(z) + b(z) = \widetilde{f}(z).$$

Note that here we are extending f to $(-\infty, 0)$ to be identically equal to zero so that we may Laplace transform it. Assume that $\mathrm{Re}(q) > 0$. (If this weren't the case, just swap q and $-q$). To be able to invert the Laplace transform and get the solution to our PDE, we will not want $\widetilde{u}(x,z) \to \infty$ when $x \to \infty$. Hence, we throw out the e^{qx} solution and just use

$$\widetilde{u}(x,z) = b(z)e^{-qx}.$$

Therefore, $b(z) = \widetilde{f}(z)$. So, our Laplace-transformed solution is

$$\widetilde{u}(x,z) = \widetilde{f}(z)e^{-qx}.$$

By the properties of the Laplace transform, if we can find $g(x,t)$ such that

$$\widetilde{g}(x,z) = e^{-qx},$$

then the solution to this PDE will be

$$u(x,t) = f * g(x,t) = \int_{\mathbb{R}} f(t-s)\Theta(t-s)g(x,s)\Theta(s)ds = \int_0^t f(t-s)g(x,s)ds. \tag{10.3.1}$$

The reason for those heavyside functions is that $f(*) = 0$ for $* < 0$ and $g(x,*) = 0$ for $* < 0$. To guarantee that this holds, we multiply $f(t-s)$ by $\Theta(t-s)$ and multiply $g(x,s)$ by $\Theta(s)$.

Now, recalling the definition of q, we are looking for

$$g(x,t) \text{ with } \widetilde{g(x,z)} = e^{-x\sqrt{\alpha z^2+\beta z+\gamma}}.$$

To find such a g, we would like to invert the Laplace transform.

10.3.1 Inverting the Laplace transform

The Laplace transform is closely related to the Fourier transform, and it is this fact, together with the FIT, that will light our way to the LIT (Laplace Inverse Theorem).

$$\widetilde{f}(z) = \int_0^\infty f(t)e^{-zt}dt = \int_0^\infty f(t)e^{-\operatorname{Re}(z)t - i\operatorname{Im}(z)t}dt.$$

For this reason, let's define

$$g(t) = e^{-\operatorname{Re}(z)t}f(t),$$

so we also have

$$f(t) = e^{\operatorname{Re}(z)t}g(t).$$

Then

$$\mathfrak{L}f(z) = \hat{g}(\operatorname{Im}(z)) = \int_{\mathbb{R}} f(t)e^{-\operatorname{Re}(z)}e^{-i\operatorname{Im}(z)t}dt,$$

because $f(t) = 0$ for all $t < 0$. Let's apply the FIT to the function, g:

$$g(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)e^{i\xi t}d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} \mathfrak{L}f(\operatorname{Re}(z)+i\xi)e^{i\xi t}d\xi.$$

To make this look less imposing, let us write $y = \xi$. So, we have

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}(\operatorname{Re}(z)+iy)e^{iyt}dy.$$

Since $f(t) = e^{\operatorname{Re}(z)t}g(t)$, we have

$$f(t) = e^{\operatorname{Re}(z)t}\frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}(\operatorname{Re}(z)+iy)e^{iyt}dy = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}(\operatorname{Re}(z)+iy)e^{\operatorname{Re}(z)t+iyt}dy.$$

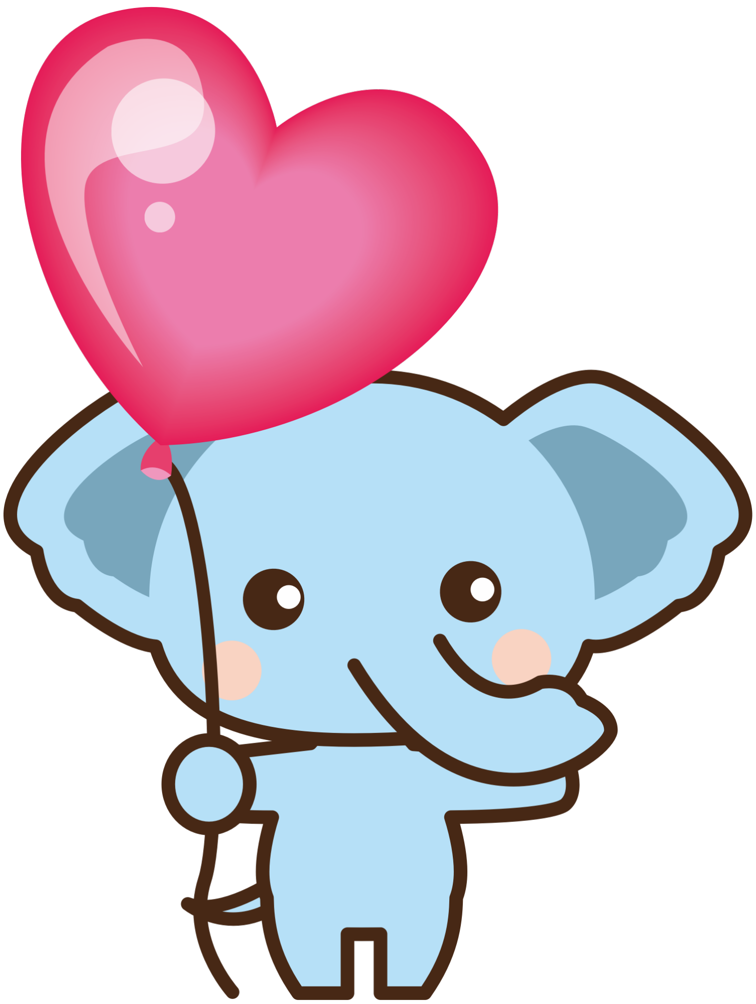


Figure 10.4: Inverting the Laplace transform is tricky business. This little elephant is here to remind us to stay on the heavyside as we apply the Fit to light our way to the LIT (Laplace Inverse Theorem).

We would like to understand this as a complex integral. If we parametrize the vertical path for $w \in \mathbb{C}$ with $\text{Re}(w) = \text{Re}(z)$ by $w = \text{Re}(z) + iy$, then $dw = idy$. We do not have an i. Hence, what we are computing is

$$f(t) = \frac{1}{2\pi i} \int_{\Gamma_z} \widetilde{f}(w) e^{wt} dw,$$

where Γ_z is the upward vertical path along the line $\text{Re}(w) = \text{Re}(z)$. This is the LIT: Laplace inversion formula:

$$f(t) = \frac{1}{2\pi i} \int_{\Gamma_z} \widetilde{f}(w) e^{wt} dw.$$

By definition of the Laplace transform, this should hold for $z \in \mathbb{C}$ with $\text{Re}(z) > a$ where a comes from the growth estimate on f, that is $|f(t)| \le Ce^{at}$ for all $t \ge 0$ for constants a and C. If we naively look at this equation, we see that the left side is *independent of* z. So, the right side ought to be as well. It is.

Theorem 138 (LIT)**.** *Assume that f is Laplace-transformable. Denote by $\widetilde{f}$ its Laplace transform. Then for $b > a$,*

$$f(t) = \frac{1}{2\pi i} \int_{b-i\infty}^{b+i\infty} \widetilde{f}(z) e^{zt} dz.$$

Conversely, assume that $F(z)$ is analytic in $\text{Re}(z) > a$. For $b > a$, $R > 0$, and $t \in \mathbb{R}$, let

$$f_{R,b}(t) = \frac{1}{2\pi i} \int_{b-iR}^{b+iR} F(z) e^{zt} dz.$$

Assume that for some $\alpha > 1/2$ and $C > 0$ we have

$$|F(z)| \le C(1 + |z|)^{-\alpha}, \quad \forall z \in \mathbb{C} \text{ with } \text{Re}(z) > a,$$

and assume that for some $b > a$, $f_{R,b}(t)$ converges pointwise as $R \to \infty$ to $f(t)$ for a Laplace transformable f. Then

$$\lim_{R \to \infty} f_{R,b}(t) = f(t) \quad \forall b > a,$$

and

$$F(z) = \mathfrak{L} f(z).$$

Proof: Let us draw and define a contour, with our amazing tikz skillz yo.
By assumption the function F is analytic inside the box, and e^{zt} is an entire function. Therefore

$$\int_{\Gamma_R} F(z) e^{zt} dz = 0.$$

So, we wish to show that the limit as $R \to \infty$ of the top and bottom integrals is zero. To obtain this, we either wave our hands like Folland or actually estimate:

$$\int_{b \pm iR}^{c \pm iR} |F(z)||e^{zt}| dz \le |c - b| e^{ct} \max_{b \le x \le c} \frac{C}{(1 + |x \pm iR|)^\alpha}.$$

Above we used the fact that between $b \pm iR$ and $c \pm iR$, $|e^{zt}| \le e^{ct}$ together with the estimate assumed on F. Next, we note that

$$|x \pm iR| = \sqrt{x^2 + R^2} \ge R.$$

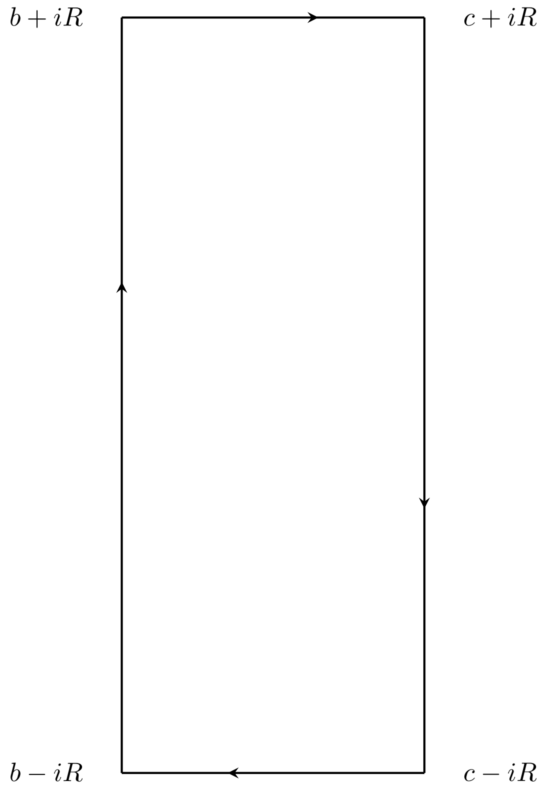


Figure 10.5: The contour over which we integral. Call the contour Γ_R. As one can see, we assume that $c > b$.

Therefore we estimate from above by

$$|c-b|e^{ct}\frac{C}{(1+R)^\alpha} \to 0 \text{ as } R\to\infty.$$

Therefore, if for some $b>a$,

$$\lim_{R\to\infty} f_{R,b}(t) = f(t),$$

this means that

$$\lim_{R\to\infty} \int\limits_{b-iR}^{b+iR} F(z)e^{zt}dz - \int\limits_{c-iR}^{c+iR} F(z)e^{zt}dz = 0.$$

To see this, observe that

$$\int\limits_{\Gamma_R} F(z)e^{zt}dz = 0 \quad \forall R.$$

Moreover, the top and bottom integrals go to zero as $R\to\infty$. Hence the sum of the left and right integrals also tends to zero as $R\to\infty$. So,

$$\lim_{R\to\infty} \int\limits_{b-iR}^{b+iR} F(z)e^{zt}dz = \lim_{R\to\infty} \int\limits_{c-iR}^{c+iR} F(z)e^{zt}dz \implies \lim_{R\to\infty} f_{R,b}(t) = f(t) = \lim_{R\to\infty} f_{R,c}(t).$$

Now, let us parametrize the complex integral. We use $\gamma(s) = b+is$ so $\dot\gamma(s) = ids$. Hence

$$\int\limits_{b-iR}^{b+iR} F(z)e^{zt}dz = \int\limits_{-R}^{R} F(b+is)e^{(b+is)t}ids = ie^{bt}\int\limits_{-R}^{R} F(b+is)e^{ist}ds.$$

Moreover, we have assumed that

$$\lim_{R\to\infty} f_{R,b}(t) = \lim_{R\to\infty} \frac{ie^{bt}}{2\pi i}\int\limits_{-R}^{R} F(b+is)e^{ist}ds = f(t)$$

so

$$\lim_{R\to\infty} \int\limits_{-R}^{R} F(b+is)e^{ist}ds = 2\pi e^{-bt}f(t).$$

Let us define here

$$g_{R,b}(s) = \begin{cases} F(b+is) & |s|\le R \\ 0 & |s|>R \end{cases}.$$

Then

$$\int\limits_{-R}^{R} F(b+is)e^{ist}ds = \int\limits_{\mathbb{R}} g_{R,b}(s)e^{ist}ds = \widehat{g_{R,b}}(-t).$$

Moreover,

$$\lim_{R\to\infty} \widehat{g_{R,b}(-t)} = 2\pi e^{-bt}f(t).$$

Similarly

$$\lim_{R\to\infty} \widehat{g_{R,b}(t)} = 2\pi e^{bt}f(-t).$$

On the other hand,

$$\lim_{R\to\infty} g_{R,b}(s) = F(b+is).$$

By the FIT,

$$F(b+it) = \frac{1}{2\pi}\int_{\mathbb{R}} 2\pi e^{bs} f(-s)e^{its}ds.$$

It is more natural to do a change of variables, letting $\sigma = -s$, so $d\sigma = -ds$, and we get

$$F(b+it) = \int_{\sigma=\infty}^{\sigma=-\infty} e^{-b\sigma} f(\sigma)e^{-it\sigma}(-d\sigma) = \int_{-\infty}^{\infty} e^{-\sigma(b+it)} f(\sigma)d\sigma$$

$$= \int_{0}^{\infty} e^{-\sigma(b+it)} f(\sigma)d\sigma = \mathfrak{L}f(b+it).$$

Here we use the fact that f satisfies the growth estimate needed to be Laplace transformable.

10.4 Computing an inverse Laplace transform to solve the heat equation

For the case in which our telegraph equation is the heat equation, we have $\alpha = \gamma = 0$, and $\beta = 1$. Consequently, the square rooted polynomial in z we had named q is of the simple form:

$$q = \sqrt{z}.$$

Our Laplace-transformed solution is:

$$\widetilde{f}(z)e^{-\sqrt{z}x}.$$

Since the Laplace transform turns convolutions into multiplication, we would like to find $g(x,t)$ so that

$$\widetilde{g}(x,z) = e^{-\sqrt{z}x}.$$

Then, the solution will be given as in (10.3.1).

We are therefore looking for $g(x,t)$ so that

$$\widetilde{g}(x,z) = e^{-\sqrt{z}x}.$$

If we try to apply the LIT directly, we should compute

$$\int_{b-i\infty}^{b+i\infty} e^{-x\sqrt{z}}e^{zt}dz.$$

Do you know how to integrate that? I do not. It is pretty scary looking. For starters, there is the $\sqrt{z}$. This really needs to be understood using the complex logarithm which is, as the name suggests, complex.

Idea!

Always be careful with $\log(z)$ in $\mathbb{C}$. It is *not* entire. It is a log. Logs come from trees which have branches. Complex logs always have branches and branch cuts. You have been warned.

So, since trying to compute the inverse Laplace transform directly seems impossible, let us try to make a reasonable guess at a function whose Laplace transform might be what we need to solve the heat equation. To solve the heat equation on $\mathbb{R}$ we used

$$e^{-x^2/(4t)}(4\pi t)^{-1/2}.$$

So, since the Laplace and Fourier transforms are closely related, and we are solving the heat equation on $[0,\infty)$, which is an unbounded interval, this is a good candidate. We shall compute its Laplace transform and see what we get. If we are super lucky, it will just give us the function we want. If we are less lucky, but still pretty lucky, the process of computing the Laplace transform together with the properties of the Laplace transform will show us how to get $g(x,t)$ whose Laplace transform is $\widetilde{g}(x,z) = e^{-\sqrt{z}x}$.

Let us therefore define:

$$\star = \int_0^\infty e^{-tz}e^{-x^2/(4t)}(4\pi t)^{-1/2}dt.$$

We are computing the Laplace transform of $\Theta(t)h(x,t)$ where

$$h(x,t) = e^{-x^2/(4t)}(4\pi t)^{-1/2}.$$

Now, we see that

$$\star = \int_0^\infty (4\pi t)^{-1/2}\exp\left(-(\sqrt{tz})^2 - \left(\frac{x}{2\sqrt{t}}\right)^2\right)dt.$$

We do the completing the square trick in the exponent:

$$\star = \int_0^\infty (4\pi t)^{-1/2}\exp\left(-\left(\sqrt{tz}-\frac{x}{2\sqrt{t}}\right)^2 - x\sqrt{z}\right)dt$$

$$= e^{-x\sqrt{z}}\int_0^\infty \frac{1}{2\sqrt{\pi t}}\exp\left(-\left(\sqrt{tz}-\frac{x}{2\sqrt{t}}\right)^2\right).$$

To compute this we need to use a very very clever trick. First, let us clean up our integral just a little bit to remove that pesky $\sqrt{t}$ which is getting divided. We let $s=\sqrt{t}$. Then

$$ds = \frac{dt}{2\sqrt{t}}$$

So,

$$\star = \frac{e^{-x\sqrt{z}}}{\sqrt{\pi}}\int_0^\infty e^{-(s\sqrt{z}-x/(2s))^2}ds.$$

Theorem 139 (Cauchy & Schlömilch transform)**.**

$$\int_0^\infty af((as-b/s)^2)ds = \int_0^\infty f(y^2)dy.$$

Proof: The proof is so clever.[1]

[1] I don't know if Cauchy and Schlömilch actually had anything to do with this formula. Oscar Schlömilch was elected a foreign member of the Royal Swedish Academy of Sciences in 1862. He was a German mathematician who lived from April 13, 1823 until February 7, 1901. Augustin-Louis Cauchy was a French mathematician who lived August 21, 1789 until May 23, 1857. Did they ever meet? Why is this named after them?

We do a substitution in the integral. Let $t = \frac{b}{as}$. Then

$$dt = -\frac{b}{as^2} ds \implies -\frac{as^2}{b} dt = ds.$$

We see that

$$t^2 = \frac{b^2}{a^2 s^2} \implies \frac{a^2 s^2}{b^2} = t^{-2} \implies \frac{as^2}{b} = \frac{b}{at^2}.$$

Next, when $s \to 0$ and $s > 0$ we see that $t \to \infty$. On the other hand, when $s \to \infty$, $t \to 0$. We also see that

$$as = \frac{t}{b}, \quad -\frac{b}{s} = -ta.$$

So, let us call

$$\heartsuit = \int_0^\infty af((as - b/s)^2) ds = \int_\infty^0 af((t/b - ta)^2) \left(-\frac{b}{at^2} \right) dt$$

$$= \int_0^\infty f((t/b - at)^2) \frac{b}{t^2} dt.$$

Note that

$$(t/b - at)^2 = (-(at - t/b))^2 = (at - t/b)^2.$$

Hence we have computed

$$\heartsuit = \int_0^\infty f((at - t/b)^2) \frac{b}{t^2} dt.$$

Therefore

$$2\heartsuit = \int_0^\infty af((as - b/s)^2) ds + \int_0^\infty f((at - t/b)^2) \frac{b}{t^2} dt$$

$$= a \int_0^\infty f((as - b/s)^2) ds + b \int_0^\infty f((as - b/s)^2) \frac{ds}{s^2}.$$

As a consequence,

$$\heartsuit = \frac{1}{2} \int_0^\infty f((as - b/s)^2) \left(a + \frac{b}{s^2} \right) ds.$$

Now we let

$$y = as - \frac{b}{s} \implies dy = \left(a + \frac{b}{s^2} \right) ds.$$

We note that when $s \to 0$, $y \to -\infty$, and on the flip side, when $s \to \infty$, $y \to \infty$. Therefore

$$\heartsuit = \frac{1}{2} \int_{-\infty}^\infty f(y^2) dy = \int_0^\infty f(y^2) dy,$$

since $f(y^2)$ is an even function.

We will use the Cauchy & Schlömilch transform with

$$a = \sqrt{z}, \quad b = \frac{x}{2}, \quad f(s) = e^{-s^2}.$$

Then, it says that

$$\int_0^\infty \sqrt{z}\exp(-(as - b/s)^2)ds = \int_0^\infty \sqrt{z}\exp\left(-\left(s\sqrt{z} - \frac{x}{2s}\right)^2\right)ds$$

$$= \int_0^\infty e^{-y^2}dy = \frac{\sqrt{\pi}}{2}.$$

Now we were computing

$$\star = \frac{e^{-x\sqrt{z}}}{\sqrt{\pi}} \int_0^\infty e^{-(s\sqrt{z}-x/(2s))^2}ds = \frac{e^{-x\sqrt{z}}}{\sqrt{\pi z}} \int_0^\infty \sqrt{z}e^{-(s\sqrt{z}-x/(2s))^2}ds$$

$$= \frac{e^{-x\sqrt{z}}}{2\sqrt{z}}.$$

So, we have computed

$$\mathfrak{L}\left(\Theta(t)h(x,t)\right)(z) = \frac{e^{-x\sqrt{z}}}{2\sqrt{z}}.$$

This is almost what we wanted, except for the $2\sqrt{z}$ in the denominator. Here we use the properties of the Laplace transform. Consider the function:

$$\int_z^\infty \frac{e^{-x\sqrt{w}}}{2\sqrt{w}}dw = -\frac{e^{-x\sqrt{w}}}{x}\bigg|_{w=z}^\infty = \frac{e^{-x\sqrt{z}}}{x}.$$

By the properties of the Laplace transform

$$\mathfrak{L}(t^{-1}f(t))(z) = \int_z^\infty \tilde{f}(w)dw.$$

So,

$$\mathfrak{L}(t^{-1}\Theta(t)h(x,t))(z) = \int_z^\infty \frac{e^{-x\sqrt{w}}}{2\sqrt{w}}dw = \frac{e^{-x\sqrt{z}}}{x}.$$

because we computed

$$\mathfrak{L}\left(\Theta(t)h(x,t)\right)(z) = \frac{e^{-x\sqrt{z}}}{2\sqrt{z}}.$$

We can simply multiply both sides by x to get

$$\mathfrak{L}(t^{-1}x\Theta(t)h(x,t))(z) = e^{-x\sqrt{z}}$$

as desired. Let us summarize this phenomenal calculation as a theorem for future reference.

Theorem 140. *The Laplace transform of*

$$g(x,s) := \frac{x}{s}\Theta(s)h(x,s), \quad h(x,s) = \frac{1}{\sqrt{4\pi s}}e^{-\frac{x^2}{4s}}, \quad \Theta(s) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

in the variable s is

$$\mathfrak{L}(g)(x,z) = e^{-x\sqrt{z}}.$$

Therefore going back to our problem, the solution

$$u(x,t) = (f(s) * (s^{-1} x \Theta(s) h(x,s))(t) = \int_{\mathbb{R}} f(t-s) g(x,s) ds$$

$$= \int_0^t \frac{f(t-s)}{2\sqrt{\pi} s^{3/2}} x e^{-\frac{x^2}{4s}} ds.$$

This is because f is zero for negative times since we are remembering to *stay on the heavyside.*

Remark 4. One of the things I love about this class is that you begin to approach actual research mathematics. I think that must be exciting for you, because calculus (envariabelanalys) is like 300 years old. Cauchy's complex analysis is also a few hundred years old. That's not very close to actual current math! Here is an example of how the Cauchy-Schlömilch transform is super awesome https://arxiv.org/abs/1004.2445. This paper is recent, showing that this transform is interesting even from a modern day research perspective!

10.5 The Laplace transform in mathematical biology: how many elephants?

We have seen how the Fourier transform can be used to solve integral equations for unknown functions, like

$$u(x) + \int_{\mathbb{R}} f(x-y) u(y) dy = g(x),$$

where u is the unknown function we wish to find, and f and g are specified. To solve this equation using the Fourier transform, we take the Fourier transform of both sides with respect to x, obtaining the transformed equation

$$\hat{u}(\xi) + \hat{f}(\xi)\hat{u}(\xi) = \hat{g}(\xi).$$

We solve this equation for $\hat{u}(\xi)$ and obtain

$$\hat{u}(\xi) = \frac{\hat{g}(\xi)}{1 + \hat{f}(\xi)}.$$

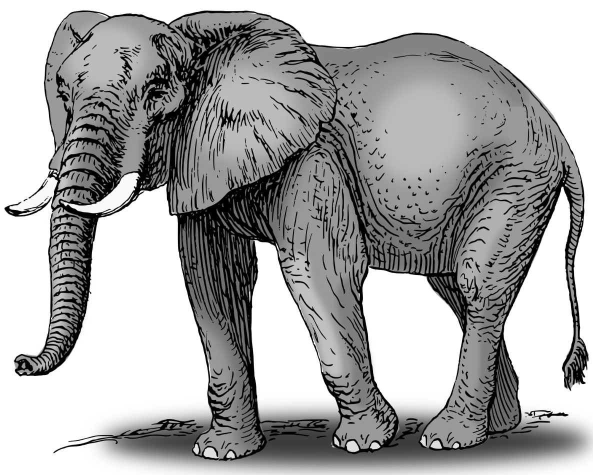
The function u is then obtained by taking the inverse Fourier transform, and so

$$u(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \frac{\hat{g}(\xi)}{1 + \hat{f}(\xi)}.$$

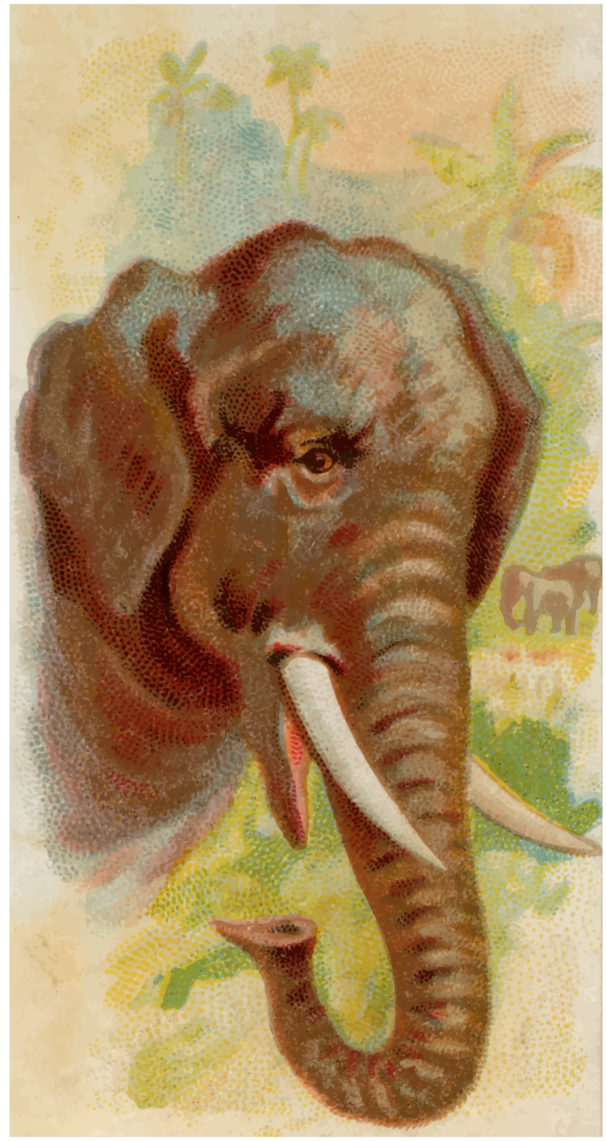
Since the Laplace transform is simply the Fourier transform evaluated at a complex value, for functions that live on the heavy side, that is vanish on the negative real axis, it is reasonable to expect that the Laplace transform could also be used to solve integral equations. A large class of integral equations are called *Volterra* equations, and have the general form

$$u(t) - \lambda \int_a^t K(t,s) u(s) ds = f(t).$$

The functions K and f would be specified, and the task at hand is to solve for the unknown function u. Equations of this type are ubiquitous in physics, chemistry, biology, finance, and engineering. Volterra equations were first studied in the beginning of the twentieth century but remain a hot research topic to this day. We explore one of Volterra's first such equations, that is known as the renewal equation, because it describes renewal processes. In biology, this could be birth and death, whereas in engineering, this could describe mechanical equipment wearing out and being replaced.



(a) An African elephant can be recognized due to its large ears in comparison to those of Asian elephants.



(b) Asian elephants can be distinguished from African elephants by their generally smaller ears.

Figure 10.6: African and Asian elephants are equally wonderful beings. They have the longest gestation period of any mammal: 22 months. Images are licensed under creative commons 1.0 obtained from `https://openclipart.org`.

Let's think about our heavyside elephants; this is quite important because currently Asian elephants are classified as endangered, and African elephants are vulnerable. Suppose at time $t = 0$ we have a community of N_0 elephants. We can set up an equation to predict the elephant population at later times if we collect some information:

1. For $t > 0$ the probability $p(t)$ that an elephant will live for at least t units of time after it is born.

2. At time $t = 0$, we need to know the ages of the elephants in our community. We call this $a(t)$.

3. For the sake of simplicity, we will assume the birth rate R is constant, and we need to know this information as well. We could relax this assumption by calculating the average birth rate over time, and if this average over long periods of time does not change drastically, then it is not really a loss of generality to assume the birth rate is just a constant.

At time t, the population will be comprised of the elephants from the community that have survived since $t = 0$, minus those who have died, and plus those cute new elephants that have been born. With all considerations in mind, the population

$$N(t) = N_0 \int_0^\infty p(s+t)a(s)ds + R\int_0^t p(t-s)N(s)ds.$$

Since the functions p and a are known from the information we collected, the first part of the equation can be calculated from these known functions by computing the integral. So, we call this part

$$n(t) := N_0 \int_0^\infty p(s+t)a(s)ds.$$

Then, our equation for N is

$$N(t) = n(t) + R\int_0^t p(t-s)N(s)ds.$$

Note that we can re-write the integral as a convolution by introducing some heavyside functions

$$\int_0^t p(t-s)N(s)ds = \int_{\mathbb{R}} p(t-s)\Theta(t-s)N(s)\Theta(s)ds.$$

This works because $\Theta(s)$ vanishes for all $s < 0$, and $\Theta(t-s)$ vanishes for all $t - s < 0 \iff t < s$. Then we use the properties of the Laplace transform to conclude that the transformed equation is simply

$$\mathfrak{L}N(z) = \mathfrak{L}n(z) + R\mathfrak{L}p(z)\mathfrak{L}N(z).$$

We solve the equation:

$$\mathfrak{L}N(z) = \frac{\mathfrak{L}n(z)}{1 - R\mathfrak{L}p(z)}.$$

Consequently, we obtain our equation for the elephant population is illuminated by the LIT

$$N(t) = \mathfrak{L}^{-1}\left(\frac{\mathfrak{L}n(z)}{1 - R\mathfrak{L}p(z)}\right)(t).$$

10.5.1 If the Laplace and Fourier transform are so closely related, why do we need both?

The Laplace and Fourier transforms apply in different contexts. They are therefore both useful and complementary; in contexts where we cannot apply the Fourier transform, we might be able to use Laplace, and vice-versa! For example, one reason to solve the Volterra equation for the elephant population is because populations can grow exponentially, and so the equation need not be Fourier transformable. A second reason is that the functions all depend on time, and therefore we can simply set everything to zero for negative time and guarantee that all functions in the equation *live on the heavyside.*

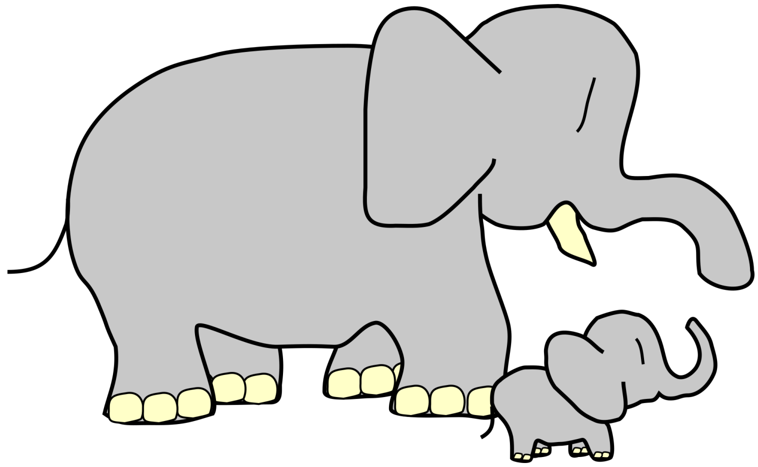


Figure 10.7: The Laplace transform has numerous applications; this is one of the reasons there are entire books dedicated to their calculation like [17]. Here we investigate the application to solving the renewal equation, to understand elephant populations.

Fourier transform can be used	Laplace transform can be used
Functions are in $\mathcal{L}^1(\mathbb{R})$ or $\mathcal{L}^2(\mathbb{R})$	Functions live on the heavyside and have at most exponential growth

Table 10.3: On the one hand, we can Fourier transform functions that live on the entire real line, whereas we can only Laplace transform functions that live on the heavyside, meaning they must be identically zero on the negative real line. On the other hand, we can Laplace transform functions that grow as much as $\sim Ce^{ax}$ as $x \to \infty$ for constants $a, C > 0$. So, this includes polynomials and functions that grow exponentially, none of which are Fourier-transformable. The Fourier and Laplace transform are complementary in this sense; sometimes we can use Fourier, sometimes we can use Laplace, and it depends on the precise problem at hand. We still need to do a 'sound check' to determine which method to use!

10.6 Exercises

1. [4, 8.4.2] Find the temperature in a semi-infinite rod, mathematicized as the half line $[0, \infty)$ if its initial temperature is 0 and the end temperature at $x = 0$ is held at temperature 42 for $0 < t < 1$ and thereafter held at temperature zero.

2. [4, 8.4.4] A semi-infinite rod is initially at temperature 1. Its end is in contact with a medium at temperature zero and loses heat according to Newton's law of cooling:
$$u_t = ku_{xx} \text{ for } x > 0,\ u(x,0) = 1,\ u_x(0,t) = cu(0,t).$$
Show that
$$\mathfrak{L}u(x,z) = \frac{1}{z} - \frac{c\sqrt{k}}{z(c\sqrt{k} + \sqrt{z})} e^{-x\sqrt{z/k}}.$$
Show that the temperature at the end is given by $u(0,t) = e^{c^2kt}\,\mathrm{erfc}(c\sqrt{kt})$, where erfc is the complementary error function. Hint: multiply and divide by $\sqrt{z} - c\sqrt{k}$.

3. [4, 8.4.1] Solve:
$$u_t = ku_{xx} - au, \quad x > 0, \quad u(x,0) = 0, \quad u(0,t) = f(t).$$

4. [4, 8.4.3] Consider heat flow in a semi-infinite rod when heat is supplied to the end at a constant rate c:
$$u_t = ku_{xx} \text{ for } x > 0, \quad u(x,0) = 0, \quad u_x(0,t) = -c.$$

With the aid of the computation:

$$\mathfrak{L}\left(\frac{1}{\sqrt{\pi t}}e^{-a^2/(4t)}\right)(z) = \frac{e^{-a\sqrt{z}}}{\sqrt{z}},$$

show that

$$u(x,t) = c\sqrt{\frac{k}{\pi}}\int_0^t s^{-1/2}e^{-x^2/(4ks)}ds.$$

By substituting

$$\sigma = \frac{x}{\sqrt{4ks}}$$

and then integrating by parts, show that

$$u(x,t) = c\sqrt{\frac{4kt}{\pi}}e^{-x^2/(4kt)} - cx\,\mathrm{erfc}\left(\frac{x}{\sqrt{4kt}}\right).$$

5. [4, 8.5.3] Solve the following equation for u

$$u(t) + 2\int_0^t u(t-s)\cos(as)ds = \sin(at), \quad a > 0.$$

6. [4, 8.5.4] Solve the following equation for u:

$$\int_0^t u(s)u(t-s)ds = t^5 e^{-3t}.$$

7. [4, 8.1.1] Compute the Laplace transforms of $\sinh(at)$ and $\cosh(at)$.
8. [4, 8.1.2] Compute the Laplace transform of $\cos^2 t$. (Hint: double angle formula!)
9. [4, 8.1.3] Compute the Laplace transform of $e^{-a^2t^2}$.
10. [4, 8.1.4] Compute the Laplace transforms of $(t+a)^{-1}$ and $(t+a)^{-2}$.
11. [4, 8.1.5] Compute the Laplace transform of

$$f(t) := \begin{cases} t & 0 \le t \le 1 \\ e^{1-t} & t > 1. \end{cases}$$

12. [4, 8.1.6] Compute the Laplace transform of $t^{-1/2}e^{-\sqrt{at}}$.
13. [4, 8.1.7] Compute the Laplace transform of $t^\alpha L_n^\alpha(t)$ for the Laguerre polynomial L_n^α.
14. [4, 8.1.8] Use partial fraction decompositions to find the exponential polynomial whose Laplace transform is
   (a) $\frac{2(z+1)}{z^2+2z}$
   (b) $\frac{4}{z(z+2)^2}$
   (c) $\frac{1}{z(z^2+1)}$

15. [4, 8.1.9] Let $f(t) = e^{at}$ and $g(t) = e^{bt}$. Compute $f * g$ per definition and by using the Laplace transform.

16. [4, 8.1.10] Compute $f * t$ using the Laplace transform for

    (a) $f(t) = g(t) = J_0(t)$

    (b) $f(t) = t^{a-1}$, $g(t) = t^{b-1}$ for $a, b > 0$

    (c) $f(t) = \sin(t)$, $g(t) = \sin(2t)$.

17. [4, 8.1.11] Assume that f is a-periodic on the positive real axis. Show that

$$\mathfrak{L}f(z) = \frac{F(z)}{1 - e^{-az}}, \quad F(z) = \int_0^a f(t)e^{-zt}dt.$$

    (Hint: one way to do this is to use the fact $\int_0^\infty = \sum_0^\infty \int_{na}^{(n+1)a}$.)

18. [4, 8.1.12] Use the preceding exercise to compute the Laplace transforms of the 2π periodic functions that are equal to:

    (a) $f(t) = t$ for $0 < t < \pi$ and $t - 2\pi$ for $\pi < t < 2\pi$

    (b) $f(t) = 1$ for $0 < t < \pi$, and -1 for $\pi < t < 2\pi$

    (c) $f(t) = t$ for $0 < t < \pi$ and $2\pi - t$ for $\pi < t < 2\pi$.

19. [4, 8.2.2–7] Use the residue theorem to evaluate the inverse Laplace transforms

    (a) $\frac{3z^2+12z+8}{(z+2)^2(z+4)}$

    (b) $\frac{4z^2+z+15}{z(z^2-2z+5}$

    (c) $z^2 \tanh(\pi z/2)$

    (d) $(z \cosh(\sqrt{z}))^{-1}$

    (e) $e^{-a\sqrt{z}}$ and $z^{-1}e^{-a\sqrt{z}}$.

    (f) $z^{-1} \log(1 + z)$.

20. [4, 8.2.10] Let $F(z) = e^{z^2}$ and $f(t) = \frac{1}{2\sqrt{\pi}} e^{-t^2/4}$. Show that for any real b

$$f(t) = \frac{1}{2\pi i} \int_{b-i\infty}^{b+i\infty} F(z)e^{zt}dz,$$

    but F is *not* the Laplace transform of f. Show that F cannot be the Laplace transform of any function that satisfies the conditions in the definition of the Laplace transform. This example shows that one must be careful in applying the LIT to a function F if you do not know in advance that F is the Laplace transform of a Laplace transformable function. If you're going to apply the LIT, make sure your F is legit!

21. [4, 8.3.1–5] Solve the initial value problems using the Laplace transform.

    (a) $u'' + 4u = \sin(\omega t)$, $u(0) = u'(0) = 0$, with $\omega > 0$.

    (b) $u'' + 4u' + 4u = f(t)$, $u(0) = c_0$, $u'(0) = c_1$.

    (c) $u'' + 2u' + 2u = \Theta(t - \pi) - \Theta(t - 2\pi)$, $u(0) = 0$, $u'(0) = 1$.

    (d) $u''' - u' = f(t)$, $u(0) = 1$, $u'(0) = -1$, $u''(0) = 0$.

(e) $u^{(4)} - u = t$, $u^{(j)}(0) = 0$, $j = 0, 1, 2, 3$.

22. [4, 8.3.9] Consider the differential equation $t^2 u'' - 2u = 2t$ for $t > 0$. Use the Laplace transform to find a family of solutions containing one arbitrary constant. What is the general solution of this equation, and why does the procedure you just did not yield it? (Hint: $t^2 u'' - 2u = 0$ is of Euler type.)

23. [4, 8.3.10] Solve the equation $tu'' - (1+t)u' + u = 0$ by Laplace transform.

24. [4, 8.4.5] Consider heat flow in a rod of length ℓ with initial temperature zero when one end is held at temperature zero, and the other end at a variable temperature $f(t)$,
$$u_t = ku_{xx}, \quad u(x,0) = 0, \quad u(0,t) = 0, \quad u(\ell,t) = f(t).$$
Let $v(x,t)$ be the solution to this problem in the special case $f(t) = 1$. Use the Laplace transform to obtain *Duhamel's formula* for u in terms of v,
$$u(x,t) = \frac{\partial}{\partial t} \int_0^t f(s) v(x, t-s) ds.$$

25. [4, 8.4.6] Consider heat flow in an infinite rod, $u_t = ku_{xx}$ for $x \in \mathbb{R}$ with $u(x,0) = f(x)$. Obtain the differential equation $kU_{xx} = zU - f(x)$ for $U = \mathfrak{L}u$, and show that the only solution of this equation that tends to zero as $\operatorname{Re} z \to \infty$ is
$$U(x,z) = \frac{1}{\sqrt{4kz}} \left[\int_{-\infty}^{x} e^{(y-x)\sqrt{z/k}} f(y) dy + \int_{x}^{\infty} e^{(x-y)\sqrt{z/k}} f(y) dy \right].$$
Invert the Laplace transform to obtain the solution
$$u(x,t) = \int_{\mathbb{R}} (4\pi kt)^{-1/2} e^{-(x-y)^2/(4kt)} f(y) dy.$$

26. [4, 8.4.7] Consider flow in a rod of length ℓ: $u_t = ku_{xx}$, $u(x,0) = 0$, $u(0,t) = 0$, $u(\ell,t) = A$. Show that
$$\mathfrak{L}u(x,z) = \frac{A \sinh(x\sqrt{z/k})}{z \sinh(\ell\sqrt{z/k})}.$$

27. [4, 8.5.1] Solve:
$$u(t) - a^2 \int_0^t (t-s) u(s) ds = t^2.$$

28. [4, 8.5.2] Solve:
$$u(t) - \frac{1}{6} \int_0^t (t-s)^3 u(s) ds = f(t).$$

29. [4, 8.5.7] A company buys N_0 new light bulbs at time $t = 0$ and thereafter buys new bulbs at the rate $r(t)$. Suppose the probability that a light bulb will last T units of time after purchase is $p(T)$. Let $N(t)$ be the number of light bulbs in use at time t. Show that
$$N(t) = N_0 p(t) + \int_0^t r(s) p(t-s) ds.$$

30. [4, 8.5.8] In the preceding exercise, suppose that $p(T) = e^{-cT}$. What should the replacement rate $r(t)$ be if the number of light bulbs needed for use at time t is $N(t) = N_0$? Same question but with $N(t) = N_0(2 - e^{-t})$?

Appendix A

Linear algebra

We recall some fundamental facts from linear algebra. Some concepts are well-defined for Hilbert spaces, that are the infinite dimensional analogue of finite dimensional vector spaces. When this is the case we present the concepts in their most general form.

A.1 Linear independence and bases

Definition 141. Let H be a Hilbert space. Then $u, v \in H$ are *linearly independent* if they are both non-zero and the only complex numbers λ, μ such that

$$\lambda u + \mu v = 0$$

are $\lambda = \mu = 0$. A finite set of non-zero vectors in a Hilbert space, $\{v_n\}_{n=1}^N$ are *linearly independent* if the only complex numbers $\{\lambda_n\}_{n=1}^N$ such that

$$\sum_{n=1}^N \lambda_n v_n = 0,$$

are $\lambda_n = 0$ for all $n = 1, \ldots, N$. A collection of non-zero vectors $\{v_n\}_{n \in \mathbb{Z}} \subset H$ is linearly independent if $\{v_n\}_{|n| \le N}$ is linearly independent for all $N \in \mathbb{N}$.

Proposition 142. *Assume that $\{v_k\}_{k=1}^n \subset H$ for some Hilbert space H are non-zero and orthogonal, that is*

$$\langle v_k, v_j \rangle = 0 \text{ for any } k \neq j.$$

Then $\{v_k\}_{k=1}^n$ is linearly independent.

Proof: Assume that

$$\sum_{k=1}^n \lambda_k v_k = 0.$$

We take the scalar product with v_j, for some $j \in \{1, \ldots, n\}$,

$$\langle \sum_{k=1}^n \lambda_k v_k, v_j \rangle = \sum_{k=1}^n \lambda_k \langle v_k, v_j \rangle = \lambda_j ||v_j||^2.$$

On the other hand since we assumed that the sum is zero, this vanishes because

$$\langle \sum_{k=1}^n \lambda_k v_k, v_j \rangle = \langle 0, v_j \rangle = 0.$$

Since $v_j \neq 0$ its length is positive, that is $||v_j||^2 > 0$, and so this forces $\lambda_j = 0$. This argument works for each $j \in \{1, \ldots, n\}$ hence all $\lambda_j = 0$.

Corollary 143. *Assume that $\{v_k\}_{k\in\mathbb{Z}} \subset H$ for some Hilbert space H are non-zero and orthogonal. Then they are linearly independent.*

Proof: Similarly, let $N \in \mathbb{N}$, and assume that for some complex numbers $\{c_n\}_{|n|\leq N}$ we have

$$\sum_{|n|\leq N} c_n v_n = 0.$$

Let $k \in \mathbb{Z}$ with $|k| \leq N$. We take the scalar product of the sum with v_k,

$$0 = \langle \sum_{|n|\leq N} c_n v_n, v_k\rangle = \sum_{|n|\leq N} c_n \langle v_n, v_k\rangle = c_k ||v_k||^2.$$

Since $v_k \neq 0$, and thus $||v_k||^2 \neq 0$ this forces $c_k = 0$. The same argument works for any $N \in \mathbb{N}$ and $k \in \mathbb{Z}$ with $|k| \leq N$. This shows the definition of linearly independent is satisfied.

Definition 144. An orthogonal base for a Hilbert space H is a set of non-zero orthogonal vectors $\{v_n\}$ such that the three conditions in Theorem 28 are satisfied for

$$\left\{\frac{v_n}{||v_n||}\right\}.$$

A basis for a Hilbert space H is a set of non-zero linearly independent vectors $\{v_n\}$ such that every $v \in H$ is equal to

$$v = \sum \langle v, v_n\rangle \frac{v_n}{||v_n||^2}.$$

Proposition 145. *Every basis of $\mathbb{C}^N$ is comprised of exactly N linearly independent non-zero vectors, and any set of N linearly independent non-zero vectors in $\mathbb{C}^N$ is a basis for $\mathbb{C}^N$.*

Proof: Consider first a set of N linearly independent non-zero vectors in $\mathbb{C}^N$, $\{v_n\}$. Let these vectors be the column vectors of a matrix M. Then for any $c \in \mathbb{C}^N$ with $c = (c_1, \ldots, c_n)$ the product

$$Mc = \sum_{n=1}^{N} c_n v_n$$

is a linear combination of these vectors. By the assumption that these vectors are linearly independent, the only solution to $Mc = 0$ is $c = 0$. Consequently the dimension of the kernel of the linear operator defined by matrix multiplication by M is zero. By the Rank-Nullity Theorem, the dimension of the image of this operator is therefore N. This means that for any $v \in \mathbb{C}^N$ there exists some $c \in \mathbb{C}^N$ such that $Mc = v$. Therefore every $v \in \mathbb{C}^N$ can be expressed as a linear combination of the vectors $\{v_n\}$ and they therefore comprise a basis.

The elements of a basis are by definition linearly independent and non-zero. Assume that we have K linearly independent non-zero vectors in $\mathbb{C}^N$ for some $K < N$. Let these vectors be the columns of a matrix A. Then the dimension of the image of $\mathbb{C}^N$ under A is $K < N$. In particular we have the closed subspace,

$$A\mathbb{C}^N := \{Ac : c \in \mathbb{C}^N\} \subset \mathbb{C}^N, \quad A\mathbb{C}^N \cong \mathbb{C}^K \ncong \mathbb{C}^N.$$

These vectors therefore do not form a basis as they cannot span the whole space; there are vectors in $\mathbb{C}^N$ that cannot be expressed as a linear combination of them.

On the other hand if we have K vectors in $\mathbb{C}^N$ for some $K > 0$, then consider the matrix B with these vectors as its column vectors. By the dimension of the image $B\mathbb{C}^N$ is at most N, but the number of columns of B is $K > N$. Consequently, the dimension of the kernel of B is positive, so there exists $c \neq 0$, $c \in \mathbb{C}^N$ such that

$$Bc = 0.$$

Since Bc is a linear combination of the column vectors, and $c \neq 0$, this shows that the vectors are not linearly independent, and therefore they cannot be a basis.

Corollary 146. *Any set of N (non-zero) orthogonal vectors is a basis for $\mathbb{C}^N$.*

Proof: A set of N non-zero orthogonal vectors is linearly independent. The corollary therefore follows immediately from the preceding proposition.

A.2 The spectral theorem for hermitian matrices

An $n \times n$ matrix M defines a linear function that sends vectors in $\mathbb{C}^n$ to vectors in $\mathbb{C}^n$. That is, it is a *linear map* from $\mathbb{C}^n$ to $\mathbb{C}^n$, defined by

$$M(\mathbf{v}) = M\mathbf{v}.$$

So, M sends $\mathbf{v}$ to the matrix-vector product $M\mathbf{v}$. Since M is $n \times n$, and $\mathbf{v} \in \mathbb{C}^n$, the result is also a vector in $\mathbb{C}^n$. This map is linear because if we multiply two vectors by scalars and add them:

$$M(a\mathbf{v} + b\mathbf{w}) = M(a\mathbf{v} + b\mathbf{w}) = aM\mathbf{v} + bM\mathbf{w} = aM(\mathbf{v}) + bM(\mathbf{w}).$$

When the matrix is hermitian, the spectral theorem says that there exists an orthogonal basis of $\mathbb{C}^n$ that consists of eigenvectors of M. Recall that eigenvectors are vectors that satisfy

$$M\mathbf{v} = \lambda\mathbf{v}, \qquad \text{for some } \lambda \in \mathbb{C}.$$

If we have a orthogonal basis of eigenvectors for $\mathbb{C}^n$, this means that by normalizing, we also have an orthonormal basis of eigenvectors $\{\mathbf{v}_k\}_{k=1}^n$. We can therefore express every $\mathbf{x} \in \mathbb{C}^n$ as

$$\mathbf{x} = \sum_{k=1}^n \langle \mathbf{x}, \mathbf{v}_k \rangle \mathbf{v}_k.$$

Moreover, if the eigenvalue for $\mathbf{v}_k$ is λ_k, then it becomes much more simple to compute

$$M(\mathbf{x}) = M\mathbf{x} = \sum_{k=1}^n \langle \mathbf{x}, \mathbf{v}_k \rangle \lambda_k \mathbf{v}_k.$$

Theorem 147 (Spectral Theorem for $\mathbb{C}^n$)**.** *Assume that A is a Hermitian matrix. Then there exists an orthonormal basis of $\mathbb{C}^n$ which consists of eigenvectors of A. Moreover, each of the eigenvalues is real.*

Proof: Remember what Hermitian means. It means that for any $u, v \in \mathbb{C}^n$, we have

$$\langle Au, v \rangle = \langle u, Av \rangle.$$

By the Fundamental Theorem of Algebra, the characteristic polynomial

$$p(x) := \det(A - xI)$$

factors over $\mathbb{C}$. The roots of p are $\{\lambda_k\}_{k=1}^n$. These are by definition the eigenvalues of A. First, we consider the case when A has zero as an eigenvalue. If this is the case, then we define

$$K_0 := \mathrm{Ker}(A) = \{u \in \mathbb{C}^n : Au = 0\}.$$

We note that all nonzero $u \in K_0$ are eigenvectors of A for the eigenvalue 0. Since K_0 is a k-dimensional subspace of $\mathbb{C}^n$, it has an ONB $\{v_1, \ldots, v_k\}$. If $k = n$, we are done. So, assume that $k < n$. Then we consider

$$K_0^\perp = \{u \in \mathbb{C}^n : \langle u, v \rangle = 0 \forall v \in K_0\}.$$

Note that if $u \in K_0^\perp$ then

$$\langle Au, v \rangle = \langle u, Av \rangle = 0 \quad \forall v \in K_0.$$

Hence $A : K_0^\perp \to K_0^\perp$. Moreover, if

$$u \in K_0^\perp, \quad Au = 0 \implies u \in K_0 \cap K_0^\perp \implies u = 0.$$

Hence A is bijective from $K_0^\perp$ to itself. Since A has eigenvalues $\{\lambda_j\}_{j=1}^n$, and 0 appears with multiplicity k, $\lambda_{k+1} \neq 0$. It has some non-zero eigenvector. Let's call it u. Since it is an eigenvector it is not zero, so we define

$$v_{k+1} := \frac{u}{||u||}.$$

Proceeding inductively, we define K_1 to be the span of the vectors $\{v_1, \ldots, v_{k+1}\}$. We look at A restricted to $K_1^\perp$. We note that A maps K_1 to itself because if

$$v = \sum_1^{k+1} c_j v_j \implies Av = \sum_1^{k+1} c_j A v_j = \sum_1^{k+1} c_j \lambda_j v_j \in K_1.$$

Similarly, if $w \in K_1^\perp$,

$$\langle Aw, v \rangle = \langle w, Av \rangle = 0 \forall v \in K_1.$$

So, A maps $K_1^\perp$ into itself. Since the kernel of A is in K_1, A is a surjective and injective map from $K_1^\perp$ into itself. We note that A restricted to $K_1^\perp$ satisfies the same hypotheses as A, in the sense that it is still Hermitian, and it has a characteristic polynomial of degree equal to the dimension of $K_1^\perp$ So, there is an eigenvalue λ_{k+2}, for A as a linear map from $K_1^\perp$ to itself. It has an eigenvector, which we may assume has unit length, contained in $K_1^\perp$. Call it v_{k+2}. Continue inductively until we reach in this way $\{v_1, \ldots, v_n\}$ to span $\mathbb{C}^n$.

Why are the eigenvalues all real? This follows from the fact that if λ is an eigenvalue with eigenvector u then

$$\langle Au, u \rangle = \lambda ||u||^2 = \langle u, Au \rangle = \overline{\lambda} ||u||^2.$$

Since u is an eigenvector it is not zero, so this forces $\lambda = \overline{\lambda}$.

Appendix B

Computer code

Here we include a collection of computer code that can be used to visualize some of the topics from the text.

B.1 Matlab code to graph a Fourier series

To create Figure 4.6, Anton Rosén used the following Matlab code:

```
%Settings to enable saving as vectorized pdf
fig1=figure(1);
fig1.Renderer='Painters';

%Define function that given n and x computes the n+1 first terms in the
%Fourier expansion of abs(x)
syms x k n
absF = @(x,n) pi/2 - 4/pi*symsum(cos((2*k-1)*x)/((2*k-1)^2),k,1,n);

%Plot the Fourier expansion of abs(x) against abs(x)
fplot(absF(x,8),[-8 8])
hold on
f = abs(x);
fplot(f,[-4.5 4.5])
hold off
set(gca,'Xcolor','none')
set(gca,'Ycolor','none')
orient(fig1,'landscape')
print(fig1,'absFourier', '-dpdf', '-bestfit')
```

B.2 Python code

The initial value problem for the homogeneous heat equation on a circular rod can be visualized using python. This physically represents a rod supplied with an initial temperature distribution function $f(x)$ that is then left to disperse throughout the rod with no sinks or sources of heat.

```
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt

# Set plot params
```

```python
plt.rc('font', size=16)          # controls default text sizes
plt.rc('axes', titlesize=16)     # fontsize of the axes title
plt.rc('axes', labelsize=18)     # fontsize of the x and y labels
plt.rc('xtick', labelsize=16)    # fontsize of the tick labels
plt.rc('ytick', labelsize=16)    # fontsize of the tick labels
plt.rc('legend', fontsize=16)    # legend fontsize

def alpha0(x,f):
    return np.trapz(y=f, x=x) / (2*np.pi)

def alpha(n,x,f):
    return np.trapz(y=f*np.cos(n*x), x=x) / np.trapz(y=np.abs(np.cos(n*x))**2, x=x)

def beta(n,x,f):
    return np.trapz(y=f*np.sin(n*x), x=x) / np.trapz(y=np.abs(np.sin(n*x))**2, x=x)

def un(x,t,n,k,f):
    if n == 0:
        alpha_n = alpha0(x,f)
        beta_n = 0
    else:
        alpha_n = alpha(n,x,f)
        beta_n = beta(n,x,f)
    return np.exp(-n**2 * t*k) * (alpha_n * np.cos(n*x) + beta_n * np.sin(n*x))

def u(x,t,n_max,k,f):
    u_xt = np.zeros(len(x))
    for n in range(n_max):
        u_xt += un(x,t,n,k,f)
    return u_xt


def circular_rod_heat_solution(n_max, x, f, flabel, filename):
    '''
    Plots the solution to the heat equation on a circular one dimensional rod, for a given
    and an initial function f defined on those x-values with a corresponding label flabel.
    Args:
      n_max (number): Integer for the maximum n (i.e. the number of basis functions) that
      x (list): x-domain, vector of values in the range [-pi, pi].
      f (list): an initial value function defined on the x-domain.
      flabel (string): Label of the initial value function for naming the output file.
    '''
    # Define constants - here one can control the number of ns, i.e. the accuracy of the s
    k = 1
    ts = [0,1,100]

    # Calculate heat equation solution for the timesteps
    sols = np.zeros((len(ts), len(x)))

    for i,t in enumerate(ts):
        sol = u(x,t,n_max,k,f)
        sols[i,:] = sol
```

```python
    # Plot solutions
    fig, ax = plt.subplots(figsize=(8,6))

    cs = ['r', 'y', 'b']
    styles = ['-', '-.', ':']
    ax.plot(x, f, c='k', linestyle='--', linewidth=2, alpha=1, label=r'$f(x)=$'+flabel)
    for i,sol in enumerate(sols):
        ax.plot(x, sol, c=cs[i], alpha=0.6, linestyle=styles[i], linewidth=2, label=r'$u(x
    ax.legend(loc='best')
    ax.set_xlabel(r'$x$')
    ax.set_ylabel(r'$u(x,t)$, arb. units')
    tick_places = [-np.pi, -np.pi/2, 0, np.pi/2, np.pi]
    my_xticks = [r'$-\pi$',r'$-\frac{\pi}{2}$','0',r'$\frac{\pi}{2}$',r'$\pi$']
    plt.xticks(tick_places, my_xticks)

    # Disable y ticks
    # frame1 = plt.gca()
    # frame1.axes.get_yaxis().set_ticks([])

    # Save figure
    plt.tight_layout()
    plt.savefig(f'./images/pdf/f_{filename}_n={n_max}_k={k}.pdf') # Save as pdf
    plt.savefig(f'./images/png/f_{filename}_n={n_max}_k={k}.png') # and png
    plt.close()


def main():
    n_max = 10 # number of ns - i.e. NOT including n = n_max
    x = np.linspace(-np.pi, np.pi, 500)
    # Initial function values - add/modify these and the labels/filenames below appropriat
    fs = [
        2*np.ones(len(x)),          # Constant
        x,                          # Linear
        np.sin(x),                  # Sine
        np.heaviside(x, 0.5),       # Heaviside - x2 is the value at f(x=0)
        x**2,                       # Quadratic
        np.cosh(x),                 # Cosh
        np.abs(x),                  # Absolute value
    ]
    flabels = [
        'constant',
        r'$x$',
        r'$\sin(x)$',
        r'$H(x)$',
        r'$x^2$',
        r'$\cosh(x)$',
        r'$|x|$'
    ]
    filenames = [
        'constant',
        'linear',
```

```
        'sine',
        'heaviside',
        'quadratic',
        'cosh',
        'abs'
    ]

    for i, f in enumerate(fs):
        circular_rod_heat_solution(n_max, x, f, flabels[i], filenames[i])


if __name__ == "__main__":
    main()

#  |\__/,|   (`\
#  |_ _  |.--.) )
#  ( T   )     /
# (((^_(((/(((_/
#
# Thank you for checking out this script!
# Author: Eric Lindgren, F16, 2021
```

Appendix C

Distributions and weak solutions to PDEs

The mathematical concept of a *distribution,* or, as they are sometimes called, *generalized function,* has been badly abused not only by physicists but also by mathematicians. You may have already heard about the so-called "delta function." It's not really a function. It's not a 'generalized function.' It has its very own terminology, and that is that it is a distribution. Now, distributions are not as mysterious and weird as the mystique in which they are often shrouded.

Distributions are functions which themselves take as input a function. A particularly nice class of distributions are the *tempered distributions.* These distributions take in a Schwarz class function and spit out a number.

Definition 148. Assume that f is a smooth function on $\mathbb{R}$. Then, we say that $f \in \mathcal{S}$ if for all k and for all n,

$$\lim_{|x| \to \infty} x^n f^{(k)}(x) = 0.$$

In other words, f and all its derivatives decay rapidly at $\pm\infty$. There are quite a few functions which satisfy this. For example, all smooth functions which live on a bounded interval (compactly supported) satisfy this property.

Exercise 149. *Show that if $f \in \mathcal{S}$ then all of the derivatives of f are in $\mathcal{S}$. Show that if $f \in \mathcal{S}$ then its Fourier transform is also in $\mathcal{S}$.*

Definition 150. A *tempered distribution* is a function which maps $\mathcal{S}$ to $\mathbb{C}$, which satisfies the following conditions:

- It is linear, so for a distribution denoted by L, we have

  $$L(\alpha f + \beta g) = \alpha L(f) + \beta L(g),$$

  for all f and g in $\mathcal{C}_c^\infty(\mathbb{R})$ and for all complex numbers α and β.

- There is a non-negative integer N and a constant $C \geq 0$ such that for all $f \in \mathcal{S}$

  $$|L(f)| \leq C \sum_{j+k \leq N} \sup_{x \in \mathbb{R}} |x^j f^{(k)}(x)|.$$

Let's do an example. We define a distribution in the following way. For $f \in \mathcal{C}_c^\infty(\mathbb{R})$,

$$L(f) := f(0).$$

That is, the distribution takes in the function, f, and spits out the value of f at the point $0 \in \mathbb{R}$. This distribution satisfies for any f and g in $\mathcal{C}_c^\infty(\mathbb{R})$ and for any α and $\beta \in \mathbb{C}$,

$$L(\alpha f + \beta g) = \alpha L(f) + \beta L(g).$$

Moreover, we have the estimate that

$$|L(f)| \leq |f(0)| \leq \sup_{x \in \mathbb{R}} |f(x)|.$$

So the estimate required is satisfied with $N = 0$ and $C = 1$. This distribution has a name. It is called the *delta* distribution. It is usually written with the letter δ. It is nothing other than a function which takes a function as its input and spits out a number as its output.

Exercise 151. *Assume that $f \in \mathcal{C}_c^\infty(\mathbb{R})$. Show that by defining*

$$L_f(g) = \int_{\mathbb{R}} f(x)g(x)dx, \quad g \in \mathcal{C}_c^\infty(\mathbb{R}),$$

L_f is a tempered distribution.

In fact, the assumption that $f \in \mathcal{C}_c^\infty(\mathbb{R})$ wasn't even necessary. You can show that for $f \in \mathcal{L}^2(\mathbb{R})$ or $f \in \mathcal{L}^1(\mathbb{R})$, the distribution, L_f defined above (it takes in a function $g \in \mathcal{C}_c^\infty(\mathbb{R})$ and integrates the product with f over $\mathbb{R}$), is a distribution. So, here's something which is rather cool. The elements in $\mathcal{L}^2(\mathbb{R})$ and $\mathcal{L}^1(\mathbb{R})$ are in general *not* differentiable at all. However, the *distributions* we can make out of them *are* differentiable. Here's how we do that.

Definition 152. The derivative of a tempered distribution, L is another tempered distribution, denoted by $L' \in \mathcal{D}(\mathbb{R})$, which is defined by

$$L'(g) = -L(g'), \quad g \in \mathcal{S}.$$

To see that this definition makes sense, we think about the special case where $L = L_f$, and $f \in \mathcal{S}$. Then, we *can* take the derivative of f, and it is also an element of $\mathcal{S}$. So, we can define $L_{f'}$ in the analogous way. Let's write it down when it takes in $g \in \mathcal{S}$,

$$L_{f'}(g) = \int_{\mathbb{R}} f'(x)g(x)dx.$$

We can do integration by parts. The boundary terms vanish, so we get

$$L_{f'}(g) = \int_{\mathbb{R}} f'(x)g(x)dx = -\int_{\mathbb{R}} f(x)g'(x)dx.$$

So,

$$L_{f'}(g) = -L_f(g') = (L_f)'(g).$$

This is why it makes a lot of sense to define the derivative of a distribution in this way. For the heavyside function, we define

$$L_H, \quad L_H(g) = \int_0^\infty g(x)dx.$$

Then, we compute that

$$L_H'(g) = -L_H(g') = -\int_0^\infty g'(x)dx.$$

Due to the fact that $g \in \mathcal{S}$,

$$\lim_{x\to\infty} g(x) = 0.$$

Hence, we have

$$-\int_0^\infty g'(x)dx = -(0 - g(0)) = g(0) = \delta(g).$$

So, we see that the derivative of L_H is the δ distribution! Pretty neat!

In this way, distributions can solve differential equations! For example, we'd say that a distribution L satisfies the equation

$$L'' + \lambda L = 0$$

if, for every $g \in \mathcal{S}$ we have

$$L''(g) + \lambda L(g) = 0.$$

This turns out to be incredibly useful and important in the theory of partial differential equations. However, the way it usually works is that instead of actually finding a distribution which solves the PDE, one shows by abstract mathematics that there *exists* a distribution which solves the PDE. Then, one can use clever methods to show that the mere existence of a distribution solving the PDE, which is called a *weak solution,* actually implies that there exists a genuinely differentiable solution to the PDE. We don't want to get ahead of ourselves here, so conclude with one last exercise, which proves that you can differentiate distributions as many times as you like!

Exercise 153. *Use induction to show that you can differentiate a distribution as many times as you like, by defining*

$$L^{(k)}(g) := (-1)^k L(g^{(k)}).$$

In a similar way, we can define the Fourier transform of a distribution.

Definition 154. Assume that L is a tempered distribution. The Fourier transform of L is the distribution, $\hat{L}$ which for $f \in \mathcal{S}$ acts as follows

$$\hat{L}(f) := L(\hat{f}).$$

In this way, we can compute the Fourier transform of our favorite distribution, δ.

$$\hat{\delta}(f) := \delta(\hat{f}) = \hat{f}(0) = \int_{\mathbb{R}} f(x)dx.$$

So, we could think of the Fourier transform of δ as the distribution which acts by

$$\hat{\delta} : f \in \mathcal{S} \mapsto \int_{\mathbb{R}} f(x)dx.$$

On the other hand, by the FIT,

$$\delta(f) = f(0) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\xi)d\xi = \frac{1}{2\pi}\hat{\delta}(\hat{f}) = \frac{1}{2\pi}\hat{\hat{\delta}}(f).$$

So that's kind of cute. It says that

$$\delta = \frac{1}{2\pi}\hat{\hat{\delta}}.$$

Appendix D

The Lebesgue and Hausdorff measures

Measures are as the name suggests: a method for measuring the size of sets. They are defined on collections of sets known as sigma algebras.

D.1 Algebras, sigma algebras, and measures

To get started, we define the set of all sets:

$$P(X) = \text{the set of all subsets of } X.$$

Definition 155. Let X be a set. A subset $\mathcal{A} \subset P(X)$ is called an *algebra* if

1. $X \in \mathcal{A}$
2. $Y \in \mathcal{A} \implies X \setminus Y =: Y^c \in \mathcal{A}$
3. $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$

$\mathcal{A}$ is a σ-algebra if in addition

$$\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}.$$

Remark 5. First, note that since $X \subset \mathcal{A}$, and algebras are closed under complementation, (yes it is a real word), one always has

$$\emptyset = X^c \in \mathcal{A}.$$

Moreover, we note that algebras are always closed under intersections, since for $A, B \in \mathcal{A}$,

$$A \cap B = (A^c \cup B^c)^c \in \mathcal{A},$$

since algebras are closed under complements and unions. Consequently, σ-algebras are closed under countable intersections.

We will often use the symbol σ in describing countably-infinite properties.

Exercise 156. *What is the smallest possible algebra? What is the next-smallest algebra? Continue building up algebras. Now, let X be a topological space. The Borel σ-algebra is defined to be the smallest σ-algebra which contains all open sets. What other kinds of sets are contained in the Borel σ-algebra?*

With the notion of σ-algebra, we can define a *measure.*

Definition 157. Let X be a set and $\mathcal{A} \subset P(X)$ a σ-algebra. A *measure* μ is a countably additive, set function which is defined on the σ-algebra, $\mathcal{A}$, such that $\mu(\emptyset) = 0$. The elements of $\mathcal{A}$ are known as *measurable sets.* We will only work with non-negative measures, but there is such a thing as a signed measure. Just so you know those beasties are out there. Countably additive means that for a countable disjoint collection of sets in the σ-algebra

$$\{A_n\} \subset \mathcal{A} \text{ such that } A_n \cap A_m = \emptyset \forall n \neq m \implies \mu\left(\bigcup A_n\right) = \sum \mu(A_n).$$

We shall refer to $(X, \mathcal{A}, \mu)$ as a *measure space.* What this means is that a measure space is comprised of a big set, X, and a certain collection of subsets of X, which is the σ-algebra, $\mathcal{A}$. Moreover, there is a *measure,* μ, which is a countably additive set function that is defined on all elements of $\mathcal{A}$.

Proposition 158 (Measures are monotone)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then μ is finitely additive, that is if $A \cap B = \emptyset$ for two elements A, $B \in \mathcal{A}$, we have*

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Moreover, μ is monotone, that is for any $A \subset B$ which are both elements of $\mathcal{A}$ we have

$$\mu(A) \leq \mu(B).$$

Proof: First we make the rather trivial observation that if A and B are two elements of $\mathcal{A}$ with empty intersection, then

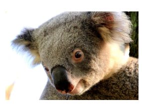
$$A \cup B = \cup A_j, \quad A_1 = A, \quad A_2 = B, \quad A_j = \emptyset \forall j \geq 3.$$

Then we have

$$\mu(A \cup B) = \mu(\cup A_j) = \sum_j \mu(A_j) = \mu(A) + \mu(B),$$

since $\mu(\emptyset) = 0$. For the monotonicity, if $A \subset B$ are two elements of $\mathcal{A}$, then

$$\mu(B) = \mu(B \setminus A \cup A) = \mu(B \setminus A) + \mu(A) \geq \mu(A),$$

since $\mu \geq 0$. 

So, in layman's terms, when we've got a measure space, we have a big set, X, together with a collection of subsets of X (note that X is a subset of itself, albeit not a *proper* subset), for which we have a notion of *size.* This size is the value of the function μ. So, if $Y \in \mathcal{A}$, then $\mu(Y)$ is the *measure* of Y. Roughly speaking, $\mu(Y)$ tells us how much space within X the set Y is occupying. For the case of the Lebesgue measure on $\mathbb{R}^n$, and the n-dimensional Hausdorff measure, we shall see that measure coincides with our usual notion of n-dimensional volume.

Proposition 159 (How to disjointify sets and countable sub-additivity)**.** *If $\{A_n\} \subset \mathcal{A}$ is a countable collection of sets, then we can find a disjoint collection $\{B_n\} \subset \mathcal{A}$ such that*

$$\cup A_n = \cup B_n.$$

Let μ be a measure defined on the σ-algebra, $\mathcal{A}$. Then countable sub-additivity holds for not-necessarily-disjoint countable collections of sets, which means that for all such $\{A_n\}$ as above,

$$\mu(\cup A_n) \leq \sum \mu(A_n).$$

Proof: We do this by setting

$$B_1 := A_1, \quad B_n := A_n \setminus \cup_{k=1}^{n-1} B_k, \qquad n \geq 2.$$

Then for $m > n$, note that

$$B_m \cap B_n = \left(A_m \setminus \cup_{k=1}^{m-1} B_k\right) \cap B_n = \emptyset$$

since

$$B_n \subset \cup_{k=1}^{m-1} B_k,$$

since $n \le m-1$. Thus they are in fact disjoint. Moreover,

$$B_1 = A_1, \quad B_2 \cup B_1 = A_2 \setminus A_1 \cup A_1 = A_2 \cup A_1.$$

Similarly, by induction, assuming that

$$\cup_{k=1}^n B_k = \cup_{k=1}^n A_k,$$

we have

$$\cup_{k=1}^{n+1} B_k = B_{n+1} \cup \cup_{k=1}^n B_k = A_{n+1} \cup \cup_{k=1}^n B_k = A_{n+1} \cup \cup_{k=1}^n A_n$$

where in the last equality we used the induction hypothesis. Thus,

$$\cup_{n \ge 1} B_n = \cup_{n \ge 1} A_n.$$

Moreover, the way we have defined B_n together with the definition of the σ-algebra, $\mathcal{A}$, shows that $B_n \in \mathcal{A}$ for all n. By the monotonicity of μ,

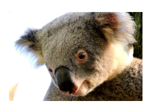
$$B_n \subset A_n \forall n \implies \mu(B_n) \le \mu(A_n).$$

By the countable additivity for the disjoint sets, $\{B_n\}$, and since $\cup B_n = \cup A_n$

$$\mu(\cup A_n) = \mu(\cup B_n) = \sum \mu(B_n) \le \sum \mu(A_n).$$

So, for not-necessarily disjoint sets, we have countable subadditivity, which means that

$$\mu(\cup A_n) \le \sum \mu(A_n),$$

for all countable collections of sets $\{A_n\} \subset \mathcal{A}$. 

Definition 160. A measure space $(X, \mathcal{A}, \mu)$ is σ-finite if there exists a collection of sets $\{A_n\} \subset \mathcal{A}$ such that

$$X = \cup A_n, \quad \text{and} \quad \mu(A_n) < \infty \quad \forall n.$$

Exercise 161. *What are some examples of σ-finite measure spaces? What are some examples of measure spaces which are* not *σ-finite?*

One unfortunate fact about measures is that they're not defined on arbitrary sets, only on measurable sets (remember, those are the ones in the associated σ algebra). However, there is a way to define a set function which is almost like a measure and is defined for *every imaginable or unimaginable set.* This thing is called an *outer measure.*

Definition 162. Let X be a set. An outer measure μ^* on X is a map from $P(X) \to [0, \infty]$ such that

$$\mu^*(\emptyset) = 0, \quad A \subset B \implies \mu^*(A) \le \mu^*(B),$$

and

$$\mu^*(\cup A_n) \le \sum \mu^*(A_n).$$

Whenever things are indexed with n or some other letter and are not obviously indicated to be uncountable or finite, we implicitly are referring to a set indexed by the natural numbers.

D.1.1 Carathéodory's outer measures

We will require techniques from a Greek mathematician, Constantin Carathéodory.

Proposition 163 (Outer Measures). *Let $E \subset P(X)$ such that $\emptyset \in E$. Let ρ be a map from elements of E to $[0,\infty]$ such that $\rho(\emptyset) = 0$. Then we can define for every element $A \in P(X)$*

$$\rho^*(A) := \inf \left\{ \sum \rho(E_j) : E_j \in E, A \subset \cup E_j \right\},$$

where we assume that $\inf\{\emptyset\} = \infty$, so that if it is impossible to cover a set A by elements of E then $\rho^*(A) := \infty$. So defined, ρ^* is an outer measure.

Proof: Note that ρ^* is defined for every set. Now since $\emptyset \subset \emptyset = \cup E_j$, taking all $E_j = \emptyset \in E$ we have the cover for $\emptyset$ given by this particular choice of $\{E_j\} \subset E$. Therefore, since $\rho \geq 0$, we have that $\rho^* \geq 0$, and on the other hand since it is an infimum,

$$0 \leq \rho^*(\emptyset) \leq \sum_j \rho(\emptyset) = 0 \implies \rho^*(\emptyset) = 0.$$

This is the first condition an outer measure must satisfy.

Next, let's assume $A \subset B$. (By $\subset$ we always mean $\subseteq$). Then, since any covering of B by elements of E is also a covering of A by elements of E, it follows that the infimum over coverings of A is an infimum over a potentially *larger set of objects (namely coverings)* as compared with the infimum over coverings of B. Hence we have

$$\rho^*(A) = \inf\{\sum \rho(E_j) : E_j \in E, A \in \cup E_j\} \leq \inf\{\sum \rho(E_j) : E_j \in E, B \in \cup E_j\} = \rho^*(B).$$

This is the second condition.

Finally, we must show that ρ^* is countably subadditive. So, let $\{A_n\}$ be a collection of sets in $P(X)$. If for any n we have no cover of A_n by elements of E, then since

$$A_n \subset \cup_k A_k,$$

there is no cover of $\cup_k A_k$ by elements of E either. Hence we have

$$\rho^*(\cup A_n) = \infty, \quad \rho^*(A_n) = \infty \leq \sum \rho^*(A_k) \implies \rho^*(\cup A_n) = \infty = \sum \rho^*(A_n).$$

Thus countable subadditivity is verified in this case.

So, to complete the proof, we assume that each A_n admits at least one covering by elements of E. Let $\varepsilon > 0$. Since the definition of ρ^* is by means of an infimum, for each $j \in \mathbb{N}$ there exists a countable collection of sets $\{E_j^k\}_{k=1}^\infty$ where each $E_j^k \in E$, such that

$$\rho^*(A_j) \geq \sum_{k \geq 1} \rho(E_j^k) - \frac{\epsilon}{2^j} \implies \rho^*(A_j) + \frac{\epsilon}{2^j} \geq \sum_{k \geq 1} \rho(E_j^k).$$

Well then, the collection $\{E_j^k\}$ is a countable collections of elements of E which covers

$$\cup A_j.$$

Therefore by the definition of ρ^* as the infimum over such covers, we have

$$\rho^*(\cup A_j) \leq \sum_{j,k \geq 1} \rho(E_j^k).$$

Since for each E_j^k we have

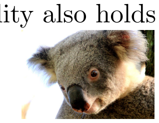
$$\rho^*(A_j) + \frac{\epsilon}{2^j} \geq \sum_{k \geq 1} \rho(E_j^k),$$

summing over k we have

$$\sum_{j,k\geq 1} \rho(E_j^k) \leq \sum_{j\geq 1} \rho^*(A_j) + \frac{\epsilon}{2^j} = \epsilon + \sum_{j\geq 1} \rho^*(A_j).$$

Thus following all the inequalities we have

$$\rho^*(\cup A_j) \leq \epsilon + \sum_j \rho^*(A_j).$$

Since this inequality holds for arbitrary $\epsilon > 0$, we may let $\epsilon \to 0$, and the inequality also holds without that pesky ϵ. Hence we have verified countable subadditivity in this last case as well. 

For each measure space, there is a canonically associated outer measure.

Corollary 164. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then, there is a canonically associated outer measure induced by μ defined by*

$$\mu^*(A) := \inf\{\sum \mu(E_j), \quad \{E_j\} \subset \mathcal{A},\ A \subset \cup E_j\}.$$

Proof: By the definition of measure space, we have that $\emptyset \in \mathcal{A}$, and $\mu(\emptyset) = 0$. Moreover, $\mu : \mathcal{A} \to [0, \infty]$. Finally, we note that since for any $A \in P(X)$, $A \subset X \in \mathcal{A}$, we can always find a covering of such A by elements of $\mathcal{A}$. (In particular, one covering is to take $E_j = X$ for all j). Thus, μ^* is defined for all $A \in P(X)$. Moreover, μ and $\mathcal{A}$ satisfy the hypotheses of the preceding proposition. Therefore, since μ^* is defined in an analogous way to ρ^*, by the preceding proposition we also have that μ^* is an outer measure. 

Remark 6. For a measure space $(X, \mathcal{A}, \mu)$, we shall use μ^* to denote the *canonically associated outer measure,* which is defined according to the corollary. One of the reasons we require the notion of an outer measure is because it is used to define what it means for a measure space to be *complete.*

D.1.2 Completeness

If our notion of size (volume) defined in terms of the measure of sets belonging to a sigma algebra is a good notion, then if a certain set has size zero, anything contained within that set ought to also have size zero. It is precisely this observation that motivates the definition of a *complete* measure, which can be formulated in two different but equivalent ways.

Proposition 165 (Completeness Proposition)**.** *The following are equivalent for a measure space $(X, \mathcal{M}, \mu)$. If either of these hold, then μ is called* complete.

1. *If there exists $N \in \mathcal{M}$ with $\mu(N) = 0$, and $Y \subset N$, then $Y \in \mathcal{M}$.*

2. *If $\mu^*(Y) = 0$ then $Y \in \mathcal{M}$.*

Proof:

First let us assume (1) holds. Then if $Y \subset X$ with $\mu^*(Y) = 0$, by the definition of μ^* for each $k \in \mathbb{N}$ there exists

$$\{E_n^k\}_{n\geq 1} \subset \mathcal{M}, \quad Y \subset \cup_n E_n^k, \quad \sum_n \mu(E_n^k) < 2^{-k}.$$

Well, then

$$Y \subset N := \cap_k \cup_n E_n^k \in \mathcal{M},$$

where the containment holds because $\mathcal{M}$ is a σ-algebra. Since $N \subset \cup_n E_n^k$ for each $k \in \mathbb{N}$, by monotonicity of the measure

$$\mu(N) \leq \mu(\cup_n E_n^k) < 2^{-k} \forall k \in \mathbb{N} \implies \mu(N) = 0.$$

By the assumption of (1) since $Y \subset N \in \mathcal{M}$ and $\mu(N) = 0$, it follows that $Y \in \mathcal{M}$. So, every set with outer measure zero is measurable (that's what (2) says!)

Next, we assume (2) holds. Then if there exists $N \in \mathcal{M}$ with $\mu(N) = 0$ and $Y \subset N$, then

$$Y \subset \cup A_j, \quad A_1 := N, \quad A_j = \emptyset \forall j \geq 2,$$

and $\{A_j\} \subset \mathcal{M}$. So, by definition of outer measure,

$$0 \leq \mu^*(Y) = \inf \ldots \leq \sum \mu(A_j) = \mu(N) = 0.$$

Consequently $\mu^*(Y) = 0$, and by the assumption (2), $Y \in \mathcal{M}$. This shows that (2) $\implies$ (1). Hence, they are equivalent. 

D.1.3 Exercises: Measure theory basics

1. Let X be a finite set. How many elements does $P(X)$ contain? Prove your answer!
2. Given a measure space $(X, \mathcal{A}, \mu)$ and $E \in \mathcal{A}$, define
$$\mu_E(A) = \mu(A \cap E)$$
for $A \in \mathcal{A}$. Prove that μ_E is a measure.
3. Prove that the intersection of arbitrarily many σ-algebras is again a σ-algebra. Does the same hold for unions?
4. Let A be an infinite σ-algebra. Prove that A contains uncountably many elements.
5. Let $X = \mathbb{N}$, and define the algebra $\mathcal{A} = P(X)$. Prove that all elements of $\mathcal{A}$ are either countably infinite, finite, or empty. Define the measure to be 1 on a single element of $\mathbb{N}$ and 0 on the empty set. Prove that this satisfies the definition of a measure space. Will it also work to take $X = \mathbb{R}$, and let $\mathcal{A} = P(\mathbb{R})$, using the same definition of the measure? Do we get a measure space? Why or why not?

D.2 Completion of a measure, creating a measure from an outer measure, and pre-measures

There is a natural way to complete a measure; this is the content of the following theorem.

Theorem 166 (Completion of a measure)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $\mathcal{N} := \{N \in \mathcal{M} \mid \mu(N) = 0\}$ and*

$$\bar{\mathcal{M}} = \{E \cup F \mid E \in \mathcal{M} \text{ and } F \subset N \text{ for some } N \in \mathcal{N}\}.$$

Then $\bar{\mathcal{M}}$ is a σ-algebra and $\exists!$ extension $\bar{\mu}$ of μ to a complete measure on $\bar{\mathcal{M}}$. Moreover, if $\mathcal{A}$ is a σ-algebra which contains $\mathcal{M}$, such that $(X, \mathcal{A}, \nu)$ is a complete measure space, and ν restricted to $\mathcal{M}$ is equal to μ, then $\mathcal{A} \supset \bar{\mathcal{M}}$. In this sense, $(X, \bar{\mathcal{M}}, \bar{\mu})$ is the minimal complete extension of $(X, \mathcal{M}, \mu)$ to a complete measure space.

Proof: First we show that $\overline{\mathcal{M}}$ is a σ-algebra. We observe that every element of $\mathcal{M}$ can be written as itself union with $\emptyset$, and $\emptyset \subset \emptyset \in \mathcal{N}$. So it follows that every element of $\mathcal{M}$ is an element of $\bar{\mathcal{M}}$. Next, assume that $\{A_n\} \subset \bar{\mathcal{M}}$ and $\{E_n, N_n\} \subset \mathcal{M}$ such that

$$A_n = E_n \cup F_n, \quad F_n \subset N_n \in \mathcal{N}.$$

Then

$$N := \cup N_n \in \mathcal{M}, \quad \text{and} \quad \mu(\cup N_n) \leq \sum \mu(N_n) = 0.$$

Since $\emptyset \subset N$, we have by the monotonicity of μ that

$$0 = \mu(\emptyset) \leq \mu(N) \leq \sum \mu(N_n) = 0.$$

We also have that

$$E := \cup E_n \in \mathcal{M}.$$

Then, let us define $F := \cup F_n \subset N$. It follows that

$$\cup A_n = E \cup F \in \bar{\mathcal{M}}.$$

Consequently $\bar{\mathcal{M}}$ is closed under countable unions. What about complements? If $A = E \cup F \in \bar{\mathcal{M}}$ with $F \subset N \in \mathcal{N}$ then note that

$$(E \cup F)^c = E^c \cap F^c = ((E^c \cap N) \cup (E^c \cap N^c)) \cap F^c,$$

and since $F \subset N \implies F^c \supset N^c$, the intersection of the last two terms is just $E^c \cap N^c$, so

$$(E \cup F)^c = (E^c \cap N \cap F^c) \cup (E^c \cap N^c).$$

Since $E, N \in \mathcal{M} \implies E^c \cap N^c \in \mathcal{M}$, and $E^c \cap N \cap F^c \subset N \in \mathcal{N}$ we see that $(E \cup F)^c \in \bar{\mathcal{M}}$. So, $\bar{\mathcal{M}}$ is closed under complements. Hence, we have shown that $\bar{\mathcal{M}}$ is a σ-algebra which contains $\mathcal{M}$.

Next, we must demonstrate that $\bar{\mu}$ is a well-defined, complete, and unique extension of μ. It is natural to ignore the subset of the zero-measure set, so we define

$$\bar{\mu}(E \cup F) := \mu(E).$$

If we have another representation of $E \cup F = G \cup H$ with $G \in \mathcal{M}$ and $F, H \subset N, M \in \mathcal{N}$, respectively, then

$$\bar{\mu}(E \cup F) = \mu(E).$$

Since $E \subset E \cup F = G \cup H \subset G \cup M$, with $G \cup M \in \mathcal{M}$, we have by the monotonicity of μ,

$$\mu(E) \leq \mu(G \cup M) \leq \mu(G) + \mu(M) = \mu(G).$$

Above, we have used countable subadditivity and the fact that $M \in \mathcal{N}$. Then, we note that

$$\bar{\mu}(G \cup H) = \mu(G),$$

as we have defined $\bar{\mu}$. So, following the equalities and inequalities, we have

$$\bar{\mu}(E \cup F) = \mu(E) \leq \mu(G) = \bar{\mu}(G \cup H).$$

To complete the argument, we use the Shakespeare technique: what is in a name? Would not a rose by any other name smell as sweet? Simply repeat the same argument above, replacing E by G and F by H, that is we do the same mathematical argument but we simply swap the names. Then we obtain

$$\bar{\mu}(G \cup H) \leq \bar{\mu}(E \cup F).$$

Hence we have shown that

$$\bar{\mu}(E \cup F) = \bar{\mu}(G \cup H).$$

We conclude that $\bar{\mu}$ is well-defined.

Now, let's show that $\bar{\mu}$ is a measure which extends μ. By definition, for $E \subset \mathcal{M}$

$$\bar{\mu}(E) = \bar{\mu}(E \cup \emptyset) = \mu(E).$$

So, this shows that

$$\bar{\mu}|_{\mathcal{M}} = \mu.$$

We also observe that since

$$\emptyset \in \mathcal{M} \implies \bar{\mu}(\emptyset) = \mu(\emptyset) = 0.$$

Next we wish to show monotonicity. If

$$A = E \cup F, \quad E \in \mathcal{M}, \quad F \subset N \in \mathcal{N},$$

and

$$A \subset B = G \cup H, \quad G \in \mathcal{M}, \quad H \subset M \in \mathcal{N},$$

then we have

$$E \subset A \subset B = G \cup H \subset G \cup M \implies$$

$$\bar{\mu}(A) = \mu(E) \leq \mu(G \cup M) \leq \mu(G) + \mu(M) = \mu(G) = \bar{\mu}(B).$$

We therefore have shown that $\bar{\mu}$ is monotone.

Next we wish to show that $\bar{\mu}$ is countably additive. Assume that $\{A_n\} = \{E_n \cup F_n\} \subset \bar{\mathcal{M}}$ are disjoint. Then

$$A_n \cap A_m = E_n \cup F_n \cap (E_m \cup F_m) \supset E_n \cap E_m,$$

which shows that

$$E_n \cap E_m = \emptyset, \quad \forall n \neq m.$$

Consequently,

$$\bar{\mu}(\cup A_n) = \mu(\cup E_n) = \sum \mu(E_n) = \sum \bar{\mu}(A_n).$$

So, $\bar{\mu}$ is countably additive. We have therefore proven that $\bar{\mu}$ is a measure on $\bar{\mathcal{M}}$.

Let's show that $\bar{\mu}$ is complete. Assume that $Y \in \bar{\mathcal{M}}$ with $\bar{\mu}(Y) = 0$. Then we can write

$$Y = E \cup F, \quad E \in \mathcal{N}, \quad F \subset N \in \mathcal{N}.$$

Hence, in particular,

$$Y \subset E \cup N \in \mathcal{N}.$$

Therefore $Z \subset Y \subset N$. We can therefore write Z as

$$Z = \emptyset \cup Z, \quad \emptyset \in \mathcal{M}, \quad Z \subset N \in \mathcal{N}.$$

It follows from the definition of $\mathcal{M}$ that $Z \in \bar{\mathcal{M}}$. Thus, any subset of a $\bar{\mathcal{M}}$ measurable set which has $\bar{\mu}$ measure zero is also an element of $\bar{\mathcal{M}}$, which is the first of the equivalent conditions required to be a complete measure.

Finally the uniqueness. Let's assume ν also extends μ to a complete measure on $\mathcal{M}$. This means that

$$\nu|_{\mathcal{M}} = \bar{\mu}|_{\mathcal{M}} = \mu.$$

For $Y = E \cup F \in \bar{\mathcal{M}}$, we also have $Y \subset E \cup N$, so by countable subadditivity,

$$\nu(Y) \leq \nu(E) + \nu(N) = \mu(E) + \mu(N) = \mu(E) = \bar{\mu}(Y).$$

Conversely

$$\bar{\mu}(Y) = \mu(E) = \nu(E) \leq \nu(E \cup F) = \nu(Y).$$

So, we've got equality all across, and in particular, $\nu(Y) = \bar{\mu}(Y)$.

Finally, let us assume that there is some other extension, φ, of μ to a complete measure on some σ-algebra $\mathcal{A}$ which contains $\mathcal{M}$. Thus, $(X, \mathcal{A}, \varphi)$ is a complete measure space, and

$$\varphi|_{\mathcal{M}} = \mu.$$

Then

$$\varphi(N) = 0 \quad \forall N \in \mathcal{N}.$$

Now, let $E \cup F \in \mathcal{M}$. Then $E \in \mathcal{M}$, and thus $E \in \mathcal{A}$ is also true. Moreover, $F \subset N \in \mathcal{N}$, and so

$$N \in \mathcal{A}, \quad \varphi(N) = \mu(N) = 0.$$

Since $\mathcal{A}$ is complete, by the completeness proposition, we have that

$$F \in \mathcal{A} \implies E \cup F \in \mathcal{A}.$$

We have therefore proven that $\bar{\mathcal{M}} \subset \mathcal{A}$. 

Proposition 167 (Null Set Proposition)**.** *Let $(X, \mathcal{M}, \mu)$ be a non-trivial measure space, meaning there exist measurable subsets of positive measure. Then*

$$\mathcal{N} := \{Y \in \mathcal{M} : \mu(Y) = 0\}$$

is not *a σ-algebra, but it* is *closed under countable unions.*

Proof: If $\{N_n\} \subset \mathcal{N}$ is a countable collection, then since $\mathcal{M}$ is a σ-algebra,

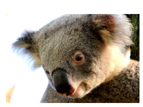
$$\cup N_n \in \mathcal{M}.$$

Moreover, we have

$$\mu(\cup N_n) \leq \sum \mu(N_n) = 0 \implies \mu(\cup N_n) = 0.$$

This shows that $\mathcal{N}$ is closed under countable unions. Why is it however, *not* a σ-algebra? It's not even an algebra! This is because it is not closed under complements. What is always an element of $\mathcal{N}$? The $\emptyset$ is always measurable and has measure zero. Hence $\emptyset \in \mathcal{N}$. What about its complement? This is where the non-triviality hypothesis plays a role. There is some $Y \in \mathcal{M}$ such that $\mu(Y) > 0$. Since $Y \subset X$, by monotonicity

$$\mu(X) \geq \mu(Y) > 0 \implies X = \emptyset^c \notin \mathcal{N}.$$



We shall now see that once we have an outer measure, we can build a sigma algebra and a measure, and obtain a complete measure space!

Theorem 168 (Carathéodory: creating a measure from an outer measure)**.** *Let μ^* be an outer measure on X. A set $A \subset X$ is called measurable with respect to $\mu^* \Leftrightarrow \forall\ E \subset X$ the following equation holds:*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c). \qquad (*)$$

Then $\mathcal{M} := \{A \subset X | A$ is μ^* measurable$\}$ is a σ-algebra and $\mu^*\big|_{\mathcal{M}}$ is a complete measure.

Proof: Note that $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$ because (*) is symmetric in A and A^c. Since $\mu^*(\emptyset) = 0$, we have

$$\mu^*(E \cap \emptyset) + \mu^*(E \cap \emptyset^c) = \mu^*(\emptyset) + \mu^*(E \cap X) = 0 + \mu^*(E) = \mu^*(E).$$

Consequently, $\emptyset \in \mathcal{M}$.

Next we will show that $\mathcal{M}$ is closed under finite unions of sets. For $A, B \in \mathcal{M}$ and $E \subset X$ we get, by multiple use of (*):

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*((E \cap A) \cap B) + \mu^*((E \cap A) \cap B^c) \\ + \mu^*((E \cap A^c) \cap B) + \mu^*((E \cap A^c) \cap B^c).$$

Furthermore, we can write $A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$, so that

$$E \cap (A \cup B) = (E \cap (A \cap B)) \cup (E \cap (A \cap B^c)) \cup (E \cap (A^c \cap B)),$$

so by countable subadditivity of outer measures, we have

$$\mu^*(E \cap A \cap B) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A \cap B^c) \geq \mu^*(E \cap (A \cup B))$$

Since $E \cap A^c \cap B^c = E \cap (A \cup B)^c$, using this inequality in the above equation gives us:

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c).$$

Moreover, by countable subadditivity of outer measures,

$$\mu^*(E) = \mu^*[(E \cap (A \cup B)) \cup (E \cap (A \cup B)^c)] \leq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)).$$

So the inequality is actually an equality, since we have shown that

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c) \geq \mu^*(E).$$

Hence $A \cup B \in \mathcal{M}$.

Next we show that μ^* is finitely-additive:

$$\forall\, A, B \in \mathcal{M},\ A \cap B = \emptyset \Rightarrow \mu^*(A \cup B) = \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) = \mu^*(A) + \mu^*(B).$$

Now we will show that $\mathcal{M}$ is actually a σ-algebra: For $\{A_j\}_{j \in \mathbb{N}} \subset \mathcal{M}$ we can define a sequence of disjoint sets $\{B_j\}_{j \in \mathbb{N}} \subset \mathcal{M}$ fulfilling $\bigcup_{j \in \mathbb{N}} A_j = \bigcup_{j \in \mathbb{N}} B_j$ by:

$$B_1 := A_1, \quad B_n := A_n \setminus \cup_{k=1}^{n-1} B_k, \quad n \geq 2.$$

Let us also define

$$\tilde{B}_n := \bigcup_{j=1}^{n} B_j.$$

Then since $\mathcal{M}$ is closed under finite unions of sets and also closed under complementation, both

$$\tilde{B}_n \in \mathcal{M}, \quad B_n \in \mathcal{M}.$$

So, we need to show that

$$\bigcup_{j \in \mathbb{N}} A_n = \bigcup_{j \in \mathbb{N}} B_j \in \mathcal{M}.$$

For $E \subset X$, since $B_n \in \mathcal{M}$,

$$\mu^*(E \cap \tilde{B}_n) \overset{(*)}{=} \mu^*(E \cap \tilde{B}_n \cap B_n) + \mu^*(E \cap \tilde{B}_n \cap {B_n}^c) = \mu^*(E \cap B_n) + \mu^*(E \cap \tilde{B}_{n-1})$$

Thus $\mu^*(E \cap \tilde{B}_n) = \mu^*(E \cap B_n) + \mu^*(E \cap \tilde{B}_{n-1})$. Repeating this argument, we have $\mu^*(E \cap \tilde{B}_{n-1}) = \mu^*(E \cap B_{n-1}) + \mu^*(E \cap \tilde{B}_{n-2})$. Continuing inductively, we have:

$$\mu^*(E \cap \tilde{B}_n) = \mu^*(E \cap B_n) + \mu^*(E \cap B_{n-1}) + \mu^*(E \cap \tilde{B}_{n-2}) = ... = \sum_{k=1}^{n} \mu^*(E \cap B_k)$$

Using this result together with the fact that $\tilde{B}_n \in \mathcal{M}$, we get:

$$\begin{aligned}\mu^*(E) = \mu^*(E \cap \tilde{B}_n) + \mu^*(E \cap \tilde{B}_n^c) &= \sum_{k=1}^{n} \mu^*(E \cap B_k) + \mu^*(E \cap \tilde{B}_n^c) \\ &\geq \sum_{k=1}^{n} \mu^*(E \cap B_k) + \mu^*(E \setminus (\bigcup_{k=1}^{\infty} B_k))\end{aligned}$$

Above, we have used that

$$E \cap \tilde{B}_n^c = E \cap (\cup_{k=1}^{n} B_n)^c = E \setminus \cup_{k=1}^{n} B_n \supset E \setminus \bigcup_{k=1}^{\infty} B_k,$$

together with the fact that outer measures are monotone. This inequality holds for any $n \in \mathbb{N}$, so we obtain

$$(**) \qquad \mu^*(E) \geq \sum_{k=1}^{\infty} \mu^*(E \cap B_k) + \mu^*(E \setminus (\bigcup_{k=1}^{\infty} B_k)).$$

Since

$$E \cap (\cup_{k=1}^{\infty} B_k) = \cup_{k=1}^{\infty} E \cap B_k,$$

by countable subadditivity of out measures,

$$\mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)) \leq \sum_{k=1}^{\infty} \mu^*(E \cap B_k).$$

We therefore obtain, combining this with the above inequality

$$\mu^*(E) \geq \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)) + \mu^*(E \setminus (\bigcup_{k=1}^{\infty} B_k)) = \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)) + \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)^c).$$

Since $E \subset (E \cap Y) \cup (E \cap Y^c)$, by countable subadditivity of outer measures, for any Y we have

$$\mu^*(E) \leq \mu^*(E \cap Y) + \mu^*(E \cap Y^c).$$

We therefore also have the inequality

$$\mu^*(E) \leq \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)) + \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)^c).$$

Combining with the reverse inequality we demonstrated above, we obtain

$$\mu^*(E) = \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)) + \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)^c).$$

This shows that $\cup B_k$ satisfies the definition of $\mathcal{M}$, so we have

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k \in \mathcal{M}.$$

Hence $\mathcal{M}$ is a σ-algebra.

Now we want to show that $\mu^*\big|_{\mathcal{M}}$ is a measure. First we note that since μ^* is an outer measure, we have $\mu^*(\emptyset) = 0$. Moreover, outer measures are also monotone, so μ^* is monotone. Thus, we only need to show that μ^* restricted to $\mathcal{M}$ is countably additive. Let $\{B_k\}_{k\in\mathbb{N}} \subset \mathcal{M}$ be pairwise disjoint sets. Defining $E := \bigcup_{k=1}^{\infty} B_k$ and using (**), we get

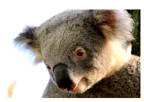
$$\mu^*(\bigcup_{k=1}^{\infty} B_k) = \mu^*(E) \overset{(**)}{\geq} \sum_{k=1}^{\infty} \mu^*(E \cap B_k) + \mu^*(\emptyset) = \sum_{k=1}^{\infty} \mu^*(B_k) \geq \mu^*(\bigcup_{k=1}^{\infty} B_k)$$

$$\implies \mu^*(\bigcup_{k=1}^{\infty} B_k) = \sum_{k=1}^{\infty} \mu^*(B_k)$$

So $\mu^*\big|_{\mathcal{M}}$ is a measure.

Finally, we show that it is a complete measure: For $Y \subset X$ such that $\mu^*(Y) = 0$, and for arbitrary $E \subset X$ we have by countable subadditivity of outer measures

$$\mu^*(E) \leq \mu^*(E \cap Y) + \mu^*(E \cap Y^c) \leq \mu^*(Y) + \mu^*(E) = \mu^*(E)$$

Therefore $Y \in \mathcal{M}$. 

Remark 7. We briefly discussed the proof of completion, and I shall add a remark here. Technically speaking, we should be considering

$$\mu^{**} : P(X) \to [0, \infty], \quad \mu^{**}(A) = \inf\{\sum_{j\geq 1} \mu^*(E_j) : A \subset \cup_{j\geq 1} E_j, \quad E_j \in \mathcal{M}\}.$$

If some set has $\mu^{**}(Y) = 0$, then for each $k \in \mathbb{N}$ there exists $\{E_j^k\} \in \mathcal{M}$ such that

$$Y \subset \cup_{j\geq 1} E_j^k, \quad \sum_{j\geq 1} \mu^*(E_j^k) < 2^{-k}.$$

Since $\mathcal{M}$ is a σ-algebra,

$$A_k := \cup_{j\geq 1} E_j^k \in \mathcal{M},$$

and

$$\mu^*(A_k) \leq \sum_{j\geq 1} \mu^*(E_j^k) < 2^{-k}.$$

Moreover, since $Y \subset A_k$ for all k, we have

$$Y \subset \cap_{k\geq 1} A_k,$$

and we also have that since $\mathcal{M}$ is a σ-algebra

$$\cap_{k\geq 1} A_k \in \mathcal{M}.$$

Since

$$\cap_{k\geq 1} A_k \subset A_n \quad \forall n \in \mathbb{N},$$

by monotonicity,

$$\mu^*(Y) \leq \mu^*(\cap_{k\geq 1} A_k) \leq 2^{-n} \quad \forall n \in \mathbb{N}.$$

This shows that $\mu^*(Y) = 0$. It is pretty straightforward to show that the converse holds as well, that is if $\mu^*(Z) = 0$ then $\mu^{**}(Z) = 0$. So, by the completeness proposition, our μ^* is complete!

Another important concept in measure theory is that of a *pre-measure.*

Definition 169. Let $\mathcal{A} \subset P(X)$ be an algebra. A function $\mu_0 : \mathcal{A} \to [0, \infty]$ is called a *pre-measure* if

1. $\mu_0(\emptyset) = 0$

2. If $\{A_j\}$ is a countable collection of disjoint elements of $\mathcal{A}$ such that

$$\cup A_j \in \mathcal{A},$$

then

$$\mu_0(\cup A_j) = \sum \mu_0(A_j).$$

Exercise 170. *We have shown how, given a measure space $(X, \mathcal{M}, \mu)$, we can obtain a minimal complete measure space, $(X, \bar{\mathcal{M}}, \bar{\mu})$. We have also shown how, given a measure, μ, we can canonically construct an outer measure, μ^*.*

1. *Using the canonically associated outer measure, μ^*, determine whether or not the set*

$$\mathcal{A} := \{A \in P(X) : \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ holds true for all } E \subset X\}$$

is equal to the set

$$\bar{\mathcal{M}} := \{E \cup F : E \in \mathcal{M}, \text{ and } F \subset N \in \mathcal{N}\},$$

where again $\mathcal{N}$ is the set of elements of $\mathcal{M}$ which have μ-measure zero.

2. *In this way, determine whether or not the spaces*

$$(X, \mathcal{A}, \mu^*)$$

and

$$(X, \bar{\mathcal{M}}, \bar{\mu})$$

are the same? My sneaking suspicion is that they are the same, but I shall not spoil your fun in investigating this question.

D.2.1 Exercises: Constructing the Lebesgue measure

The n-dimensional Lebesgue measure is the unique, complete measure which agrees with our intuitive notion of n-dimensional volume. To make this precise, first we define a generalized interval and our notion of intuitive volume.

Definition 171. A generalized interval in $\mathbb{R}^n$ is a set for which there exist real numbers $a_k \le b_k$ for $k = 1, \dots n$, such that this set has the form

$$I = \{x \in \mathbb{R}^n, x = \sum x_k e_k, \quad a_k < \text{ or } \le x_k < \text{ or } \le b_k, k = 1, \dots, n\}.$$

Above we are using e_k to denote the standard unit vectors for $\mathbb{R}^n$. The intuitive volume function on $\mathbb{R}^n$ is defined on such a set to be

$$v_n(I) = \prod (b_k - a_k).$$

Next we can extend our intuitive notion of volume to *elementary sets.*

Definition 172. An *elementary subset* of $\mathbb{R}^n$ is a set which can be expressed as a finite disjoint union of generalized intervals. The collection of all of these is denoted by $\mathcal{E}_n$.

Exercise 173. *Prove that v_n is well-defined on $\mathcal{E}_n$.*

Exercise 174. *To make an algebra containing $\mathcal{E}_n$, in particular the smallest algebra containing $\mathcal{E}_n$, it is necessary to include compliments. Define*

$$\mathcal{A} := \{Y \subseteq \mathbb{R}^n \mid Y \in \varepsilon_n \text{ or } \exists Z \in \varepsilon_n \text{ s.t. } Y = Z^c\}$$

Prove that $\mathcal{A}$ is an algebra.

Exercise 175. *Show that ν_n is well-defined on $\mathcal{A}$ where*

$$\nu_n(\prod^n Ia_i, \alpha_i I) := \begin{cases} 0, & \text{if } a_i = \alpha_i \text{ for some } i \\ \prod(\alpha_i - a_i), & \text{else} \end{cases}$$

Exercise 176. *Show that ν_n is a pre-measure on $\mathcal{A}$.*

D.2.2 Hints

1. $\emptyset = \prod Ix, xI$ for $x \in \mathbb{R}^n$. *Notation:* we use Ia, bI to denote either $]a, b[$, $[a, b]$, $]a, b]$ or $[a, b[$. Notation which is unnecessary shall be simplified when possible.

2. Show that $\mathcal{A}$ is closed under compliments

3. Let $A, B \in \mathcal{A}$. If $A, B \in \varepsilon_n$ then first consider the case where A, B are each single intervals i.e. $A = \prod Ia_i, \alpha_i I, B = \prod Ib_i, \beta_i I$ for $a_i \leq \alpha_i$, $b_i \leq \beta_i$. For each i, if $Ib_i, \beta_i I \subset Ia_i, \alpha_i I$ then note that

$$Ia_i, \alpha_i I \backslash Ib_i, \beta_i I = Ia_i, b_i I \cup I\beta_i, \alpha_i I$$

If $Ib_i, \beta_i I \not\subset Ia_i, \alpha_i I$, then either $Ib_i, \beta_i I \cap Ia_i, \alpha_i I = \emptyset$ in which case $Ia_i, \alpha_i I \backslash Ib_i, \beta_i I = Ia_i, \alpha_i I$, or $Ib_i, \beta_i I \cap Ia_i, \alpha_i I \neq \emptyset$ so that

$$Ia_i, \alpha_i I \backslash Ib_i, \beta_i I = \begin{cases} Ia_i, b_i I \text{ if } b_i \leq \alpha_i (\Rightarrow \beta_i > \alpha_i) \\ I\beta_i, \alpha_i I \text{ if } a_i \leq beta_i (\Rightarrow b_i < a_i) \end{cases}$$

In both cases $Ia_i, \alpha_i I \backslash Ib_i, \beta_i I$ is the disjoint union of intervals. Repeating for each $i = 1, ..., n$ gives $A \backslash B \in \varepsilon_n$, and similarly $B \backslash A \in \varepsilon_n$. Note that $A \cap B = \prod Ix_i, y_i I$ with $x_i = \max\{a_i, b_i\}$, $y_i = \min\{\alpha_i, \beta_i\}$ (and should $x_i \geq y_i$ then it is understood that $Ix_i, y_i I = \emptyset$. Therefore,

$$A \cup B = (A \backslash B) \cup (B \backslash A) \cup (A \cap B) \in \varepsilon_n.$$

In fact, for $A = \prod Ia_i, \alpha_i I \in \varepsilon_n$ note that

$$\begin{aligned} A^c =& \mathbb{R}^n \backslash A \\ =& \prod I - \infty, a_i I \cup \prod I\alpha_i, \infty I \end{aligned}$$

Allowing the endpoints x_i and/or y_i of $Ix_i, y_i I$ to be $\pm\infty$, the same arguments for A, B as above show that $A^c \cup B$ and $A^c \cup B^c$ are elements of $\mathcal{A}$.

More generally, for $A = \bigcup_{j=1}^k I_j \in \varepsilon_n$ with $I_j \underset{k \neq j}{\cap} I_k = \emptyset$ and $B = \bigcup_{l=1}^m J_l \in \varepsilon_n$ with $J_l \underset{m \neq l}{\cap} J_m = \emptyset$ with end points possibly $\pm\infty$, repeated application of the above arguments shows that $I_1 \cup J_1 \in \varepsilon_n$, $(I_1 \cup J_1) \cup I_2 \in \varepsilon_n$, and so forth. Therefore, $A \cup B \in \varepsilon_n$. So $\mathcal{A}$ is closed under finite unions and hence $\mathcal{A}$ is an algebra.

4. To show that ν_n is well-defined on $\mathcal{A}$ and that it is a pre-measure, first show that $\nu_n(\emptyset) = 0$.

5. Next, let $\{A_m\}_{m\geq 1} \subset \mathcal{A}$ such that $\underset{m\geq 1}{\cup} A_m \in \mathcal{A}$, $A_m \underset{k\neq m}{\cap} A_k = \emptyset$ then $\exists \{I_j\}_{j=1}^k$ disjoint in $\mathcal{A}$ such that $\bigcup\limits_{j=1}^k I_j = \bigcup\limits_{m=1}^\infty A_m$.

   By definition, $\nu_n(\bigcup\limits_{m=1}^M A_m) = \sum\limits_{m=1}^M \upsilon_n(A_m) \leq \nu_n(\bigcup\limits_{j=1}^k I_j) = \sum\limits_{j=1}^k \upsilon_n(I_j)$

   $\forall M \in \mathbb{N}$, $\sum\limits_{m=1}^M \upsilon_n(A_m) \leq \sum\limits_{j=1}^k \upsilon_n(I_j) = \nu_n(\bigcup\limits_{m=1}^\infty A_m) \leq \sum\limits_{m=1}^M \upsilon_n(A_m)$

   $\Rightarrow \nu_n(\bigcup\limits_{m=1}^\infty A_m) = \sum\limits_{m=1}^M \upsilon_n(A_m)$

D.3 Pre-measure extension theorem and metric outer measures

The name pre-measure is appropriate because it's almost a measure, it's just possibly not countably additive for *every* disjoint countable union, since these need not *always* be contained in a mere algebra (which is not necessarily a σ-algebra). However, Carathéodory can help us to extend pre-measures to measures. First, we require the following.

Proposition 177. *Let μ_0 be a pre-measure on the algebra $\mathcal{A} \subset P(X)$, and define*

$$\mu^*(Y) := \inf\{\sum_j \mu_0(A_j) : A_j \in \mathcal{A}\forall j, Y \subset \cup_j A_j\},$$

where the infimum is taken to be ∞ if there is no such cover of Y. Then we have:

1. *μ^* is an outer measure.*
2. *$\mu^*(A) = \mu_0(A) \forall A \in \mathcal{A}$.*
3. *Every set in $\mathcal{A}$ is μ^* measurable in the same sense as above, being that for arbitrary $E \subset X$, for $A \in \mathcal{A}$,*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Proof: First, note that $\emptyset \in \mathcal{A}$ since $\mathcal{A}$ is an algebra. Moreover, the map μ_0 is defined on $\mathcal{A}$, with $\mu_0 : \mathcal{A} \to [0, \infty]$, and has $\mu_0(\emptyset) = 0$. Therefore by the Outer Measure Proposition, as we have defined μ^*, it is an outer measure.

Next, we wish to show that μ^* and μ_0 are the same when we restrict to the algebra, $\mathcal{A}$. To do this we will show that (1) pre-measures are finitely additive and (2) pre-measures are monotone.

Finite additivity of pre-measures: Next, we show that pre-measures are by definition finitely additive since for $A, B \in \mathcal{A}$ with $A \cap B = \emptyset$, then

$$A \cup B = \cup A_j, \quad A_1 = A, A_2 = B, A_j = \emptyset \forall j > 2,$$

gives

$$\mu_0(A \cup B) = \mu_0(\cup A_j) = \sum \mu_0(A_j) = \mu_0(A) + \mu_0(B).$$

Monotonicity of pre-measures: Assume that $A \subset B$ are both elements of $\mathcal{A}$. Then $B \setminus A = B \cap A^c \in \mathcal{A}$, so finite additivity gives

$$\mu_0(B) = \mu_0(B \setminus A) + \mu_0(A) \implies \mu_0(A) = \mu_0(B) - \mu_0(B \setminus A) \leq \mu_0(B).$$

Showing that $\mu^* = \mu_0$ on $\mathcal{A}$: Now, let $E \in \mathcal{A}$. If $E \subset \cup A_j$ with $A_j \in \mathcal{A}$ $\forall j$, then let

$$B_n := E \cap (A_n \setminus \cup_1^{n-1} A_j).$$

Then

$$B_n \in \mathcal{A} \forall n, \quad B_n \cap B_m = \emptyset \forall n \neq m.$$

The union

$$\cup B_n = \cup E \cap (A_n \setminus \cup_1^{n-1} A_j) = E \cap \cup (A_n \setminus \cup_1^{n-1} A_j) = E \cap \cup A_n = E \in \mathcal{A}.$$

So by definition of pre-measure,

$$\mu_0(E) = \mu_0(\cup B_n) = \sum \mu_0(B_n) \leq \sum \mu_0(A_n),$$

since $B_n \subset A_n \forall n$. Taking the infimum over all such covers of E comprised of elements of $\mathcal{A}$, we have

$$\mu_0(E) \leq \mu^*(E).$$

On the other hand, $E \subset \cup A_j$ with $A_1 = E \in \mathcal{A}$, and $A_j = \emptyset \forall j > 1$. Then, this collection is considered in the infimum defining μ^*, so

$$\mu^*(E) \leq \sum \mu_0(A_j) = \mu_0(E).$$

We've shown the inequality is true in both directions, hence $\mu^*(E) = \mu_0(E)$ for any $E \in \mathcal{A}$.

Showing that $\mathcal{A}$ sets are μ^* measurable: Let $A \in \mathcal{A}$, $E \subset X$, and $\varepsilon > 0$. Since we always have by countable subadditivity

$$\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c),$$

if $\mu^*(E) = \infty$, then we also have

$$\infty \leq \mu^*(E \cap A) + \mu^*(E \cap A^c) \implies \mu^*(E \cap A) + \mu^*(E \cap A^c) = \infty,$$

so the equality holds. Now, let us assume that $\mu^*(E) < \infty$. Then, by its definition, there exists $\{B_j\} \subset \mathcal{A}$ with $E \subset \cup B_j$ and

$$\sum \mu_0(B_j) \leq \mu^*(E) + \varepsilon.$$

Since μ_0 is additive on $\mathcal{A}$,

$$\mu^*(E) + \varepsilon \geq \sum \mu_0(B_j \cap A) + \mu_0(B_j \cap A^c) = \sum \mu_0(B_j \cap A) + \sum \mu_0(B_j \cap A^c) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Above we have used the definition of μ^* as an infimum, together with the fact that since $A \in \mathcal{A}$ and $B_j \in \mathcal{A}$ for all j, we have $B_j \cap A \in \mathcal{A}$ and $B_j \cap A^c \in \mathcal{A}$ for all j, and we also have

$$E \cap A \subset \cup B_j \cap A, \quad E \cap A^c \subset \cup B_j \cap A^c.$$

This is true for any $\varepsilon > 0$, so we have

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c) \geq \mu^*(E).$$

So, these are all equal, which shows that A satisfies the definition of being μ^* measurable since E was arbitrary. 

Now we will prove that we can always extend a pre-measure to a measure. You will use this in the first exercise to complete the construction of the Lebesgue measure.

Theorem 178 (Pre-measure extension theorem)**.** *Let $\mathcal{A} \subset P(X)$ be an algebra, μ_0 a pre-measure on $\mathcal{A}$, and $\mathcal{M}$ the smallest σ-algebra generated by $\mathcal{A}$. Then there exists a measure μ on $\mathcal{M}$ which extends μ_0, namely*

$$\mu := \mu^* \text{ restricted to } \mathcal{M}.$$

If ν also extends μ_0 then $\nu(E) \leq \mu(E) \forall E \in \mathcal{M}$ with equality when $\mu(E) < \infty$. If μ_0 is σ-finite, then $\nu \equiv \mu$ on $\mathcal{M}$, so μ is the unique extension.

Proof: By its very definition, $\mathcal{M}$ is a σ-algebra, and all elements of $\mathcal{A}$ are contained in $\mathcal{M}$. Moreover, by the proposition,

$$\mu^*(A) = \mu_0(A), \quad \forall A \in \mathcal{A}.$$

Since $\emptyset \in \mathcal{A}$, we have

$$\mu^*(\emptyset) = 0.$$

Moreover, since μ^* is an outer measure, by the proposition, it is monotone. Consider the set

$$\{A \subset X | \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ holds true for all } E \subset X\}.$$

By the preceding proposition, this set contains all elements of $\mathcal{A}$. Moreover, since μ^* is an outer measure, in Caratheodory's Theorem, we proved that this set is a σ algebra, and μ^* restricted to this set is a measure. Hence, since it is a σ algebra which contains $\mathcal{A}$, it also contains $\mathcal{M}$. Therefore, μ^* restricted to $\mathcal{M}$ is countably additive, since μ^* on this larger (Carathéodory-Theorem-σ-algebra-set) is countably additive. Hence μ is a measure.

So, we only need to consider the statements about a possibly different extension ν which coincides with μ_0 on $\mathcal{A}$ and is a measure on $\mathcal{M}$. If $E \in \mathcal{M}$ and

$$E \subset \cup A_j, \quad A_j \in \mathcal{A} \forall j,$$

then

$$\nu(E) \le \sum \nu(A_j) = \sum \mu_0(A_j).$$

This holds for any such covering of E by elements of $\mathcal{A}$, so taking the infimum we have

$$\nu(E) \le \mu^*(E) = \mu(E) \text{ since } E \in \mathcal{M}.$$

If $\mu(E) < \infty$, let $\varepsilon > 0$. Then we may choose $\{A_j\} \subset \mathcal{A}$ which are WLOG (without loss of generality) disjoint (why/how can we do this?) such that

$$E \subset \cup A_j, \quad \mu(\cup A_j) = \sum \mu_0(A_j) < \mu^*(E) + \varepsilon = \mu(E) + \varepsilon,$$

since $E \in \mathcal{M}$. Note that $E \in \mathcal{M}$, $\{A_j\} \subset \mathcal{A}$, and $\mathcal{M}$ is a σ algebra containing $\mathcal{A}$. We therefore have

$$A := \cup A_j \in \mathcal{M}.$$

Then, we also have

$$\nu(A) = \lim_{n\to\infty} \nu(\cup_1^n A_j) = \lim_{n\to\infty} \sum_1^n \nu(A_j) = \lim_{n\to\infty} \sum_1^n \mu_0(A_j) = \mu(A).$$

Then we have since $E \in \mathcal{M}$, and $\{A_j\} \subset \mathcal{A}$, and $\mathcal{M}$ is a σ algebra containing $\mathcal{A}$ that $A \in \mathcal{M}$. By countable additivity of the measure μ, we have

$$\mu(\cup A_j) = \mu(A) = \mu(A \cap E) + \mu(A \setminus E) = \mu(E) + \mu(A \setminus E) < \mu(E) + \varepsilon$$

which shows that

$$\mu(A \setminus E) < \varepsilon.$$

Consequently, using monotonicity, the fact that $\mu(A) = \nu(A)$, the additivity of ν, and the fact that $\nu \le \mu$, we obtain

$$\mu(E) \le \mu(A) = \nu(A) = \nu(E \cap A) + \nu(A \setminus E) \le \nu(E) + \mu(A \setminus E) < \nu(E) + \varepsilon.$$

This holds for all $\varepsilon > 0$, so

$$\mu(E) \le \nu(E).$$

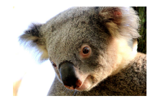
Consequently in this case $\mu(E) = \nu(E)$, whenever these are finite.

Finally, if $X = \cup A_j$ with $A_j \in \mathcal{A}$, $\mu_0(A_j) < \infty \forall j$, we may WLOG assume the A_j are disjoint. Then for $E \in \mathcal{M}$,

$$E = \cup(E \cap A_j),$$

which is a disjoint union of elements of $\mathcal{M}$. So by countable additivity

$$\mu(E) = \mu(\cup E \cap A_j) = \sum \mu(E \cap A_j) = \sum \nu(E \cap A_j) = \nu(\cup E \cap A_j) = \nu(E),$$

since $E \cap A_j \subset A_j$ shows that $\mu(E \cap A_j) \leq \mu(A_j) < \infty$, so $\mu(E \cap A_j) = \nu(E \cap A_j)$. 

D.4 Metric outer measures

To define the Hausdorff measure, we will introduce *metric outer measures.* A metric outer measure requires an addition type of structure on the big set X: we need a notion of *distance between points.* Thus, metric outer measures are only defined when the set X also carries along a distance, d, also known as a *metric.* So, for a metric space (X, d) and for $A, B \subset X$ define

$$\text{dist}(A, B) := \inf\{d(x, y) : x \in A, y \in B\}.$$

Define also the diameter of a set $A \subset X$

$$\text{diam}(A) := \sup\{d(x, y) : x, y \in A\}, \text{diam}(\emptyset) := 0.$$

Definition 179. Let μ^* be an outer measure defined on a metric space, (X, d). Then μ^* is called *metric outer measure* iff for each $A, B \subset X$ we have

$$\text{dist}(A, B) > 0 \Rightarrow \mu^*(A \cup B) = \mu^*(A) + \mu^*(B).$$

Recall: $A \subset X$ is μ^*-measurable iff for each $E \subset X$

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^C).$$

Denote by $\mathfrak{M}(\mu^*)$ the μ^*-measurable subsets. Recall that the Borel sets $\mathcal{B}(X)$ is the smallest σ-algebra generated by the topology of X (induced by the metric). In other words, it is the smallest σ-algebra which contains all open sets. We note that $\emptyset$ is both open and closed. A non-empty subset, U, of a metric space (X, d) is defined to be open precisely when

$$\forall x \in U \exists \delta > 0 \text{ such that } B_\delta(x) \subset U,$$

where

$$B_\delta(x) = \{y \in X | d(x, y) < \delta\}.$$

A subset of X is said to be closed precisely when its complement is open. We now prove a Theorem due to Carathéodory which states that the Borel sets in X are contained in $\mathfrak{M}(\mu^*)$.

Theorem 180 (Carathéodory)**.** *Let μ^* be a metric outer measure on (X, d). Then we have $\mathcal{B}(X) \subset \mathfrak{M}(\mu^*)$.*

Exercise 181. *Show that μ^* is a measure on $\mathcal{B}(X)$. Denote by μ the restriction of μ^* to $\mathcal{B}(X)$. Let $\mathcal{A}$ be defined as in the completion theorem, that is:*

$$\mathcal{A} = \{E \cup F : E \in \mathcal{B}(X), F \subset N \in \mathcal{B}(X), \quad \mu^*(N) = 0\}.$$

Define as in the completion $\bar{\mu}(E \cup F) = \mu(E)$. Is it true that $\mathfrak{M}(\mu^*) = \mathcal{A}$? Prove your answer.

D.4.1 Exercises: properties of the Lebesgue σ-algebra

1. In the previous exercises, we proved that ν_n is a pre-measure on the algebra $\mathcal{A}$. Note by the definition of $\mathcal{A}$, it is the smallest algebra which contains ε_n. By the pre-measure extension theorem, since ν_n is σ-finite on $\mathcal{A}$, there exists a unique extension of ν_n to a measure $\bar{\mathcal{M}}$ on the smallest σ-algebra containing ε_n. It is unique, because $\mathbb{R}^n = \underset{m\geq 1}{\cup}[-M,M]^n = \underset{m\geq 1}{\cup} I_M$ and $\nu_m(I_M) = (2M)^n < \infty$ for each M. Canonically completing this measure to $\mathcal{M}$ by applying the completion theorem yields the Lebesgue measure and the Lebesgue σ-algebra, the smallest σ-algebra generated by ε_n such that the extension of ν_n to a measure with respect to this σ-algebra is complete. This is the construction of the Lebesgue measure. In this exercise, the task is to review the construction of the Lebesgue measure step-by-step, and make sure it makes sense to you.

2. Prove that Borel sets are Lebesgue measurable.

3. Prove $\mathcal{B} \subsetneq \mathcal{M}$

4. It is difficult to construct sets $\not\subset \mathcal{M}$, but actually there are many natural examples... *Exercise*: Construct a subset of $\mathbb{R}^n$ which is not measurable. Recall that $f : \mathbb{R}^n \to \mathbb{R}^m$ is "measurable" usually is understood to mean that $\forall B \in \mathcal{B}^m$, $f^{-1}(B) \in \mathcal{M}^n$. More precisely, f is $(\mathbb{R}^n, \mathcal{B}^n), (\mathbb{R}^m, \mathcal{B}^m)$ measurable. In general, $f : X \to Y$ is $(X, \mathcal{A}), (Y, \mathcal{B})$ measurable if $\forall B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$, where $\mathcal{A}$ and $\mathcal{B}$ are σ-algebras.

5. Prove that all $n-1$ dimensional sets have $\mathcal{L}^n$ measure 0.

D.4.2 Hints

1. To prove that Borel sets are Lebesgue measurable, it suffices to show that open sets are Lebesgue measurable. So, let $\mathcal{O} \subset \mathbb{R}^n$ be open. Then we will show that $\mathcal{O} \in \mathcal{M}$.

   First consider $\mathcal{O} = \prod]a_i, \alpha_i[\in \varepsilon_n \subset \mathcal{M}$. For an arbitrary open set $\mathcal{O}$, for each $x \in \mathcal{O}$ there exists $\varepsilon \in \mathbb{Q}, \varepsilon > 0$ such that $x \in \prod\,]q_m - \varepsilon, q_m + \varepsilon[\subset \mathcal{O}$, $q_m \in \mathbb{Q}$, $m = 1, ..., n$.

   Taking the union of all such intervals, namely those contained in $\mathcal{O}$ such that endpoints are rational is a countable union. Countability of course follows since $\mathbb{Q}^n \subset \mathbb{R}^n$ is countable and $\mathbb{Q}$ is countable so a union of intervals with endpoints in $\mathbb{Q}^n$ is countable. Therefore, $\mathcal{O} \in \mathcal{M}$.

In the following theorem, we prove that Borel sets are always measurable with respect to metric outer measures.

Theorem 182 (Carathéodory). *Let μ^* be a metric outer measure on (X, d). Then we have $\mathcal{B}(X) \subset \mathfrak{M}(\mu^*)$.*

Proof:

Note that since $\mathfrak{M}(\mu^*)$ is a σ-algebra (by Thm. 168) it is enough to prove that every closed set is μ^*-measurable. (why does this suffice?) So let $F \subset X$ be a closed subset. Since the reverse inequality always holds, it will be enough to prove that for any set A we have

$$\mu^*(A) \geq \mu^*(A \cap F) + \mu^*(A \setminus F).$$

Define the sets

$$A_k := \left\{ x \in A : \operatorname{dist}(x, F) \geq \frac{1}{k} \right\}.$$

Then $\operatorname{dist}(A_k, A \cap F) \geq \frac{1}{k}$, so since μ^* is a metric outer measure we have

$$\mu^*(A \cap F) + \mu^*(A_k) = \mu^*(\underbrace{(A \cap F) \cup A_k}_{\subset A}) \leq \mu^*(A). \tag{+}$$

Let $x \in A \setminus F = A \cap F^c$. Since F^c is open, there exists $\delta > 0$ such that $B_\delta(x) \subset F^c$. Hence $d(x, F) \geq \delta$. So, in general, for all $x \in A \setminus F$, we have

$$\operatorname{dist}(x, F) > 0.$$

Consequently, we have

$$A \setminus F = \bigcup A_k.$$

The main and last step in the proof is to calculate the limit in (+) as $k \to \infty$. If the limit is infinity there is nothing to do, because it shows that

$$\mu^*(A) = \infty \geq \text{ anything we want, in particular } \geq \mu^*(A \cap F) + \mu^*(A \cap F^c).$$

So, let us assume that the limit in (+) is finite.

Note that $A_1 \subset A_2 \subset A_3 \subset \ldots$.

To get a bit of room between our sets, let us define

$$B_1 := A_1, \quad B_n := A_n \setminus A_{n-1}, \quad n \geq 2.$$

By definition, $A_k \subset A$, so we also have $B_k \subset A$ for all A. By definition of A_k and B_k, for all $x \in B_k$ we have

$$\frac{1}{k} \leq \operatorname{dist}(x, F) < \frac{1}{k-1},$$

where the second inequality follows since $B_k = A_k \setminus A_{k-1}$. Therefore if $j \geq k+2$, for all $y \in B_j$ we have

$$\frac{1}{j} \leq \operatorname{dist}(y, F) < \frac{1}{j-1} \leq \frac{1}{k+1} < \frac{1}{k} \leq \operatorname{dist}(x, F).$$

Let $\varepsilon > 0$ such that

$$\frac{1}{j-1} + \varepsilon < \frac{1}{k}.$$

By definition, there exists $z \in F$ such that

$$d(y, z) \leq \operatorname{dist}(y, F) + \varepsilon,$$

so

$$d(y, z) < \frac{1}{j-1} + \varepsilon < \frac{1}{k} \leq \operatorname{dist}(x, F) \leq d(x, z).$$

We therefore have by the triangle inequality,

$$d(x, y) \geq d(x, z) - d(z, y) \geq \frac{1}{k} - \left(\frac{1}{j-1} + \varepsilon \right) > 0.$$

Since $x \in B_k$ and $y \in B_j$ are arbitrary, and $\varepsilon > 0$ is fixed, we therefore have proven that $\operatorname{dist}(B_j, B_k) > 0$.

This means we can apply the metric outer measure property (for even and odd indices) and by induction we conclude that

$$\mu^* \left(\bigcup_{k=1}^n B_{2k-1} \right) = \sum_{k=1}^n \mu^*(B_{2k-1}),$$
$$\mu^* \left(\bigcup_{k=1}^n B_{2k} \right) = \sum_{k=1}^n \mu^*(B_{2k}).$$

These unions are each contained in A_{2n}, so we have the inequalities

$$\mu^*\left(\bigcup_{k=1}^{n} B_{2k-1}\right) = \sum_{k=1}^{n} \mu^*(B_{2k-1}) \leq \mu^*(A_{2n}),$$
$$\mu^*\left(\bigcup_{k=1}^{n} B_{2k}\right) = \sum_{k=1}^{n} \mu^*(B_{2k}) \leq \mu^*(A_{2n}).$$

Since $A_1 \subset A_2 \subset \ldots$, the values $\mu^*(A_{2n})$ are non-decreasing and by assumption bounded. Hence both sums above, since they are comprised of non-negative terms, are convergent as $n \to \infty$.

Therefore we conclude for any j

$$\begin{aligned}\mu^*(A \setminus F) &= \mu^*\left(\bigcup_i A_i\right) \\ &= \mu^*\left(A_j \cup \bigcup_{k \geq j+1} B_k\right) \\ &\leq \mu^*(A_j) + \sum_{k=j+1}^{\infty} \mu^*(B_k) \\ &\leq \lim_{n\to\infty} \mu^*(A_n) + \underbrace{\sum_{k=j+1}^{\infty} \mu^*(B_j)}_{\to 0, j\to\infty}.\end{aligned}$$

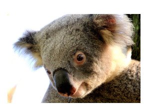
The last term tends to zero because it is comprised of the tails of two convergent series.

Since the latter sum goes to 0 by convergence we obtain

$$\mu^*(A \setminus F) \leq \lim_{n\to\infty} \mu^*(A_n).$$

Together with (+) this yields

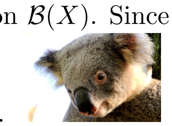
$$\mu^*(A) \geq \lim_{k\to\infty} \mu^*(A_k) + \mu^*(A \cap F) \geq \mu^*(A \setminus F) + \mu^*(A \cap F)$$

which is the desired inequality. 

Corollary 183. *Let (X, d) be a metric space, and let μ^* be a metric outer measure on X. Then μ^* restricted to the Borel sigma algebra is a measure, that is $(X, \mathcal{B}(X), \mu^*)$ is a measure space.*

Proof: By the theorem, $\mathcal{M}(\mu^*) \supset \mathcal{B}(X)$. In a previous theorem, we proved that μ^* restricted to $\mathcal{M}(\mu^*)$ is a measure. Note that $\emptyset \in \mathcal{B}(X)$ and $\mu^*(\emptyset) = 0$. If $\{A_j\} \subset \mathcal{B}(X)$ are pairwise disjoint, then since they are also contained in $\mathcal{M}(\mu^*)$ we have

$$\mu^*(\cup A_j) = \sum \mu^*(A_j).$$

Hence μ^* vanishes on the empty set and is countably additive on $\mathcal{B}(X)$. Since μ^* is defined on $\mathcal{B}(X)$ which is a σ-algebra, we have that μ^* restricted to $\mathcal{B}(X)$ is a measure. 

D.4.3 General results which shall be used to obtain the Hausdorff measure

We shall obtain the Hausdorff measure using results which can be applied much more generally to obtain metric outer measures.

Definition 184 (Countable covers)**.** Let $\mathcal{C}$ denote a collection of sets in X. Assume $\emptyset \in \mathcal{C}$. Then for each $A \subset X$ we denote by $\mathcal{CC}(A)$ the collection of sets in $\mathcal{C}$ such that there is an at most countable sequence of sets $\{E_n\}_{n\in\mathbb{N}} \in \mathcal{CC}(A)$ such that

$$A \subset \bigcup_{n=1}^{\infty} E_n.$$

These are the *countable covers* of A by sets belonging to $\mathcal{C}$.

Definition 185. With $\mathcal{C}$ a collection of sets in X, let $\nu\colon \mathcal{C} \to [0,\infty]$ with $\nu(\emptyset) = 0$. We define the following set function depending on $\mathcal{C}$ and ν

$$\mu^*_{\nu,\mathcal{C}}(A) := \inf_{\mathbb{D}\in\mathcal{CC}(A)} \sum_{D\in\mathbb{D}} \nu(D). \tag{D.4.1}$$

If the infimum is empty, then we define $\mu^*_{\nu,\mathcal{C}}(A) = \infty$.

Theorem 186. *The set function given by (D.4.1), which for simplicity we denote here by μ^*, is an outer measure μ^* on X with*

$$\mu^*(A) \le \nu(A),\ \ A \in \mathcal{C}$$

For any other outer measure $\tilde{\mu}^*$ on X with the above condition we have

$$\tilde{\mu}^*(A) \le \mu^*(A),\ \ A \subset X.$$

So in this sense, μ^* is the unique maximal outer measure on X which satisfies $\mu^*(A) \le \nu(A)$ for all $A \in \mathcal{C}$.

Proof: Let $A \in \mathcal{C}$. Then, A covers itself, so we have by definition

$$\mu^*(A) \le \nu(A).$$

Next, we need to show that this μ^* is an outer measure. We have basically already done this in the Proposition on Outer Measures! Since $\nu \ge 0$, it follows that $\mu^* \ge 0$. Moreover, since $\emptyset$ is a countable cover of itself, we have

$$0 \le \mu^*(\emptyset) \le \nu(\emptyset) = 0.$$

Hence $\mu^(\emptyset) = 0$.

Monotonicity: Assume that $A \subset B$. Then, any countable cover of B is also a countable cover of A. However, there could be covers of A which do not cover B. Hence, the set of countable covers of A contains the set of countable covers of B, so the infimum over covers of A is smaller than the infimum over covers of B, and therefore

$$\mu^*(A) \le \mu^*(B).$$

Countable sub-additivity: Let $\{A_j\}$ be pairwise disjoint. We wish to show that

$$\mu^*(\cup A_j) \le \sum \mu^*(A_j).$$

Note that if for any j we have $\mu^*(A_j) = \infty$, we are immediately done. So, assume this is not the case for any j. Let $\varepsilon > 0$. Then for each j there exists a countable cover $\{D_j^k\}$ such that

$$A_j \subset \cup_k D_j^k, \quad \mu^*(A_j) + \frac{\varepsilon}{2^j} \ge \sum_k \nu(D_j^k).$$

Hence, we also have

$$\cup_j A_j \subset \cup_{j,k} D_j^k,$$

and so

$$\mu^*(\cup A_j) = \inf ... \leq \sum_{j,k} \nu(D_j^k) \leq \sum_j \mu^*(A_j) + \frac{\varepsilon}{2^j} = \varepsilon + \sum_j \mu^*(A_j).$$

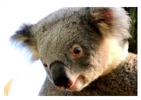
Since this holds for any $\varepsilon > 0$, we obtain the desired inequality.

Another outer measure: Assume that $\tilde{\mu}^*$ is another outer measure defined on X which has $\tilde{\mu}^*(A) \leq \nu(A)$ for all $A \in \mathcal{C}$. If $\mu^*(A) = \infty$ there is nothing to prove. So assume that this is not the case. Let $\varepsilon > 0$. Then there exists a countable cover $\{D_j\}$ which contains A such that

$$\mu^*(A) + \varepsilon \geq \sum_k \nu(D_k) = \sum_k \tilde{\mu}^*(D_k) \geq \tilde{\mu}^*(\cup D_k) \geq \tilde{\mu}^*(A).$$

Above we have used that $\nu = \tilde{\mu}^*$ on the D_k, followed by countable sub-additivity of the outer measure $\tilde{\mu}^*$, followed by monotonicity of the outer measure $\tilde{\mu}^*$. Since this inequality holds for any $\varepsilon > 0$, we get that

$$\mu^*(A) \geq \tilde{\mu}^*(A).$$



Now, we shall specify to the case in which (X, d) is a metric space. For this, we recall that for a non-empty set $A \subset X$ we define its *diameter,*

$$\text{diam}(A) := \sup\{d(x, y) : x \in A, \quad y \in A\}.$$

With this in mind, we can define the countable covers of diameter less than ϵ.

Definition 187. Let $\mathcal{C}$ be as above. For $\epsilon > 0$, define

$$\mathcal{C}_\epsilon := \{A \in \mathcal{C} : \text{diam}(A) < \epsilon\}.$$

Now define the outer measure depending on this cover as a special case of (D.4.1), in particular we set

$$\mu_\epsilon^*(A) := \mu_{\nu, \mathcal{C}_\epsilon}(A).$$

If $\epsilon' < \epsilon$, then all covers which have diameter less than ϵ' also have diameter less than ϵ, so $\mathcal{C}_{\epsilon'} \subset \mathcal{C}_\epsilon$. Consequently, when we take the infimum to obtain μ_ϵ^* and $\mu_{\epsilon'}^*$, there are more elements considered in the infimum for $\mathcal{C}_\epsilon$ (i.e. more covers), so the infimum is smaller, and

$$\mu_\epsilon^*(A) \leq \mu_{\epsilon'}^*(A).$$

The following theorem shows how, starting from any arbitrary set function ν which has $\nu(\emptyset) = 0$, we can construct a "canonical metric outer measure." We shall later see that for a particular choice of ν, we obtain the Hausdorff measure.

Theorem 188 (A canonical metric outer measure)**.** *The limit $\mu_0^*(A) := \lim_{\epsilon \to 0} \mu_\epsilon^*(A), A \subset X$ defines a metric outer measure.*

D.4.4 Exercises

1. Prove that for any interval $I \subset \mathbb{R}^n$, there exists a series $\{B_j\}_{j \geq 1}$ such that
   (a) Each B_j is a ball in I.
   (b) It is $B_j \cap B_k = \emptyset$ for all $j \neq k$.
   (c) We have $\mathcal{L}_n(I \setminus \bigcup B_j) = 0$ (and therefore $\mathcal{L}_n(I) = \mathcal{L}_n(\bigcup B_j)$).
2. Now for a bit of combinatorial fun... Let X be a non-empty set. Let $\{A_j\}_{j=1}^n$ be distinct, non-empty, proper subsets of X. How many elements does $\mathcal{A}$, the smallest algebra which contains $\{A_j\}_{j=1}^n$, have?

D.4.5 Hints

First note that $\mathcal{L}_n(I \setminus \mathring{I}) = 0$. So without loss of generality we can assume that I is open. For $x \in I$, there is $\delta \in \mathbb{Q}, \delta > 0$ such that $B_\delta(x) \subset I$. Also there exists $q \in \mathbb{Q}^n$ such that $|x - q| < \delta \cdot 10^{-6}$. This implies for every y with $|y - q| < (1 - 10^{-6})\delta$,

$$|y - x| \le |y - q| + |x - q| < \delta \implies y \in B_\delta(x) \subset I.$$

So we have

$$B_1 := B_{(1-10^{-6})\delta}(q) \subset I.$$

For $N \ge 1$ and $x \in I$, it is either $x \in \bigcup_{k=1}^N \overline{B_k}$ or not. We are assuming $\{B_k\}^N \subset I$ are disjoint balls with rational radii and rational centers (centers are elements of $\mathbb{Q}^n$). If $x \in \bigcup_{k=1}^N \overline{B_k}$ we consider $x \in I \setminus \bigcup_{k=1}^N \overline{B_k}$. Note that this set is *open*. So, if there exists $x \in I \setminus \bigcup_{k=1}^N \overline{B_k}$, then the same argument shows that there is a new ball,

$$x \in \; B_{N+1} \subset I \setminus \bigcup_{k=1}^N \overline{B_k}$$

with the center and radius of B_{N+1} rational (same argument as above). Then we note further that the set of balls

$$\{B_\delta(q) : \quad \delta \in \mathbb{Q}, \quad \text{and } q \in \mathbb{Q}^n\}$$

is countable. Consequently, we require at most countably many of these balls to ensure that

$$I \subset \bigcup_{k=1}^\infty \overline{B_k} \;\text{ and }\; \mathcal{L}_n(\overline{B_k} \setminus B_k) = 0 \text{ for all } k \;\Rightarrow\; \mathcal{L}_n(\bigcup(\overline{B_k} \setminus B_k)) = 0.$$

So we get

$$\mathcal{L}_n(I) = \mathcal{L}_n(I \cap \bigcup B_k) + \mathcal{L}_n(I \setminus \bigcup B_k) = \mathcal{L}_n(\bigcup B_k) + \mathcal{L}_n(\bigcup \overline{B_k} \setminus B_k) = \mathcal{L}_n(\bigcup B_k).$$

D.5 Canonical metric outer measures and Hausdorff measure

Theorem 189 (A canonical metric outer measure)**.** *The limit $\mu_0^*(A) := \lim_{\epsilon \to 0} \mu_\epsilon^*(A), A \subset X$ defines a metric outer measure.*

Proof: Since μ_ε^* is non-decreasing as $\epsilon \downarrow 0$, the limit exists (since we allow ∞ as a limit value). We have already proven that each μ_ϵ^* is an outer measure.

Exercise 190. *Prove that the outer measure property is preserved under the limit as $\epsilon \to 0$, to show that μ_0^* is indeed an outer measure.*

Metric outer measure: Let $A, B \subset X$ be such that $\operatorname{dist}(A, B) > 0$. Since μ^* is an outer measure, by countable subadditivity,

$$\mu_0^*(A \cup B) \le \mu_0^*(A) + \mu_0^*(B).$$

We would like to prove the reverse inequality. The idea is that since A and B are at a positive distance away from each other, we can take ϵ small enough so that our μ_ϵ^* cover of the union splits into two disjoint covers. (Draw a picture!)

Let us make this precise. Since the distance between A and B is positive, there exists $n_0 \in \mathbb{N}$ such that

$$\operatorname{dist}(A, B) > \frac{1}{n}, \quad \text{for} \quad n > n_0.$$

Let $\delta > 0$ be some arbitrary positive number (this is our fudge factor which we shall later banish to zero). Then, cover the union $A \cup B$ with sets E_k^n such that

$$\mu^*_{\frac{1}{n}}(A \cup B) + \delta \geq \sum_{k=1}^{\infty} \nu(E_k^n)$$

and such that for each k we have $\operatorname{diam}(E_k^n) \leq \frac{1}{n}$. Let us delete any E_k^n which has empty intersection with $A \cup B$, that is we delete any unneeded, extraneous, superfluous covers. Still denote this set by E_k^n for notational simplicity. We then still have

$$\mu^*_{\frac{1}{n}}(A \cup B) + \delta \geq \sum_k \nu(E_k^n).$$

Since the diameter of E_k^n is less than or equal to $\frac{1}{n}$ which is smaller than the distance between A and B, we have that the E_k^n intersect either A or B and not both in the sense that

$$E_k^n \cap A \neq \emptyset \Rightarrow E_k^n \cap B = \emptyset, \; E_k^n \cap B \neq \emptyset \Rightarrow E_k^n \cap A = \emptyset.$$

To see this, draw a picture. If some E_k^n intersected both A and B, then it would have to contain at least one point in A and at least one point in B. The distance between those points is strictly greater than $\frac{1}{n}$. Hence the diameter of such a set would need to exceed $\frac{1}{n}$, which is a contradiction.

So, with this consideration, let

$$E^n(A) := \{E_k^n : E_k^n \cap A \neq \emptyset\}, \quad E^n(B) := \{E_k^n : E_k^n \cap B \neq \emptyset\}.$$

Then, $E^n(A)$ and $E^n(B)$ have no sets in common and together they yield the sequence $(E_k^n)_{k=1}^{\infty}$. Since

$$A \cup B \subset \cup E_k^n \implies A \subset \cup_{E^n(A)}, \quad B \subset \cup_{E^n(B)}.$$

We therefore have

$$\mu^*_{\frac{1}{n}}(A) \leq \sum_{E_k^n \in E^n(A)} \nu(E_k^n), \quad \mu^*_{\frac{1}{n}}(B) \leq \sum_{E_k^n \in E^n(B)} \nu(E_k^n),$$

so the sum

$$\mu^*_{\frac{1}{n}}(A) + \mu^*_{\frac{1}{n}}(B) \leq \sum_{E_k^n \in E^n(A)} \nu(E_k^n) + \sum_{E_k^n \in E^n(B)} \nu(E_k^n).$$

Now, the sum on the right side is just

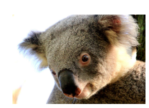
$$\sum_k \nu(E_k^n) \leq \mu^*_{\frac{1}{n}}(A \cup B) + \delta.$$

So, we have proven that

$$\mu^*_{\frac{1}{n}}(A) + \mu^*_{\frac{1}{n}}(B) \leq \mu^*_{\frac{1}{n}}(A \cup B) + \delta.$$

This holds for all $n \geq n_0$. So, letting $n \to \infty$, we obtain

$$\mu^*_0(A) + \mu^*_0(B) \leq \mu^*_0(A \cup B) + \delta.$$

Finally, we let $\delta \downarrow 0$, which completes the proof that μ^*_0 is a metric outer measure. 

Remark 8. Making a special choice of the function ν, we shall obtain the Hausdorff measure, below. However, our preceding results are super general. If you are so inclined, it could be pretty interesting to play around with different functions, ν, satisfying the hypotheses, and thereby obtain different metric outer measures according to the theorem above.... Once you've got a metric outer measure, then you can use our results to obtain its sigma algebra of measurable sets. Moreover, our results prove that this sigma algebra contains the Borel sigma algebra. Our results also prove that this metric outer measure together with its sigma algebra of measurable sets yields a complete measure. So, now you have quite a collection of tools to build all kinds of different measures!

D.5.1 The Hausdorff measure

We shall use the general results from the preceding lecture to obtain the Hausdorff measure.

Definition 191. Let (X, d) be a metric space, $\delta > 0$ and $t \in (0, \infty)$. Then for $S \subset X$, define the set function

$$\mathcal{H}^t_\delta(S) := \inf\left\{\sum_{i=1}^{\infty} (\operatorname{diam} U_i)^t | S \subset \bigcup_{i=1}^{\infty} U_i, \operatorname{diam}(U_i) < \delta\right\}$$

where the infimum is taken over all countable covers of S by sets $U_i \subset X$ with $\operatorname{diam}(U_i) < \delta$.

Remark 9. In the definition if one requires the U_i's to be closed in this definition, the result is the same because

$$\operatorname{diam}(U_i) = \operatorname{diam}(\overline{U_i}).$$

If one requires the U_i's to be open, call the corresponding thing $\tilde{\mathcal{H}}^t_\delta$. Note that the infimum is now taken over fewer covers, since the U_i need to be open. So á priori one has $\tilde{\mathcal{H}}^t_\delta \geq \mathcal{H}^t_\delta$. For S such that

$$\mathcal{H}^t_\delta(S) = \infty,$$

then one also has

$$\tilde{\mathcal{H}}^t_\delta(S) = \infty,$$

so there is nothing to do. Let us assume this is not the case. Fix $\eta > 0$. Let $\{U_j\}$ be a cover which has

$$\mathcal{H}^t_\delta(S) + \eta \geq \sum_j \operatorname{diam}(U_j)^t.$$

Then let

$$B_j = \{x \in X : d(x, U_j) < \epsilon_j 2^{-j-1}\},$$

Choose $\epsilon_j > 0$ so that the diameter of B_j, which is at most $\operatorname{diam}(U_j) + \epsilon 2^{-j}$ is still less than δ. Since the diameter of U_j is strictly less than δ this is always possible. Without loss of generality assume that $\epsilon_j \leq 1$ for all j. Then the B_j are an open cover of S, with diameter less than δ, so we have

$$\tilde{\mathcal{H}}^t_\delta(S) \leq \sum_j (\operatorname{diam}(B_j))^t \leq \sum_j (\operatorname{diam}(U_j) + \epsilon_j 2^{-j})^t.$$

As the $\epsilon_j \to 0$, the right side converges to $\sum_j \operatorname{diam}(U_j)^t$. So, let this happen, to obtain

$$\tilde{\mathcal{H}}^t_\delta(S) \leq \sum_j (\operatorname{diam}(U_j))^t \leq \mathcal{H}^t_\delta(S) + \eta.$$

Since $\eta > 0$ was arbitrary, letting now $\eta \to 0$ we obtain that $\tilde{\mathcal{H}}^t_\delta \leq \mathcal{H}^t_\delta$. So it's still the same. Thus, if it's more convenient to consider (1) closed covers in definition of Hausdorff measure or (2) open covers in definition of Hausdorff measure, DO IT! There is no loss of generality.

Corollary 192 (Hausdorff measure)**.** *The set function $\mathcal{H}^t_\delta$ is an outer measure. Moreover,*

$$\mathcal{H}^t := \lim_{\delta \to 0} \mathcal{H}^t_\delta$$

is a metric outer measure. All Borel sets are $\mathcal{H}^t$ measurable, and these sets form a σ-algebra.

Proof: First, set

$$\nu(U) := \operatorname{diam}(U)^t.$$

Then note that

$$\mathcal{H}^t_\delta(S) = \mu^*_{\nu, \mathcal{C}_\delta}(S)$$

is just a special case of the "canonical outer measure" theorem. By that theorem, we therefore obtain that

$$\mathcal{H}^t(S) := \lim_{\delta\to 0} \mathcal{H}^t_\delta(S)$$

is a metric outer measure. By an earlier theorem (168, all the Borel sets are $\mathcal{H}^t$-measurable. These Borel sets are contained in the σ-algebra of "$\mathcal{H}^t$-measurable sets from Theorem 189. Moreover, by this same theorem, $\mathcal{H}^t$ on this σ-algebra is a complete measure. 

We shall call $\mathcal{H}^t$ the t-dimensional Hausdorff measure. The reason for this is that if $t \in \mathbb{N}$ and A is t-dimensional, then the amount of A contained in a region of diam $= r$" should be proportional to r^t. This is because a ball in t-dimensional space has volume proportional to r^t. What exactly is the volume of a ball in $\mathbb{R}^n$ anyways?

D.5.2 The volume of the unit ball in $\mathbb{R}^n$

Proposition 193. *The volume of the unit ball in $\mathbb{R}^n$ is*

$$w_n = \mathrm{vol}(B_1(0)) = \frac{2\pi^{\frac{n}{2}}}{n \cdot \Gamma(\frac{n}{2})}$$

Proof: Our goal is to compute

$$\int\limits_{S_1(0)} \int\limits_0^1 r^{n-1} \text{ dr } \text{ d } \sigma.$$

For starters, we would like to compute

$$\sigma_n := \int\limits_{S_1(0)} \text{ d } \sigma,$$

that is the surface area of the unit ball. Let us start by computing a famous integral. Define

$$I_n := \int\limits_{\mathbb{R}^n} e^{-\pi|x|^2} dx.$$

Note that $I_n = (I_1)^n$ by the Fubini-Tonelli theorem, since everything converges beautifully. So, in particular,

$$I_n = (I_2)^{2/n}.$$

I_2 is particularly lovely to compute:

$$I_2 = \int\limits_0^{2\pi} \int\limits_0^\infty e^{-\pi r^2} r \text{ dr } \text{ d } \theta = 2\pi \int\limits_0^\infty e^{-s^2} \frac{sds}{\pi} = \int\limits_0^\infty e^{-s^2} 2sds$$

$$= -e^{-s^2}\Big|_0^\infty = 1.$$

We have used the substitution $s = \sqrt{\pi} r$. So we see that $I_k = 1$ for all $k \in \mathbb{N}$. Then, we can apply this to compute σ_n.

$$1 = \int\limits_{\mathbb{R}^n} e^{-\pi|x|^2} \, dx = \int\limits_{S_1(0)} \int\limits_0^\infty e^{-\pi r^2} r^{n-1} \text{ dr } \text{ d } \sigma = \sigma_n \int\limits_0^\infty e^{-\pi r^2} r^{n-1} \text{ dr } .$$

Well, the latter integral we may be able to compute, because it is one-dimensional. Let $s = r^2\pi$. Then ds $= 2r\pi dr$, so

$$r^{n-1}dr = \left(\frac{s}{\pi}\right)^{(n-1)/2} \frac{ds}{2\pi\sqrt{s/\pi}} = \frac{s^{n/2-1}}{2\pi^{n/2}}.$$

So,

$$1 = \frac{\sigma_n}{2\pi^{n/2}} \int_0^\infty e^{-s} s^{n/2-1} ds.$$

This looks familiar... Recall:

$$\Gamma(z) = \int_0^\infty s^{z-1}e^{-s} \text{ ds }, \quad z \in \mathbb{C}, \quad \Re(z) > 1.$$

So,

$$\sigma_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

We compute using integration by parts:

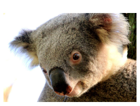
$$\Gamma(s+1) = \int_0^\infty t^s e^{-t} \text{ dt } = \left[-t^s e^{-t}\right] - \int_0^\infty -e^{-t} s t^{s-1} dt = s\Gamma(s).$$

Exercise 194. *Prove that the* Γ *function admits a meromorphic continuation to* $\mathbb{C}$ *which is holomorphic with the exception of simple poles at* $0 \cup -\mathbb{N}$.

Finally, we compute the volume of the ball:

$$\int_{B_1(0)} dx = \text{vol}(B_1(0)) = \int_{S_1(0)} \int_0^1 r^{n-1} \text{ dr } \text{ d } \sigma = \sigma_n \int_0^1 r^{n-1} \text{ dr } = \left[\sigma_n \frac{r^n}{n}\right]_0^1 = \frac{\sigma_n}{n} = w_n$$

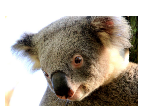
Therefore, we have $w_n = \frac{2\pi^{n/2}}{n \cdot \Gamma(\frac{n}{2})}$, which finishes our proof.



Corollary 195. *$\forall x \in \mathbb{R}^n$ and $r > 0$, the area of $S_r(x)$ is $r^{n-1}\sigma_n$ and* $\text{vol}(B_r(x)) = w_n r^n$.

Proof:

$$\int_{S_r(x)} \text{ d } \sigma = \int_{S_r(0)} \text{ d } \sigma = \int_{S_1(0)} r^{n-1} \text{ d } \sigma = r^{n-1}\sigma_n$$

Analogously for $B_r(x)$. 

Exercise 196. *Compute* $\Gamma(n/2)$ *for* $n \in \mathbb{N}$.

D.6 Hausdorff dimension

If the notion of Hausdorff dimension is to be well-defined, then it should be invariant under isometries. We prove that the Hausdorff measure is indeed invariant under isometries, and therefore the Hausdorff dimension, which we shall define using the Hausdorff measure, will similarly enjoy this invariance. Let $\mathcal{H}^p$ denote p-dimensional Hausdorff measure. We first prove a more general fact. Before proceeding to that proof, there is an exercise which will allow us to be a little sloppy (or for a more positive connotation, allow us to be a little more mellow and groovy).

Exercise 197. *Change the definition of C_ϵ covers to require diameters less than or equal to ϵ. Show that the corresponding μ_0^* remains unchanged. Thus, in the definition of Hausdorff outer measure (and Hausdorff measure), it does not require if our $\mathcal{H}_\delta^p = \mu_{\delta,\nu}^*$ for $\nu(A) = \operatorname{diam}(A)^p$ is for covers with diameter $< \delta$ or $\leq \delta$. Either way one obtains the same outer measure $\mathcal{H}_\delta^p$. Therefore, either way one also obtains the same $\mathcal{H}^p$.*

Proposition 198. *Let (X, d) be a metric space, and f, g be maps from some set Y into X. If $f, g : Y \to X$ satisfy $d(f(y), f(z)) \leq Cd(g(y), g(z))\ \forall y, z \in Y$, then $\mathcal{H}^p(f(A)) \leq C^p \mathcal{H}^p(g(A))$.*

Proof: Let $\varepsilon > 0$, $A \subset Y$. Then $g(A) \subset X$. If $\mathcal{H}^p(g(A)) = \infty$, there is nothing to prove. So, we assume this is not the case. Then, for all $\delta > 0$ small, we can find $\{B_j\}_{j\geq 1} \subset X$ such that

$$g(A) \subset \bigcup_{j=1}^{\infty} B_j, \quad \operatorname{diam}(B_j) < \frac{\delta}{C},$$

and

$$\sum_{j\geq 1} \operatorname{diam}(B_j)^p \leq \mathcal{H}^p(g(A)) + \frac{\varepsilon}{C^p}.$$

Of course, the particular collection B_j does depend on the particular small value of δ, but we shall suppress this dependence for notational convenience.

Let us define

$$\tilde{B}_j := f(g^{-1}(B_j)).$$

We claim that these are going to cover $f(A)$. Let $y \in A$ so that $f(y) \in f(A)$. Then, since $y \in A$, we also have $g(y) \in g(A) \subset \cup B_j$. So, in particular, $g(y) \in B_j$ for some j. Hence

$$y \in g^{-1}(B_j) = \{z \in Y : g(z) \in B_j\}.$$

Therefore $f(y) \in f(g^{-1}(B_j)) = \tilde{B}_j$.

We therefore have

$$f(A) \subset \cup_{j=1}^{\infty} \tilde{B}_j.$$

Now, if $f(y)$ and $f(z)$ are both in $\tilde{B}_j = f(g^{-1}(B_j))$, this means that y and z are both in $g^{-1}(B_j)$, so there exist x and x' in B_j with $g(y) = x \in B_j$ and $g(z) = x' \in B_j$. Then

$$d(f(y), f(z)) \leq Cd(g(y), g(z)) \leq C \operatorname{diam}(B_j) < C\frac{\delta}{C} = \delta.$$

Consequently $\operatorname{diam}(\tilde{B}_j) < \delta$. So,

$$\mathcal{H}_\delta^p(f(A)) \leq \sum_{j\geq 1} \operatorname{diam}(\tilde{B}_j)^p.$$

Moreover, by the same calculation as above, we also see that

$$\operatorname{diam}(\tilde{B}_j) \leq C \operatorname{diam}(B_j) \implies \operatorname{diam}(\tilde{B}_j)^p \leq C^p \operatorname{diam}(B_j)^p.$$

Consequently,

$$\mathcal{H}_\delta^p(f(A)) \leq \sum_{j\geq 1} \operatorname{diam}(\tilde{B}_j)^p \leq C^p \sum_{j\geq 1} (\operatorname{diam}(B_j))^p \leq C^p \mathcal{H}^p(g(A)) + \epsilon.$$

This holds for any $\epsilon > 0$, so we obtain the desired result:

$$\mathcal{H}^p(f(A)) \leq C^p \mathcal{H}^p(g(A)).$$

Corollary 199. *$\mathcal{H}^p$ is invariant under isometries.*

Proof: Let $I : (X, d) \to X$ be an isometry. Let $\mathrm{id} : X \to X$ be the identity map. Then since I is an isometry, we have

$$d(I(x), I(z)) = d(x, z) = d(\mathrm{id}(x), \mathrm{id}(z)), \quad \forall x, z \in X.$$

Hence the hypotheses of the proposition hold true taking $X = Y$, $f = I$, $g = \mathrm{id}$, and $C = 1$. So, we obtain

$$\mathcal{H}^p(I(A)) \leq \mathcal{H}^p(\mathrm{id}(A)) = \mathcal{H}^p(A).$$

On the other hand, we also have

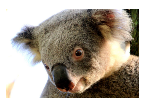
$$d(\mathrm{id}(x), \mathrm{id}(z)) = d(x, z) = d(I(x), I(z)) \leq d(I(x), I(z)).$$

So, we apply the same proposition taking $X = Y$, $f = \mathrm{id}$, $g = I$, and $C = 1$. We therefore obtain

$$\mathcal{H}^p(A) = \mathcal{H}^p(\mathrm{id}(A)) \leq \mathcal{H}^p(I(A)).$$

Thus, the inequality goes in both directions, and we have in fact an equality,

$$\mathcal{H}^p(A) = \mathcal{H}^p(I(A)).$$



Proposition 200 (Hausdorf dimension)**.** *If $\mathcal{H}^p(A) < \infty$, then $\mathcal{H}^q(A) = 0$ $\forall q > p$. If $\mathcal{H}^q(A) > 0$, then $\mathcal{H}^p(A) = \infty$ $\forall p < q$.*

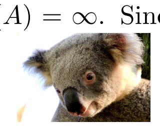
Proof: For the first statement, assume $\mathcal{H}^p(A) < \infty$. Then, for any sufficiently small $\delta > 0$, we can find a cover of A by $\{B_j\}_{j \geq 1}$ with $\mathrm{diam}(B_j) < \delta$, and

$$\mathcal{H}^p_\delta(A) \leq \sum_{j \geq 1} \mathrm{diam}(B_j)^p \leq \mathcal{H}^p(A) + 1.$$

If $q > p$, then

$$\begin{aligned} \mathcal{H}^q_\delta(A) \leq \sum_{j \geq 1} \mathrm{diam}(B_j)^q &= \sum_{j \geq 1} \mathrm{diam}(B_j)^{p+q-p} \leq \sum_{j \geq 1} \mathrm{diam}(B_j)^p \delta^{q-p} \\ &= \delta^{q-p} \sum_{j \geq 1} \mathrm{diam}(B_j)^p \leq \delta^{q-p}(\mathcal{H}^p(A) + 1), \end{aligned}$$

which tends to zero as $\delta \to 0$. Hence we can show that $\mathcal{H}^q_\delta(A)$ tends to zero as $\delta \to 0$, thus it follows that $\mathcal{H}^q(A) = 0$.

The second statement is the contrapositive. To see this let us first fix $q > p$. We shall write $\star$ to denote the statement $\mathcal{H}^p(A) < \infty$, and $\heartsuit$ to denote the statement $\mathcal{H}^q(A) = 0$. We have proven: if $\star$ then $\heartsuit$. The contrapositive says: if not $\heartsuit$ then not $\star$. It is well known from elementary logic that a statement is true if and only if its contrapositive is true. In this case, not $\heartsuit$ says that $\mathcal{H}^q(A) \neq 0$. Since $\mathcal{H}^q(A) \geq 0$, we have $\mathcal{H}^q(A) > 0$. This should imply not $\star$. Not $\star$ is the statement that $\mathcal{H}^p(A) = \infty$. Since the $q > p$ was arbitrary, we have shown that if $\mathcal{H}^q(A) > 0$, then $\mathcal{H}^p(A) = \infty$ for any $p < q$. 

Corollary 201 (Definition of Hausdorff dimension)**.** *Let $A \subset X$, where (X, d) is a metric space. Then the following infimum and supremum are equal*

$$\delta = \inf\{p \geq 0 \mid \mathcal{H}^p(A) = 0\} = \sup\{p \geq 0 \mid \mathcal{H}^p(A) = \infty\}$$

This is how we define the Hausdorff dimension *of A, δ, denoted by* $\dim(A)$*. If for some p we have*

$$\mathcal{H}^p(A) \in (0, \infty)$$

then $p = \dim(A)$.

Proof: Let $\{p_n\}$ be a sequence which converges to the infimum on the left. Then, $\mathcal{H}^{p_n}(A) = 0$ for all n. Let $\{q_n\}$ be a sequence which converges to the supremum on the right. Then, $\mathcal{H}^{q_n}(A) = \infty$ for all n. By the second statement of the preceding proposition, since $\mathcal{H}^{q_n}(A) > 0$, $\mathcal{H}^p(A) = \infty$ for all $p < q$. This shows that $p_n \geq q_m$ for all n and m. Therefore

$$\liminf p_n \geq \limsup q_m.$$

Since in these cases the limits exist, we have

$$\liminf p_n = \lim p_n, \quad \limsup q_m = \lim q_m.$$

This shows that

$$\inf\{p \geq 0 \mid \mathcal{H}^p(A) = 0\} \geq \sup\{p \geq 0 \mid \mathcal{H}^p(A) = \infty\}.$$

For the sake of contradiction, let us assume that this inequality is strict, so that

$$\inf\{p \geq 0 \mid \mathcal{H}^p(A) = 0\} > \sup\{p \geq 0 \mid \mathcal{H}^p(A) = \infty\}.$$

Then, there is some number, x which lies precisely between these two values,

$$\inf\{p \geq 0 \mid \mathcal{H}^p(A) = 0\} > x > \sup\{p \geq 0 \mid \mathcal{H}^p(A) = \infty\}.$$

Since x is less than the infimum, we cannot have $\mathcal{H}^x(A) = 0$, (because then x would be included in the infimum, so the infimum would be $\leq x$ which by assumption it is not). So we must have $\mathcal{H}^x(A) > 0$. By the proposition, it follows that $\mathcal{H}^p(A) = \infty$ for all $p < x$. Hence, the supremum on the right side is taken over a set of p which contains all $p < x$. Therefore, by definition of the supremum, the supremum is greater than or equal to x. This is a contradiction. Hence, we cannot have

$$\inf\{p \geq 0 \mid \mathcal{H}^p(A) = 0\} > \sup\{p \geq 0 \mid \mathcal{H}^p(A) = \infty\},$$

as it leads to a contradiction. Thus, since the infimum is greater than or equal to the supremum, both sides must be equal.

Finally, assume that for some p we have $\mathcal{H}^p(A) \in (0, \infty)$. Then, by the proposition, $\mathcal{H}^q(A) = 0$ for all $q > p$. This shows that all $q > p$ are considered in the infimum, hence the infimum must be less than or equal to p. On the other hand, by the same proposition, $\mathcal{H}^q(A) = \infty$ for all $q < p$. Hence, the supremum is taken over a set which includes all $q < p$, hence the supremum must be greater than or equal to p. So, we get $\inf \leq p \leq \sup$, but since the infimum and supremum are equal, we have an equality all the way across.

This shows that the supremum here is less than or equal to p. Since the supremum and infimum equivalently define $\dim(A)$, we have $\dim(A) \geq p$ and $\dim(A) \leq p$. Hence we have $\dim(A) = p$.

If our notion of dimension is a good one, then it ought to be monotone. We see below that this is the case.

Lemma 202 (Monotonicity of Hausdorff dimension)**.** *If $A \subset B$, then* $\dim(A) \leq \dim(B)$.

Proof:

If $A \subset B$, and $\mathcal{H}^p(B) = 0$, then $\mathcal{H}^p(A) = 0$. This is because $\mathcal{H}^p$ is an outer measure, which we proved, and outer measures are by definition monotone.

Therefore

$$\dim(B) = \inf\{p \geq 0 | \mathcal{H}^p(B) = 0\} \geq \inf\{p \geq 0 | \mathcal{H}^p(A) = 0\} = \dim(A).$$

If our definition of dimension is a good one, then we know what the dimension of $\mathbb{R}^n$ should be... To prove this, we shall prove a general fact about Hausdorff dimension.

Lemma 203. *The dimension of a countable union of sets, E_j,*

$$E = \cup E_j$$

is equal to

$$\dim(E) = \sup\{\dim(E_j\}.$$

Proof: We note that

$$E_j \subset E \forall j \implies \dim(E_j) \leq \dim(E) \quad \forall j,$$

so

$$\sup\{\dim(E_j)\} \leq \dim(E).$$

If the supremum on the left is infinite, there is nothing to prove, because both sides are therefore infinite and equal. Let us assume that it is not infinite. So, let us call this supremum δ. By the definition of $\dim(E_j) \leq \delta$, we have

$$\mathcal{H}^p(E_j) = 0 \quad \forall p > \delta.$$

Consequently, for all $p > \delta$, we have by countable subadditivity of Hausdorff outer measure

$$0 \leq \mathcal{H}^p(E) \leq \sum_j \mathcal{H}^p(E_j) = 0.$$

Thus

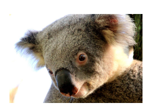
$$\mathcal{H}^p(E) = 0.$$

Since

$$\dim(E) = \inf\{p \geq 0 | \mathcal{H}^p(E) = 0\},$$

and $\mathcal{H}^p(E) = 0$ for all $p > \delta$, we have

$$\dim(E) \leq \delta.$$

Since $\dim(E) \geq \sup\{\dim(E_j)\} = \delta$, we obtain the equality. 

D.6.1 Exercises

1. Compute the Hausdorff measure of the curve $\{(x, \sin(1/x)) : 0 < x < 1\} \subset \mathbb{R}^2$.
2. Compute the Hausdorff measure of the curve $\{(x, \sin(1/x)) : 1/2 < x < 1\} \subset \mathbb{R}^2$.
3. Compute the Hausdorff measure of the unit sphere sitting in $\mathbb{R}^3$.

4. We shall see that a set whose Hausdorff dimension is positive is uncountable. Is the converse true, that is if the Hausdorff dimension of s set is zero, then is that set necessarily countable? Prove or give a counter example.

5. Is it always true that $\mathcal{H}^{\dim(A)}(A) \in (0, \infty)$? Prove or a give a counter example. What if you assume that $\dim(A) \in (0, \infty)$, then is it always true that $\mathcal{H}^{\dim(A)}(A) \in (0, \infty)$?

6. How should one define the Hausdorff dimension of the empty set? Philosophically and mathematically justify your answer.

7. What is the Hausdorff dimension of a product of sets? How should this work? Figure it out and rigorize your answer.

D.7 Properties of Hausdorff dimension

Any set with positive Hausdorff dimension is uncountable!

Corollary 204. *Let $E \subset X$. If $\dim(E) > 0$, then E is uncountable.*

Proof: If E is countable, then $E = \bigcup_j e_j$, where $e_j \in X$ is a point. Therefore, we have proven that

$$0 \leq \dim(E) = \sup \dim(\{e_j\}).$$

Now let $p > 0$. Note that a single point is contained in a ball of radius δ for any $\delta > 0$. Thus by definition

$$\mathcal{H}^p_\delta(e_j) \leq 2^p \delta^p.$$

Letting $\delta \to 0$, we obtain

$$\mathcal{H}^p(e_j) = 0.$$

Therefore the Hausdorff dimension of a point is equal to $\inf\{p : p > 0\} = 0$. By the result we proved, the dimension of E is the supremum over the dimension of e_j, and this is the supremum over zero, hence it is zero. 

Corollary 205 (Hausdorff dimension of $\mathbb{R}^n$)**.** *The Hausdorff dimension of $\mathbb{R}^n$ is n.*

Proof: We can write the euclidian space $\mathbb{R}^n$ as $\mathbb{R}^n = \bigcup_{m \geq 1} B_m$, where B_m are balls of radius m centered at the origin. Here is where we are going to use some teamwork. In the exercises, you have proven that

$$\mathcal{H}^n(B_m) = c_n \mathcal{L}^n(B_m) = c_n m^n w_n,$$

where c_n is a constant that depends only on n, and w_n is the volume of the unit ball in $\mathbb{R}^n$, and $\mathcal{L}^n$ is n-dimensional Lebesgue measure. (i.e. our usual human notion of n-dimensional volume). By a corollary proven today, the Hausdorff dimension of a ball in $\mathbb{R}^n$ is equal to n, since the Hausdorff measure of a ball of radius m is a positive, finite number. Moreover, a ball is an open set, so it is therefore contained in the Borel sigma algebra which is contained in the Hausdorff sigma algebra. So, since $\mathbb{R}^n$ is the union of these balls, and these balls are all Hausdorff measurable sets, the dimension of $\mathbb{R}^n$ is equal to the supremum of the dimensions of the balls. That is the supremum over the constant number n. Hence the supremum is n which gives the dimension of $\mathbb{R}^n$. 

Corollary 206. *For any $A \subset \mathbb{R}^n$, we have $\dim(A) \leq n$.*

Proof: This follows immediately taking $B = \mathbb{R}^n$ in the lemma showing monotonicity of Hausdorff dimension. 

Lemma 207. *Let $E \subset \mathbb{R}^n$ such that* $\dim(E) < n$. *Then* $\overset{\circ}{E} = \emptyset$.

Proof: If $\overset{\circ}{E} \neq \emptyset$, then there $\exists r > 0$ and $x \in E$ such that $B_r(x) \subset E$. $\Rightarrow \dim(E) \geq \dim(B_r(x)) = n$ So we get $n \geq \dim E \geq n \Rightarrow \dim E = n$. 

Remark 10. The Hausdorff Dimension of a subset $E \subset \mathbb{R}^n$ is the same if we consider E as a subset of $\mathbb{R}^m$ for any $m \geq n$ via the canonical embedding, $\mathbb{R}^n \mapsto \mathbb{R}^n \times \{0\}$. In this sense, if we have a set E which naturally lives in k-dimensions, if we view the set E as living in 10 zillion dimensions, the Hausdorff dimension of E remains the same. This is simply because the Hausdorff dimension, which is determined by the Hausdorff (outer) measure is defined in terms of *diameter,* and the diameter of sets does not change if we embed the sets into higher dimensional Euclidean space. That is another reason the Hausdorff dimension is "a good notion of dimension," because it is invariant of the ambient space.

D.8 A comparison of the Hausdorff and Lebesgue measures

Here we determine the relationship between the Hausdorff and Lebesgue measures. First, define

$$\mathcal{H}^0(Z) = \#Z = \text{ the number of elements of the set, } Z.$$

Theorem 208 (Hausdorff and Lebesgue measures)**.** *For all $n \in \mathbb{N}$ we have*

$$\mathcal{H}^n = \frac{2^n}{w_n}\mathcal{L}^n,$$

where w_n is the n-dimensional volume of a unit ball in $\mathbb{R}^n$.

Proof: Let B_r be a ball of radius $r > 0$. Fix $\epsilon > 0$. Then, by the definition of the Lebesgue measure (and outer measure), there exist countably many hypercubes, denoted by R_j such that

$$B_r \subset \cup_j R_j,$$

and

$$\mathcal{L}^n(B_r) + \epsilon \geq \sum_j \mathcal{L}^n(R_j).$$

Next, fix $\delta > 0$.

Claim 209. *There exist countably many open balls $\{B_j^k\}$ which are disjoint, and satisfy*

$$\mathcal{L}^n(R_j \setminus \cup_k B_j^k) = 0.$$

Moreover, given $\delta > 0$, we may choose these balls to have diameters at most equal to δ.

The proof of the claim is an exercise! From the claim it follows that $\mathcal{L}^n(R_j) = \mathcal{L}^n(\cup B_j^k)$. Therefore we have the inequality

$$\mathcal{L}^n(B_r) + \epsilon \geq \sum_j \mathcal{L}^n(R_j) = \sum_{j,k} \mathcal{L}^n(B_j^k) = \frac{w_n}{2^n}\sum_{j,k} \operatorname{diam}(B_j^k)^n.$$

By the absolute continuity of Lebesgue and Hausdorff measures with respect to each other,

$$\mathcal{H}^n(R_j \setminus \cup_k B_j^k) = 0 \implies \mathcal{H}^n_\delta(R_j \setminus \cup_k B_j^k = 0) \quad \forall \delta > 0.$$

This shows that

$$\mathcal{H}^n_\delta(R_j) = \mathcal{H}^n_\delta(\cup_k B_j^k),$$

and

$$\mathcal{H}^n_\delta(\cup R_j) = \mathcal{H}^n_\delta(\cup_{j,k} B_j^k).$$

Then, we also have by monotonicity, since $B_r \subset \cup_j R_j$,

$$\mathcal{H}^n_\delta(B_r) \leq \mathcal{H}^n_\delta(\cup R_j) = \mathcal{H}^n_\delta(\cup_{j,k} B_j^k).$$

Since $\cup B_j^k$ covers itself, by definition of Hausdorff measure

$$\mathcal{H}^n_\delta(\cup_{j,k} B_j^k) \leq \sum_{j,k} \operatorname{diam}(B_j^k)^n.$$

Thus we get

$$\mathcal{H}^n_\delta(B_r) \leq \sum_{j,k} \operatorname{diam}(B_j^k)^n \implies \frac{w_n}{2^n}\mathcal{H}^n_\delta(B_r) \leq \frac{w_n}{2^n}\sum_{j,k} \operatorname{diam}(B_j^k)^n \leq \mathcal{L}^n(B_r) + \epsilon.$$

Letting $\delta \to 0$, we get

$$\frac{w_n}{2^n}\mathcal{H}^n(B_r) \leq \mathcal{L}^n(B_r) + \epsilon,$$

and then letting $\epsilon \to 0$, we get

$$\frac{w_n}{2^n}\mathcal{H}^n(B_r) \leq \mathcal{L}^n(B_r).$$

To complete the proof, we just need to get a lower bound for the Hausdorff measure in terms of the Lebesgue measure.

There is a nifty shortcut one can use here:

Proposition 210 (Isodiametric Inequality)**.** *For any $A \subset \mathbb{R}^n$, one has*

$$\mathcal{L}^n(A) \leq \frac{w_n \operatorname{diam}(A)^n}{2^n}.$$

Exercise 211. *Locate a proof of this fact! Note that when $A = B_r$ the ball of radius r and hence diameter $2r$, the isodiametric inequality states that*

$$\mathcal{L}^n(B_r) = w_n \frac{\operatorname{diam}(B_r)^n}{2^n}.$$

Thus in that case, equality holds. This is a geometric fact which says that the ball of a specified diameter contains the largest volume amongst all sets of the same diameter. A proof can be found in Lawrence Evans & Ronald Gariepy's Measure theory and fine properties of functions, *or even earlier on p. 32 in* Littlewood's miscellany.

So, now let $\epsilon > 0$. Then, there exists a cover of B_r by $\{B_j\}$ of diameter at most δ such that

$$\mathcal{H}^n(B_r) + \epsilon \geq \sum_j \operatorname{diam}(B_j)^n.$$

Then, by the isodiametric inequality,

$$\operatorname{diam}(B_j)^n \geq \mathcal{L}^n(B_j)\frac{2^n}{w_n}.$$

So, we have

$$\mathcal{H}^n(B_r) + \epsilon \geq \frac{2^n}{w_n} \sum \mathcal{L}^n(B_j) \geq \frac{2^n}{w_n} \mathcal{L}^n(B_r),$$

where we have used in the last inequality the countable sub-additivity of the Lebesgue outer measure, since the B_j cover B_r. Since this can be done for any $\epsilon > 0$, we obtain

$$\mathcal{H}^n(B_r) \geq \frac{2^n}{w_n} \mathcal{L}^n(B_r).$$

Combining with the reverse inequality, we get

$$\mathcal{H}^n(B_r) = 2^n w_n \mathcal{L}^n(B_r).$$

Since this holds for all balls which generate the Borel sigma algebra, it holds for all Borel sets. Then, the completion is the same in both cases, so we obtain both the equality of the Hausdorff and Lebesgue sigma algebras, as well as the equality of the Hausdorff and Lebesgue measures. 

Appendix E

Solutions to selected exercises from the text

E.1 Exercises in §2.5

3. (a) $X'' = cX$, $Y' = -cyY$.
   (b) $x^2X'' + xX' = cX$, $Y'' = -(1+c)Y$.
   (c) not possible
   (d) $X'' = c(X' + X)$, $Y' = -cY$.

8. $u_n(x,t) = e^{-(2n+1)^2\pi^2 kt/(4\ell^2)} \sin((2n+1)\pi x/(2\ell))$.

14. (c) $c = 1$ or $c = 0$.

E.2 Exercises in §3.12

1. Follow the proof of the Cauchy & Schwarz inequality for the first statement. For the second statement, it is probably easiest to square both sides and proceed.

4. Check the property concerning scalars, that is what happens if you compute $\langle cf, g\rangle$?

5. $||f_n||^2 = \frac{\pi}{n+1}$.

6. $\frac{\pi}{4}(\coth \pi - \pi \mathrm{csch}^2 \pi)$.

7. $\frac{\pi^6}{960}$

8. Since $\{1, e^{ix}, e^{-ix}\}$ are orthogonal in $\mathcal{L}^2(-\pi,\pi)$ we project onto them the function e^x that is also in $\mathcal{L}^2(-\pi,\pi)$, and obtain the best approximation is

$$c_0 + c_1 e^{ix} + c_{-1}e^{-ix}, \quad c_k = \frac{\int_{-\pi}^{\pi} e^x \overline{e^{ikx}} dx}{2\pi}, \quad k = 0, 1, -1,$$

with the 2π denominator because

$$||e^{ikx}||^2 = \int_{-\pi}^{\pi} |e^{ikx}|^2 dx = 2\pi.$$

11. By definition

$$\langle f,g\rangle = \int_0^1 (1+ix)\overline{2+ix^2}dx = \int_0^1 (1+ix)(2-ix^2)dx = \int_0^1 (2+2ix+x^3-ix^2)dx$$

$$= 2+i+\frac{1}{4}-\frac{i}{3} = 2+\frac{1}{4}+\frac{2}{3}i.$$

Similarly

$$\langle g,f\rangle = \int_0^1 (2+ix^2)(1-ix)dx = \int_0^1 (2-2ix+ix^2+x^3)dx = 2-i+\frac{i}{3}+\frac{1}{4} = 2+\frac{1}{4}-\frac{2}{3}i.$$

Pretty neat, these are complex conjugates of each other just as they should be.

13. $\frac{\pi^4}{90}$.

14. $\frac{a^2(\pi-a)^2}{6}$.

15. See exercise # 8 and proceed similarly.

16. Use the definition and properties of the scalar product.

17. Such an f is in $\mathcal{L}^2$ as is its derivative because they are continuous. Using the 2π periodicity you can show that f is orthogonal to f'.

20. There are many such examples. Here is a very simple way to approach this. Define for $n \geq 1$

$$f_n(x) := \begin{cases} \frac{1}{n} & 0 \leq x \leq n^2 \\ 0 & x > n^2. \end{cases}$$

Then $f_n(x)$ converges to 0 uniformly on $\mathbb{R}$ as $n \to \infty$ since $0 \leq f_n(x) \leq \frac{1}{n}$, but we compute

$$\int_0^\infty |f_n(x)|^2 dx = \frac{1}{n^2}\int_0^{n^2} 1dx = 1 \quad \forall n.$$

E.3 Exercises in §4.9

1. We will use some trigonometric identities to compute

$$a_0 = 0, \quad a_1 = \frac{1}{\pi}\int_0^\pi \cos^2(x)dx = \frac{1}{\pi}\int_0^\pi \frac{\cos(2x)-1}{2}dx = -\frac{1}{2},$$

having used the double angle formula for the cosine

$$\cos(2x) = \cos^2(x) - \sin^2(x) \overset{\cos^2+\sin^2=1}{=} 2\cos^2 x - 1.$$

Since

$$\sin(x)\cos(x) = \frac{1}{2}\sin(2x) \implies b_1 = \frac{1}{\pi}\int_0^\pi \frac{1}{2}\sin(2x)dx = 0.$$

For the other coefficients we use further trigonometric identities that yield for $n > 1$

$$a_n = \frac{1}{\pi}\int_0^\pi \cos(nx)\cos(x)dx = \frac{1}{\pi}\left(\frac{\sin(n-1)x}{2(n-1)} + \frac{\sin(n+1)x}{2(n+1)}\right)\Bigg|_0^\pi = 0.$$

$$\begin{aligned} b_n = \frac{1}{\pi}\int_0^\pi \sin(nx)\cos(x)dx &= \frac{1}{\pi}\left(-\frac{\cos(n-1)x}{2(n-1)} - \frac{\cos(n+1)x}{2(n+1)}\right)\Bigg|_0^\pi \\ &= -\frac{1}{2\pi}\left(\frac{(-1)^{n-1}}{n-1} - \frac{1}{n-1} + \frac{(-1)^{n+1}}{n+1} - \frac{1}{n+1}\right) \\ &= \frac{1}{\pi}\begin{cases}\left(\frac{1}{n-1} + \frac{1}{n+1}\right) & n \text{ is even} \\ 0 & n \text{ is odd.}\end{cases} = \frac{1}{\pi}\begin{cases}\frac{2n}{n^2-1} & n \text{ is even} \\ 0 & n \text{ is odd.}\end{cases}\end{aligned}$$

The Fourier series is therefore

$$-\frac{1}{2}\cos(x) + \frac{1}{\pi}\sum_{n\geq 2,\text{ even}} \frac{2n}{n^2-1}\sin(nx).$$

2. The hyperbolic cosine is an even function, so we will compute the Fourier series with sines and cosines, since the sines drop out.

$$\int_0^\pi \cosh(x)\cos(nx)dx = \mathrm{Re}\int_0^\pi \cosh(x)e^{inx}dx = \mathrm{Re}\int_0^\pi e^{(1+in)x} + e^{(in-1)x}dx$$

$$\begin{aligned} &= \mathrm{Re}\left(\frac{e^{(1+in)\pi}-1}{1+in} + \frac{e^{(in-1)\pi}-1}{in-1}\right) = \mathrm{Re}\left(\frac{(in-1)(e^\pi(-1)^n-1) + (1+in)((-1)^n e^{-\pi}-1)}{-(n^2+1)}\right) \\ &= \frac{-(e^\pi(-1)^n-1) + (-1)^n e^{-\pi} - 1}{-(n^2+1)} = \frac{(-1)^n(e^{-\pi}-e^\pi)}{-(n^2+1)} = (-1)^n\frac{e^\pi - e^{-\pi}}{n^2+1} \\ &= (-1)^n 2\frac{\sinh(\pi)}{n^2+1}\end{aligned}$$

and so dividing by π we get

$$a_n = (-1)^n\frac{2\sinh(\pi)}{\pi(n^2+1)}.$$

We also compute that the

$$\frac{1}{2\pi}\int_{-\pi}^\pi \cosh(x)dx = \frac{1}{\pi}\int_0^\pi \cosh(x)dx = \frac{1}{\pi}\sinh(\pi).$$

The Fourier series is therefore

$$\frac{\sinh(\pi)}{\pi} + \sum_{n\geq 1}(-1)^n\frac{2\sinh(\pi)}{\pi(n^2+1)}\cos(nx).$$

3. There are a few cases to distinguish. If $2a \geq \pi$ then the c_n Fourier coefficients are for $n \neq 0$

$$c_n = \frac{1}{2\pi}\int_{-\pi}^\pi f(x)e^{-inx}dx = \frac{1}{2\pi}\int_{-a}^a e^{-inx}dx = \frac{e^{-ina} - e^{ina}}{-in2\pi}.$$

If $a < \frac{\pi}{2}$, then when we integrate, the integral goes up to either $4a$ if $4a < \pi$ or up to π if $4a > \pi$, so we compute

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx = \frac{1}{2\pi}\int_{-a}^{a} e^{-inx}dx + \frac{1}{2\pi}\int_{2a}^{\min\{4a,\pi\}} -e^{-inx}dx$$
$$= \frac{1}{2\pi}\left(\frac{e^{-an} - e^{an} - e^{-\min\{4a,\pi\}in} + e^{-2an}}{-in}\right), \quad n \neq 0,$$

In all cases

$$c_0 = 0.$$

4. The Fourier series of $\sin^2 x$ is an element of the table, so to get $\cos^2 x$ we can use the fact that $\sin^2 x + \cos^2 x = 1$, so

$$\cos^2 x = 1 - \left(\frac{1}{2} - \frac{1}{2}\cos(2x)\right) = \frac{1}{2} + \frac{1}{2}\cos(2x).$$

7.

$$\frac{8m}{\pi^2}\sum_{n\geq 1}\frac{(-1)^{n+1}}{(2n-1)^2}\sin((2n-1)\pi x/\ell)\cos((2n-1)\pi ct/\ell).$$

8.

$$\frac{2m\ell^2}{\pi^2 a(\ell - a)}\sum_{n\geq 1}\frac{1}{n^2}\sin(n\pi a/\ell)\sin(n\pi x/\ell)\cos(n\pi ct/\ell).$$

12.

$$\frac{\pi}{4} - \frac{2}{\pi}\sum_{n\geq 1}\frac{\cos((4n-2)x)}{(2n-1)^2}.$$

13.

$$\frac{4}{\pi}\sum_{n\geq 1}\frac{(-1)^{n+1}}{2n-1}\cos((2n-1)\pi x/4).$$

14.

$$(e-1)\sum_{n\in\mathbb{Z}}\frac{e^{2\pi inx}}{1-2\pi in}.$$

15.

$$f(t) = \frac{2}{3} - \frac{3}{\pi^2}\sum_{n=1}^{\infty}\frac{1-\cos(2n\pi/3)}{n^2}\cos(2n\pi t/3).$$
$$y(t) = \frac{2}{9} - \frac{3}{\pi^2}\sum_{n\geq 1}\frac{1-\cos(2n\pi/3)}{n^2(3-\frac{4}{9}n^2\pi^2)}\cos(2n\pi t/3).$$

16.

$$50 - \frac{400}{\pi^2}\sum_{n\geq 1}\frac{1}{(2n-1)^2}\exp\left(\frac{-(2n-1)^2\pi^2(1.1)t}{10^4}\right)\cos((2n-1)\pi x/100).$$

18.

$$\frac{4\ell}{c\pi^2}\sum_{n\geq 1}\frac{(-1)^{n+1}}{(2n-1)^2}\sin((2n-1)\pi\delta/\ell)\sin((2n-1)\pi x/\ell)\sin((2n-1)\pi ct/\ell).$$

19.

$$10 - 7e^{-.576x}\cos(2\pi t - 0.576x) - 5e^{-11x}\cos(730\pi t - 11x).$$

At depth 21 cm for daily; at 4.6 meters for annual.

20. a. has 12, b has infinitely many, c has zero.

E.4 Exercises in §5.6

1. (a) $u(x,t) = \sum_{n\geq 1} b_n \exp(-(n-1/2)^2\pi^2 kt/(\ell^2)) \sin((n-1/2)\pi x/\ell)$
   (b) $b_n = \frac{200}{\pi(2n-1)}$.
2. $u(x,t) = (e^{-2t} - e^{-kt})\sin(x)/(k-2)$ if $k \neq 2$, $u(x,t) = te^{-2t}\sin(x)$ if $k = 2$.
3. $f(x) = \frac{4}{3} + \frac{2}{\pi^2}\sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{(-1)^n(1+in\pi)}{n^2} e^{in\pi x}$ and
$$y = -\frac{4}{3} + \frac{2}{\pi^2}\sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{(-1)^{n-1}(1+in\pi)}{n^2(2n^2\pi^2+in\pi+1)} e^{in\pi x}.$$
4. $u(x,t) = \frac{4R}{\pi}\sum_{n \text{ odd}} \frac{e^{-ct}-e^{-n^2\pi^2 kt/\ell^2}}{n((n^2\pi^2k/\ell^2 - c)} \sin(n\pi x/\ell)$ if $c \neq n^2\pi^2k/\ell^2$ for any odd n. If $c = N^2\pi^2k/\ell^2$ for some odd N, then the coefficient of $\sin(N\pi x/\ell)$ is $N^{-1}te^{-N^2\pi^2kt/\ell^2}$.
5. $u(x,t) = \sum_{n\geq 1}(a_n\cos\lambda_n t + b_n \sin\lambda_n t)\sin(n\pi x/\ell)$, with $\lambda_n^2 = \frac{n^2\pi^2c^2}{\ell^2} + a^2$.
6. $u(x,t) = Rt$.
7. $u(x,y) = \frac{1}{6}(y^3 - y) + \frac{2}{\pi^3}\sum_{n\geq 1}\frac{(-1)^{n-1}}{n^3\sinh(2n\pi)}(\sinh(n\pi x) + 7\sinh(n\pi(2-x))\sin(n\pi y)$.
8. $u(x,y) = -\frac{8}{\pi^3}\sin(\pi x)\frac{\sin(\sqrt{20-\pi^2}y)}{\sin(\sqrt{20-\pi^2})} - \frac{8}{\pi^3}\sum_{k\geq 1}\frac{\sin((2k+1)\pi x)}{(2k+1)^3}\frac{\sinh(\sqrt{(2k+1)^2\pi^2-20y})}{\sinh(\sqrt{(2k+1)^2\pi^2-20})}$.
9. $u(x,y) = \frac{8\ell^2}{\pi^3}\sum_{n\geq 1}\frac{1}{(2n-1)^3\sinh((2n-1)\pi)}\sin((2n-1)\pi x/\ell)\sinh((2n-1)\pi y/\ell)$.
10. $\frac{8}{\pi}\sum_{n\geq 1}\frac{n}{4n^2-1}\sin(2nx)$ and the sum is $\frac{\pi^2}{64}$.
11. $u(x,t) = C + \sum_{n\geq 1}\left(b_n - \frac{4C}{\pi(2n-1)}\right)\exp(-(n-1/2)^2\pi^2kt/\ell^2)\sin((n-1/2)\pi x/\ell)$, with
$$b_n := \frac{2}{\ell}\int_0^\ell f(x)\sin((2n-1)\pi x/(2\ell))dx.$$
15. $u(x,t) = R(1-e^{-ct})/c$.
16. $u(x,t) = \frac{b\ell}{2} - \frac{4b\ell}{\pi^2}\sum_{n\geq 1}\frac{1}{(2n-1)^2}\cos((2n-1)\pi x/\ell)\cos((2n-1)\pi ct/\ell)$.
17. $u(x,t) = \frac{gx(x-\ell)}{2c^2} + \frac{4\ell^2 g}{\pi^3c^2}\sum_{n\geq 1}\frac{1}{(2n-1)^3}\sin((2n-1)\pi x/\ell)\cos((2n-1)\pi ct/\ell)$.
18. $u(x,t) = \sum_{n\geq 1} e^{-kt}(a_n\cos\lambda_n t + b_n\sin\lambda_n t)\sin(n\pi x/\ell)$ with $\lambda_n^2 = \frac{n^2\pi^2c^2}{\ell^2} - k^2$. If $k > \pi c/\ell$, some of the λ_n will be imaginary.
20. $u(x,y) = C + \sum_{n\geq 1}\frac{\ell a_n}{n\pi\sinh n\pi}\cos(n\pi x/\ell)\cosh(n\pi y/\ell)$ where the a_n are the coefficients of the Fourier cosine series for f with $a_0 = 0$.
21. The steady state temperature is
$$u(r,\theta) = \sum_{n\in\mathbb{Z}} c_n \frac{r^n + r_0^{2n}r^{-n}}{1+r_0^{2n}}e^{in\theta}, \quad c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-in\theta}d\theta,$$
and the solution for $f(\theta) = 1 + 2\sin\theta$ is
$$u(r,\theta) = 1 + 2\frac{r^2+r_0^2}{r(1+r_0^2)}\sin\theta.$$

22. Let $g(r) = \sum_{n\geq 1} c_n \sin(n\pi \log r/(\log r_0))$ and $h(r) = \sum_{n\geq 1} d_n \sin(n\pi \log r/(\log r_0))$, then

$$u(r,\theta) = \sum_{n\geq 1} (a_n e^{n\pi\theta/\log r_0} + b_n e^{-n\pi\theta/\log r_0}) \sin(n\pi \log r/\log r_0),$$

with $a_n + b_n = c_n$, and $d_n = a_n e^{n\pi\beta/\log r_0} + b_n e^{-n\pi\beta/\log r_0}$.

E.5 Exercises in §6.7

1. The eigenvalues $\lambda_k = \nu_k^2$ where ν_k are the positive solutions to the equation

$$\tan \nu a = \frac{3\nu}{2\nu^2 - 1},$$

with eigenfunctions $\nu_k \cos(\nu_k x) + \sin(\nu_k x)$.

2. $(2/\ell)^{1/2} \cos((n-1/2)\pi x/\ell)$, $n \in \mathbb{N}$.

3. $\lambda_1 = 4 - \beta_1^2$ where β_1 is the positive root of the equation

$$\tanh \beta = \frac{\beta}{2}, \quad u_1(x) = e^{-2x} \sinh(\beta_1 x),$$

and for $n \geq 2$, $\lambda_n = 4 + \beta_n^2$, where β_n are the positive roots to the equation

$$\tan \beta = \frac{\beta}{2}, \quad u_n(x) = e^{-2x} \sin(\beta_n x).$$

4. The eigenvalues are $\lambda_n = \nu_n^2$ where the ν_n are the positive solutions of $\tan \nu = -\nu$, and the eigenfunctions are $\sin(\nu_n x)$.

5. For each t let $\sum a_n(t)\phi_n(x)$ be the expansion of $f(x,t)$ in terms of the eigenfunctions $\phi_n(x)$. Then

$$u(x,t) = \sum_{n\geq 1} \phi_n(x) e^{-\lambda_n t} \int_0^t e^{\lambda_n s} a_n(s) ds.$$

6. $c\overline{c'} = r(a)/r(b)$.

8. $(2/\ell)^{1/2} \sin((n-1/2)\pi x/\ell)$, $n \in \mathbb{N}$.

9. If $\beta < 0$ the eigenvalues are the numbers $\lambda_n = \nu_n^2$ where the ν_n are the positive solutions of $\tan(\nu\ell) = -\beta/\nu$, and the eigenfunctions are $\cos(\nu_n x)$. If $\beta = 0$, the eigenvalues are $\lambda_n = (n\pi/\ell)^2$ for $n \geq 0$, with $n \in \mathbb{Z}$, and the eigenfunctions are $\cos(\nu_n x)$. If $\beta > 0$ the eigenvalues are the squares of the positive solutions ν_n to $\tan(\nu\ell) = -\beta/\nu$ together with the square of the unique positive solution μ_0 of $\tanh(\mu\ell) = \beta/\mu$. The eigenfunctions are

$$\phi_n(x) = \cos(\nu_n x), \quad \phi_0(x) = \cosh(\mu_0 x).$$

10. If $\alpha > 0$, the eigenvalues are $\lambda_n = \nu_n^2$ where the ν_n are the positive solutions of $\tan(\nu\ell) = \alpha/\nu$, and the eigenfunctions are $\cos(\nu_n(\ell - x))$. If $\alpha \leq 0$ one proceeds similarly as in the preceding solution.

12. $\lambda_n = \left(\frac{n\pi}{\log b}\right)^2$, $\phi_n(x) = \sin(n\pi \log x/\log b)$.

13. $\lambda_n = \frac{1}{4} + \left(\frac{n\pi}{\log b}\right)^2$, $\phi_n(x) = \sin(n\pi \log x/\log b)$.

14. $u(x,t)=\sum_{n\geq 1}\frac{200\sin^2(\nu_n\ell)}{(\ell b+\sin^2(\nu_n\ell))\cos(\nu_n\ell)}e^{-\nu_n^2kt}\cos(\nu_n x)$ where ν_n are the positive solutions of $\tan(\nu\ell)=b/\nu$.

15. b.

$$u(x,t)=\frac{\ell e^{-ht}}{2}-\frac{4\ell}{\pi^2}\sum_{n\geq 1}\frac{e^{-(h+(2n-1)^2\pi^2kt/\ell^2)t}}{(2n-1)^2}\cos((2n-1)\pi x/\ell).$$

c.

$$u(x,t)=100\frac{\sinh(\alpha x)}{\sinh(\alpha\ell)}+200\pi\sum_{n\geq 1}\frac{(-1)^n ne^{-n^2\pi^2kt/\ell^2}}{\alpha^2\ell^2+n^2\pi^2}\sin(n\pi x/\ell),\quad \alpha=\sqrt{\frac{h}{k}}.$$

E.6 Exercises in §7.6

1. $u(x,t)=e^{-4\pi^2t}\sin(2\pi x)+2\pi\sin(1)\sum_{n\geq 1}(-1)^{n-1}\frac{n}{n^2\pi^2-1}\left(\frac{t}{n^2\pi^2}-\frac{1}{n^4\pi^4}(1-e^{-n^2\pi^2t})\right)\sin(n\pi x)$.

2.

$$u(r,t)=B+\frac{2\rho(A-B)}{\rho+\delta}\sum_{k\geq 1}\frac{J_1(\lambda_k\rho/(\rho+\delta))}{\lambda_k J_1(\lambda_k)^2}J_0\left(\frac{\lambda_k r}{\rho+\delta}\right)e^{-\lambda_k^2t/(\rho+\delta)^2}$$

with $J_0(\lambda_k)=0$.

3.

$$u(r,\theta)=\sum_{n\geq 0}\frac{(2(2(n+1/2)\frac{\pi}{\ln 2}(-1)^n-1)}{(n+1/2)\pi((n+1/2)^2(\pi/\ln 2)^2+1)}\frac{\sinh(n+1/2)\frac{\pi\theta}{\ln 2}}{\sinh(n+1/2)\frac{\pi^2}{4\ln 2}}\sin(n+1/2)\frac{\pi\ln 4}{\ln 2}.$$

4. $u(x,t)=-\frac{x^2}{2}+\frac{16}{\pi^3}\sum_{k\geq 0}\frac{\cos((2k+1)\pi t)}{(2k+1)^3}\sin((2k+1)\pi x/2)$.

5. $u(r,\theta)=-\frac{1}{2}r^2\cos(2\theta)+r\cos(\theta)+\frac{1}{2}$ in polar coordinates, or $\frac{1}{2}(y^2-x^2)+x+\frac{1}{2}$ in cartesian coordinates.

6. Steady state temperature is $a(1-r^2)/4$.

$$u(r,t)=\frac{a}{4}(1-r^2)-2a\sum_{k\geq 1}\frac{J_0(\lambda_k r)}{\lambda_k^3J_1(\lambda_k)}e^{-\lambda_k^2t}.$$

8.

$$u(r,t)=2A\sum_{k\geq 1}\frac{\lambda_k J_1(\lambda_k)}{(\lambda_k^2+b^2c^2)J_0(\lambda_k)^2}J_0\left(\frac{\lambda_k r}{b}\right)e^{-\lambda_k^2t/b^2},$$

with $\lambda_k J_0'(\lambda_k)+bcJ_0(\lambda_k)=0$.

9.

$$u(r,\theta,z)=\sum_{n\geq 0}\sum_{k\geq 1}(a_{kn}\cos(n\theta)+b_{kn}\sin(n\theta))J_n\left(\frac{\lambda_{kn}r}{b}\right)\sinh(\lambda_{kn}z/b),$$

where $J_n(\lambda_{kn})=0$ and a_{kn} and b_{kn} are obtained by expanding $g(r,\theta)$ in terms of the basis $J_n(\lambda_{kn}r)\sin(n\theta)$.

10.

$$u(r,z)=a_0z+\sum_{k\geq 1}a_kJ_0(\lambda_k r)\sinh(\lambda_k z)$$

with $J_0'(\lambda_k)=0$, and $a_0=2\int_0^1 rf(r)dr$, and

$$a_k=\frac{2}{J_0(\lambda_k)^2\sinh(\lambda_k)}\int_0^1 rf(r)J_0(\lambda_k r)dr,\quad k>0.$$

E.7 Exercises in §8.7

1. $c_0 = 2\pi(1-e^{-1/2})$ and $c_n = \pi(e^{1/2}-e^{-1/2})e^{-|n|}$ for $n \neq 0$.
2. $\frac{\sqrt{\pi}}{16}(3+6x-\frac{1}{2}x^2)$.
3. $3(2x^2-1)$.
4. $\frac{8}{81}(x^2+12)$.
5. $x^3-\frac{3}{2}x^2+\frac{3}{5}x-\frac{1}{20}$.
6. $H'_{2k}(0)=0$, $H'_{2k+1}(0)=2(-1)^k\frac{(2k+1)!}{k!}$.
9. The solution is $a_0[1+\sum_{k\geq 1}c_kx^{2k}]+a_1[x+\sum_{k\geq 1}d_kx^{2k+1}]$ where a_0 and a_1 are arbitrary and

$$c_k = (-1)^k\frac{\left(\prod_{j=0}^{1-k}(\nu-2j)\right)\left(\prod_{\ell=1}^{k}(\nu+2\ell-1)\right)}{(2k)!}$$

and

$$d_k = (-1)^k\frac{\left(\prod_{j=1}^{k}(\nu-2j+1)\right)\left(\prod_{\ell=1}^{k}(\nu+2\ell)\right)}{(2k+1)!}.$$

13. $y=a_0y_0+a_1y_1$ with

$$y_0 = 1+\sum_{k\geq 1}\frac{\prod_{j=1}^{k}(4j-4-\lambda)}{(2k)!}x^{2k},$$

$$y_1 = x+\sum_{k\geq 1}\frac{\prod_{j=1}^{k}(4j-2-\lambda)}{(2k+1)!}x^{2k+1}.$$

E.8 Exercises in §9.10

1. $-\frac{i\pi}{2a}\xi e^{-a|\xi|}$ and $\frac{\pi}{2a^3}(1+a|\xi|)e^{-a|\xi|}$.
3. $-\frac{4iab\xi}{(\xi^2+2b\xi+a^2+b^2)(\xi^2-2b\xi+a^2+b^2)}$, $-\frac{\pi}{2}e^{i\xi}e^{-2|\xi|}(1+2i\,\mathrm{sgn}\,\xi)$.
5. $\frac{\pi}{8}(e^2+1)$.
7. i and $\frac{i}{2\sqrt{2}}$
8. $\pi(1-e^{-1})$
9. $\frac{3}{4}e^{-2|t|}-te^{-2t}\theta(t)$, with $\theta(t)=0$ if $t<0$ and $\theta(t)=1$ if $t>0$.
10. $f(0)=1$, $\int_{\mathbb{R}}f(x)dx=1$.
11. $u(x)=g*e^{-|x|}=e^{-x}\int_{-\infty}^{x}e^{y}g(y)dy+e^{x}\int_{x}^{\infty}e^{-y}g(y)dy$.
12. $\frac{\xi}{\xi^2+k^2}$ and $\frac{k}{\xi^2+k^2}$.
13. $u(x,t)=\frac{1}{\sqrt{4\pi kt}}\int_0^\infty f(y)(e^{-(x-y)^2/(4kt)}-e^{-(x+y)^2/(4kt)})dy$.
14. $u(x,y)=\frac{2}{\pi}\int_0^\infty\frac{\cos(x\xi)\cosh(y\xi)}{(1+\xi^2)\cosh\xi}d\xi$.

15. $\hat{f}(\xi) = \frac{2\pi\sqrt{\xi}}{1+\xi}$ for $0 < \xi < 2$, and $\frac{\pi}{2}$ and $2\pi(\ln 3 - 2/3)$.

16. $u(t) = \frac{1}{2}e^{-t/\sqrt{e}}\theta(t) + \frac{1}{2}e^{\sqrt{3}t}(1-\theta(t))$.

17. $u(x,t) = \frac{4kt+1-2x^2}{(4kt+1)^{5/2}}e^{-\frac{x^2}{4kt+1}}$.

18. $u(x,y) = \frac{2}{\pi}\int_0^\infty \frac{\xi\sin(x\xi)\sinh(y\xi)}{(1+\xi^2)\sinh\xi}d\xi$.

19. Fourier transforming in the x variable gives the Fourier transform of the solution is

$$\hat{u}(\xi,y) = \frac{\sinh(\xi y)}{\sinh(\xi a)}\hat{f}(\xi).$$

Apply Plancharel's theorem to obtain the inequality. The solution

$$u(x,y) = \frac{1}{2a}\int_{\mathbb{R}} \frac{\sin(\pi y/a)}{\cosh((\pi(x-t)/a) + \cos(\pi y/a)}f(t)dt.$$

21. $\frac{e^{-ia\xi}-e^{-ib\xi}}{i\xi}$.

23. (a) in $\mathcal{L}^1$, (b) in $\mathcal{L}^2$, (c) neither, (d) both.

24. $f*f(x) = x+2$ if $-2 \le x \le 0$, or $2-x$ if $0 \le x \le 2$ and 0 otherwise. $f*f*f(x) = \frac{1}{2}(x+3)^2$ if $-3 \le x \le -1$, $3-x^2$ if $-1 \le x \le 1$, and $\frac{1}{2}(3-x)^2$ if $1 \le x \le 3$, and 0 otherwise. $f_\varepsilon * g(x) = 2x^3 + (2\varepsilon^2 - 2)x$.

25. $f*g(x) = \sqrt{\pi/3}e^{-2x^2/3}$.

39.

$$\frac{1}{2b}\left[\int_{\mathbb{R}} \frac{\sin(\pi y/b)f(t)}{\cosh(\pi(x-t)/b) - \cos(\pi y/b)}dt + \int_{\mathbb{R}} \frac{\sin(\pi y/b)g(t)}{\cosh(\pi(x-t)/b) + \cos(\pi y/b)}dt\right].$$

40.

$$u(r,z) = \frac{1}{\pi}\int_{\mathbb{R}} \frac{e^{iz\xi}\sin(\ell\xi)}{\xi}\frac{I_0(r\xi)}{I_0(a\xi)}d\xi$$

where I_0 is the modified Bessel function of order zero, equivalently, $I_0(x) = J_0(ix)$.

42. $f(t) = \sum_{n\in\mathbb{Z}} f(n\pi/\Omega)e^{i(b-a)(\Omega t - n\pi)/(2\Omega)}\frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}$.

43. $\frac{\xi}{\xi^2+k^2}$ and $\frac{k}{\xi^2+k^2}$.

46. $u(x,y) = \int_0^\infty K(x,y,z)f(z)dz + \int_0^\infty K(y,x,z)g(z)dz$ with

$$K(s,t,z) = \frac{1}{2\pi}\left[\frac{t}{(s-z)^2+t^2} - \frac{t}{(s+z)^2+t^2}\right].$$

47.

$$u(x,y) = \frac{2}{\pi}\int_0^\infty \frac{\sin(c\xi)\cos(x\xi)\cosh(y\sqrt{\xi^2+h})}{\xi\cosh(\sqrt{\xi^2+h})}d\xi.$$

E.9 Exercises in §10.6

1. $u(x,t) = \operatorname{erfc}(x/\sqrt{4kt})$ for $0 < t < 1$, $u(x,t) = \operatorname{erfc}(x/\sqrt{4kt}) - \operatorname{erfc}(x/\sqrt{4k(t-1)})$ for $t \geq 1$.

3. $u(x,t) = \frac{x}{\sqrt{4\pi k}} \int_0^t f(t-s) e^{-as} s^{-3/2} e^{-x^2/(4ks)} ds$.

5. $u(t) = \frac{ae^{-t}}{\sqrt{a^2-1}} \sin(t\sqrt{a^2-1})$ if $a > 1$, $u(t) = te^{-t}$ if $a = 1$, and
$$u(t) = \frac{ae^{-t}}{\sqrt{a^2-1}} \sinh(t\sqrt{1-a^2}), \quad a < 1.$$

6. $u = \pm\sqrt{30} t^2 e^{-3t}$.

7. $\frac{a}{z^2-a^2}$, $\frac{z}{z^2-a^2}$.

8. $\frac{z^2+2}{z(z^2+4)}$.

9. $\frac{\sqrt{\pi}}{2a} e^{z^2/(4a^2)} \operatorname{erfc}(z/2a)$.

10. $e^{az} E_1(az)$, $a^{-1} - ze^{az} E_1(az)$.

11. $z^{-2} - \frac{e^{-z}(2z+1)}{z^2(z+1)}$.

12. $\sqrt{\pi/z} e^{a/(4z)} \operatorname{erfc}(\sqrt{a/(4z)})$.

13. $\frac{\Gamma(n+\alpha+1)(z-1)^n}{n! z^{n+\alpha+1}}$.

14. (a) $1 + e^{-2t}$, and (b) $1 - e^{-2t} - 2te^{-2t}$, and (c) $1 - \cos(t)$.

15. $(a-b)^{-1}(e^{at} - e^{bt})$ if $b \neq a$ and te^{at} if $b = a$.

16. (a) $\sin(t)$, and (b) $\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} t^{a+b-1}$, and (c) $\frac{2}{3}\sin t - \frac{1}{3}\sin(2t)$.

18. (a) $z^{-2} - \frac{\pi}{z\sinh(\pi z)}$ and (b) $\frac{1}{z}\tanh(\pi z/2)$ and (c) $\frac{1}{z^2}\tanh(\pi z/2)$.

19. (a) $e^{-2t} - 2te^{-2t} - 2e^{-4t}$
    (b) $3 + e^t\cos(2t) + 4e^t\sin(2t)$
    (c) $\frac{\pi}{2} - \frac{4}{\pi}\sum_{n\geq 1} \frac{\cos((2n-1)t)}{(2n-1)^2}$
    (d) $1 + \frac{4}{\pi}\sum_{n\geq 1} \frac{(-1)^n}{2n-1} e^{-(2n-1)^2\pi^2 t/4}$
    (e) $\frac{a}{2\sqrt{\pi t^3}} e^{-a^2/(4t)}$, and $\operatorname{erfc}(a/\sqrt{4t})$.
    (f) $E_1(t)$.

21. (a) $u(t) = \frac{\omega\sin(2t) - 2\sin(\omega t)}{2(\omega^2-4)}$ if $\omega \neq 2$, and if $\omega = 2$ it is $\frac{1}{8}(\sin(2t) - 2t\cos(2t))$.
    (b) $u(t) = \int_0^t f(t-s) se^{-2s} ds + c_0 e^{-2t} + (c_1 + 2c_0)te^{-2t}$.
    (c) $u(t) = e^{-t}\sin t + \frac{1}{2}\Theta(t-\pi)(1 + e^{\pi-t}\cos t + e^{\pi-t}\sin t) - \frac{1}{2}\Theta(t-2\pi)(1 - e^{2\pi-t}\cos t - e^{2\pi-t}\sin t)$.
    (d) $u(t) = \int_0^t f(t-s)(\cosh s - 1) ds + 1 - \sinh t$.
    (e) $u(t) = -t + \frac{1}{2}\sinh t + \frac{1}{2}\sin t$.

22. $u(t) = ct^2 - t$, and generally $u(t) = c_1 t^2 + c_t t^{-1} - t$.

23. $u(t) = c_1 e^t + c_2(t+1)$.

27. $u(t) = 2a^{-2}(\cosh(at) - 1)$.

28. $u(t) = f(t) + \frac{1}{2}\int_0^t f(t-s)(\sinh s - \sin s)ds$.

30. First question is $r(t) = cN_0$, second question $r(t) = 2cN_0 - (c-1)N_0e^{-t}$.

Bibliography

[1] C. Aldana and J. Rowlett, *A polyakov formula for sectors*, J. Geom. Anal. **28** (2018), no. 2, 1773–1839.

[2] W. O. Amrein, A. M. Hinz, and D. B. Pearson (eds.), *Sturm-Liouville theory*, Birkhäuser Verlag, Basel, 2005. Past and present, Including papers from the International Colloquium held at the University of Geneva, Geneva, September 15–19, 2003. MR2132131

[3] J. Dutka, *On the early history of bessel functions*, Archive for History of Exact Sciences **49** (1995), no. 2, 105–134.

[4] G. B. Folland, *Fourier analysis and its applications*, 4th ed., Pure and Applied Undergraduate Texts 4, American Mathematical Society, 2009.

[5] J. B. J. Fourier, *Mémoire sur la propagation de la chaleur dans les corps solides*, Oeuvres de fourier: Publiées par les soins de gaston darboux, 2013, pp. 213–222.

[6] C. Gordon, D. Webb, and S. Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*, Invent. Math. **110** (1992), no. 1, 1–22. MR1181812

[7] C. Gordon, D. L. Webb, and S. Wolpert, *One cannot hear the shape of a drum*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 1, 134–138. MR1136137

[8] S. Hawking, *Zeta function regularization of path integrals in curved spacetime*, Commun. math. Phys. **55** (1977), 133–148.

[9] M. Kac, *Can one hear the shape of a drum?*, Amer. Math. Monthly **73** (1966), no. 4, part II, 1–23. MR201237

[10] L. D. Landau and E. M. Lifshitz, *Quantum mechanics (non-relativistic theory)*, Pergamon Press, 1977.

[11] N. N. Lebedev, *Special functions and their applications*, Dover Publications, Inc., New York, 1972. Revised edition, translated from the Russian and edited by Richard A. Silverman, Unabridged and corrected republication. MR0350075

[12] Z. Lu and J. Rowlett, *The sound of symmetry*, Amer. Math. Monthly **122** (2015), no. 9, 815–835. MR3418203

[13] Z. Lu and J. M. Rowlett, *One can hear the corners of a drum*, Bull. Lond. Math. Soc. **48** (2016), no. 1, 85–93. MR3455751

[14] S. Luo, *The sturm-liouville problem of two-body system*, Journal of Physics Communications **4** (2020jun), no. 6, 061001.

[15] A. Messiah, *Quantum mechanics*, Vol. 1, John Wiley, 1961.

[16] M. Nursultanov, J. Rowlett, and D. Sher, *How to hear the corners of a drum*, 2017 MATRIX annals, 2019, pp. 243–278. MR3931070

[17] F. Oberhettinger and L. Badii, *Tables of laplace transforms* (F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, eds.), Springer-Verlag, 1973.

[18] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (eds.), *NIST handbook of mathematical functions*, U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge, 2010. With 1 CD-ROM (Windows, Macintosh and UNIX). MR2723248

[19] Å k. Pleijel, *A study of certain Green's functions with applications in the theory of vibrating membranes*, Ark. Mat. **2** (1954), 553–569. MR61257

[20] B. Riemann, *über die anzahl der primzahlen unter einer gegebenen grösse*, Monatsberichte der Berliner Akademie (1859).

[21] J. Rowlett, *Blast into math! a fun and rigorous introduction to pure mathematics*, Bookboon, 2013.

[22] J. von Neumann, *Mathematical foundations of quantum mechanics*, Princeton University Press, 1955.

[23] G. Watson, *Treatise on the theory of Bessel functions*, 2nd ed., Cambridge Univ. Press, 1966.

[24] H. Weyl, *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)*, Math. Ann. **71** (1912), no. 4, 441–479. MR1511670