Linear Analysis - Math 309

Christopher McMurdie

Spring 2013

Contents

Chapter 0

Review

1 Complex Numbers

Consider the following quadratic equation:

$$x^2 + 1 = 0.$$

In high school, we learn that solving equations like this one corresponds to finding the values of x for which the function $f(x) = x^2 + 1$ crosses the x-axis. Since the graph of this function $f(x)$ is a parabola sitting "above" the x-axis, it never crosses the x-axis and so there are no solutions to this equation. Indeed, $x^2 = -1$ has no *real* solutions since $x^2 \geq 0$ for every real number x.

So we extend our set of numbers to include solutions to this equation! This may seem like a radical idea, but it is not the first time you have done this. Probably the easiest numbers to understand are positive integers. As children, we first learn the concept by counting objects. Next, we extend our numbers to include the negative integers – conceptually, we are allowing solutions to equations like $x + 5 = 2$; $x = -3$, the "opposite of 3." The next extension typically includes fractions; we do this to solve, for example, equations like $2x - 4 = 1$ ($x = \frac{5}{2}$). Finally, we extend our numbers to include solutions to certain algebraic equations like $x^2 + x - 1 = 0$, $x = \frac{-1 \pm \sqrt{5}}{2}$. The picture of the *real number line* is probably what convinces us intuitively that this "makes sense," since we can point to solutions on the line and see how different numbers have different "sizes" and negatives are on the "other side" of zero, etc.

Actually, the real number line includes even more numbers, like π, which are *not* solutions to any algebraic equation. If you want to know how to rigorously define the set of real numbers, then take an introductory class in mathematical analysis.

1.1 Rectangular Coordinates

We introduce a picture for our new set of numbers. Instead of a line, we have the plane. Each point $(a, b) \in \mathbb{R}^2$ of the plane represents the *complex number* $a + bi$. We add (subtract) and multiply these numbers exactly like we do real numbers, with the caveat that $i^2 = -1$. For example:

$$\begin{aligned}(1-i)+(2+3i) &= 1+2-i+3i \\ &= 3+2i. \\ (1-i)\cdot(2+3i) &= 1\cdot 2+1\cdot 3i-i\cdot 2-i\cdot 3i \\ &= 2+i-3i^2 = 2+i+3 \\ &= 5+i.\end{aligned}$$

The entire set of complex numbers is denoted by $\mathbb{C}$, just like the set of real numbers is denoted by $\mathbb{R}$. Notice that $\mathbb{C}$ contains the set of real numbers, since $a+0\cdot i = a \in \mathbb{R}$. More generally, if we have a complex number $a+bi$ in $\mathbb{C}$, then we call a the *real part* of $a+bi$. We call b the *imaginary part* of $a+bi$. Graphically, these parts correspond to the x and y-components of the point $(a,b) \in \mathbb{R}^2$.

If $a+bi$ is a complex number, and $a=0$, we say that the number is *purely imaginary*. For example, the number $5i$ is purely imaginary. Also, notice that the imaginary part of a complex number $a+bi$ is always a *real* number, $b \in \mathbb{R}$. Do not be confused by this; for example, the imaginary part of $5i$ is 5, and 5 is a real number.

What about division of complex numbers? How do we compute $\frac{1}{1-i}$, for example? Remember that division (fractions) is really just a way to find solutions to equations. In this case, the relevant equation is

$$1 = (a+bi)\cdot(1-i).$$

So we may solve the equation

$$1 = a+bi-ai+b = (a+b)+(b-a)i$$

for a and b. At first this seems to be one equation with two unknowns; actually, there are two equations here! The real parts of both sides must be equal, as well as their imaginary parts. So we get

$$\begin{cases} 1 = a+b \\ 0 = b-a. \end{cases}$$

This system of equations has the solution $b = 1/2 = a$, so that $\frac{1}{1-i} = \frac{1}{2}+\frac{1}{2}i$.

There is a trick for easily computing fractions. First, consider a complex number $a+bi$. Define the *complex conjugate* of $a+bi$ to be $a-bi$. This is the original complex number with the imaginary part multiplied by -1. Here's the point:

$$(a+bi)\cdot(a-bi) = a^2+ab\,i-ab\,i-b^2i^2 = a^2+b^2 \text{ is in } \mathbb{R}.$$

In words, a complex number multiplied by its conjugate is a non-negative real number. Moreover, if the original complex number is not zero, then the product is positive. Assume that $a+bi \neq 0$, so we may invert both sides of the previous equation. We get

$$\frac{1}{(a+bi)(a-bi)} = \frac{1}{a^2+b^2},$$

so that
$$\frac{1}{a+bi} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i.$$
This shows that every non-zero complex number has a multiplicative inverse in $\mathbb{C}$.

1.1 Remarks.

- Thinking of the "picture" of $\mathbb{C}$ as points $(a,b) \in \mathbb{R}^2$, then $a^2+b^2 = |(a,b)|^2$ is just the square of the length of the line segment from $(0,0)$ to (a,b). We call the length $|(a,b)|$ of that line segment the *modulus* of $a+bi$.

- The procedure above is exactly "rationalizing the denominator." For example,
$$\begin{aligned}\frac{1}{1+\sqrt{2}} &= \frac{1}{1+\sqrt{2}} \cdot \frac{1-\sqrt{2}}{1-\sqrt{2}} \\ &= \frac{1-\sqrt{2}}{-1} = -1+\sqrt{2}.\end{aligned}$$
We just have -1 under the radical instead of 2, since $i = \sqrt{-1}$.

1.2 Example. Simplify (compute) $\frac{2-3i}{1+i}$.
We first compute $\frac{1}{1+i} = \frac{1}{1+i} \cdot \frac{1-i}{1-i} = \frac{1-i}{2} = \frac{1}{2} - \frac{1}{2}i$. Then we multiply
$$\begin{aligned}\frac{2-3i}{1+i} &= \frac{1}{1+i}(2-3i) = (\frac{1}{2} - \frac{1}{2}i) \cdot (2-3i) \\ &= 1 - \frac{3}{2}i - i + \frac{3}{2}i^2 = (1-\frac{3}{2}) + (-\frac{3}{2}-1)i \\ &= -\frac{1}{2} - \frac{5}{2}i.\end{aligned}$$

Exercises. Show the following by direct computation.

i) $(\sqrt{2}-i) - i(1-\sqrt{2}i) = -2i$

ii) $\frac{1+2i}{3-4i} + \frac{2-i}{5i} = -\frac{2}{5}$

iii) $\frac{5}{(1-i)(2-i)(3-i)} = \frac{1}{2}i$

iv) $-1 = \frac{4}{(1-i)^4}$.

v) Show that $(a+bi)^2 = 1+i$ admits only the two solutions $\pm\left[\sqrt{\sqrt{2}+1/2} + i\sqrt{\sqrt{2}-1/2}\right]$.

1.2 Exp of a complex number.

Perhaps e^x is the most important function in calculus; so we should figure out how to compute e^{a+bi}. This means we should figure out how to write it as a complex number $a' + b'i$. We expect that it satisfies the usual rule of summing in the exponent, so that

$$e^{a+bi} = e^a(e^{bi}).$$

Since e^a is a real number, it is sufficient to compute e^{bi}. We do this using Taylor series. Indeed, we know that

$$e^x = 1 + x + x^2/2 + x^3/3! + \cdots$$

so that

$$\begin{aligned} e^{bi} &= 1 + bi + (bi)^2/2 + (bi)^3/3! + (bi)^4/4! + (bi)^5/5! + \cdots \\ &= 1 + bi - b^2/2! - b^3/3!\, i + b^4/4! + b^5/5!\, i + \cdots \\ &= (1 - b^2/2! + b^4/4! + \cdots) + (b - b^3/3! + b^5/5! + \cdots)i \\ &= \cos(b) + \sin(b)i. \end{aligned}$$

The previous formula is called *Euler's formula*; since it is very important, I will reproduce it here:

$$\boxed{e^{bi} = \cos(b) + \sin(b)\, i.}$$

It follows that $e^{a+bi} = e^a \cos(b) + e^a \sin(b)i$.

1.3 Example. Compute 2^{5i}.

We first rewrite 2^{5i} as $(e^{\ln(2)})^{5i} = e^{5\ln(2)i}$. Now we may use Euler's formula with $b = 5\ln(2)$. We get:

$$2^{5i} = e^{5\ln(2)i} = \cos(5\ln(2)) + \sin(5\ln(2))\, i.$$

1.4 Remark. This trick works perfectly well when the base, in this case 2, is a positive real number. We will stick to that case in this class.

1.3 Polar coordinates

Recall the polar representation of the plane $\mathbb{R}^2$. For each point $p = (a, b) \in \mathbb{R}^2$, we can associate $r = \sqrt{a^2 + b^2}$. This r is the length of the line segment $\vec{p}$ from 0 to (a, b) (or what we called the *modulus* of the complex number $a + bi$). If one draws a circle centered at the origin of radius r, then the point (a, b) will lie somewhere on that circle. But where? We measure counter-clockwise the angle that $\vec{p}$ makes with the positive x-axis, and denote that angle by θ. Basic trigonometry tells us that

$$\begin{aligned} a &= r\cos(\theta) \\ b &= r\sin(\theta). \end{aligned}$$

So the point $(a, b) \in \mathbb{R}^2$ can be specified either by its x and y-coordinates (which is how we have just specified it), or it can be specified by its modulus r and its *argument* (angle) θ. If we have a, b and want to know (r, θ), we may use the identities

$$\begin{aligned} r &= \sqrt{a^2 + b^2} \\ \theta &= \tan^{-1}(b/a). \end{aligned}$$

We apply this polar representation of $\mathbb{R}^2$ to our set of complex numbers $\mathbb{C}$. Since every complex number may be written as $a + bi$, we have

$$\begin{aligned} a + bi &= r\cos(\theta) + r\sin(\theta)i = r(\cos(\theta) + i\sin(\theta)) \\ &= re^{\theta\, i} \end{aligned}$$

using Euler's formula. Polar coordinates allows one to multiply complex numbers much more quickly than when using rectangular coordinates. For example:

$$2e^{\frac{\pi}{4}i} \cdot \frac{2}{3}e^{\frac{\pi}{4}i} = \frac{4}{3}e^{\frac{\pi}{2}i}.$$

Compare this with the same computation in rectangular coordiantes:

$$(\sqrt{2} + \sqrt{2}i) \cdot (\frac{\sqrt{2}}{3} + \frac{\sqrt{2}}{3}i) = 2/3 + 2/3i + 2/3i - 2/3 = \frac{4}{3}i.$$

More importantly, now we can conceptually understand what complex multiplication means in terms of our "picture" of $\mathbb{C}$ as the plane. Consider two complex numbers, $r_1e^{\theta_1\, i}$ and $r_2e^{\theta_2\, i}$. Then the product is

$$r_1e^{\theta_1\, i} \cdot r_2e^{\theta_2\, i} = r_1r_2e^{(\theta_1+\theta_2)i}.$$

So the argument of the product, that is the angle that the line segment from 0 to the product makes with the positive x-axis, is exactly the sum of the arguments of each factor. Also, the modulus of the product is just the product of the moduli of the factors. If we think a little bit geometrically, multiplication by a fixed complex number has the effect of scaling (by the modulus of that number) and rotating (by the argument of that number).

When we do not want to specify the coordinate system for the complex number, we will usually write z instead of $a + bi$ or $re^{i\theta}$. In all cases, z will mean a complex number, so an element in the set $\mathbb{C}$.

Exercise.

i) Find all complex numbers such that $z^3 = -1$.

ii) Find all complex z such that $z^2 = \sqrt{2}e^{i\pi/4}$. Compare with the last of the previous exercises.

Solutions.

i) $z = e^{(\pi/3)\, i}, e^{\pi\, i}, e^{(5\pi/3)\, i}$

ii) $z = 2^{1/4}e^{(\pi/8)\, i}, 2^{1/4}e^{(9\pi/8)\, i}$

Chapter 1

Linear Systems of ODEs

1 Introduction

Why should we bother solving systems of linear differential equations? Just as in the case of second-order linear equations, the motivation comes from physics.

Recall the simple spring-mass systems that we investigated in Math 307. We model the position of the mass m as a function of time. Let $x(t)$ denote the displacement from the equilibrium position of the mass at time t. There are two forces acting on the mass. From Hooke's law, we have that the spring exerts a force

$$F_{sp} = -kx,$$

where k is the spring's so-called "spring-constant." There is also a friction force:

$$F_{fr} = -\gamma x',$$

where γ is the coefficient of friction.

From Newton's second law, we have that the sum of the forces on m is equal to m times the acceleration of the mass. Thus we have $mx'' = -kx - \gamma x'$. If we rewrite this slightly, we obtain the usual equation

$$mx'' + \gamma x' + kx = 0.$$

Example. Suppose that $m = 1$ and $\gamma = k = 2$. Then the previous equation specializes to

$$x'' + 2x' + 2x = 0.$$

This is a linear ordinary differential equation with constant coefficients. In math 307, we learned to solve this by guessing that $x(t) = e^{rt}$. We plug-in this guess, and see which values of r "work" - i.e., for which values of r the guess $x = e^{rt}$ satisfies the equation $x''+2x'+2x = 0$.

Plug-in the guess, and reduce to the equation

$$r^2 + 2r + 2 = 0.$$

This is called the characteristic equation. The only solutions are the complex numbers

$$r = -1 \pm i.$$

Using Euler's formula, we have that

$$x(t) = e^{(-1+i)t} = e^{-t}\cos(t) + ie^{-t}\sin(t)$$

is a solution. Since the coefficients in the differential equation are real-numbers, we know that both the real part and the imaginary part of this solution are themselves solutions! Since $\{e^{-t}\cos(t), e^{-t}\sin(t)\}$ is independent and the equation is second-order, a general solution is given by

$$x(t) = c_1e^{-t}\cos(t) + c_2e^{-t}\sin(t).$$ □

This is all a review of the methods we see in Math 307. Now we are going to consider a more complicated physical system consisting of two masses and three springs. We want to model $x_1(t)$, and $x_2(t)$, the displacement from the equilibrium positions of the masses m_1 and m_2, respectively. The three springs have spring constants k_1, k_2, and k_3. For simplicity, we will assume in this model that there is no friction force.

We first analyze the forces acting on m_1:

$$\begin{aligned} F_{sp}^1 &= -k_1x_1 \\ F_{sp}^2 &= k_2(x_2 - x_1). \end{aligned}$$

To see the second equality, notice that the natural length of the second spring is 0 when $x_1 = x_2$. Then the difference $x_2 - x_1$ is the displacement of the second spring from its equilibrium. If $x_2 > x_1$, then the spring is stretched and so it exerts a force of $k_2 \cdot (x_2 - x_1)$ directed to the right (and so is positive).

Similarly, the forces acting on m_2 are given by:

$$\begin{aligned} F_{sp}^3 &= -k_3x_2 \\ F_{sp}^2 &= -k_2(x_2 - x_1). \end{aligned}$$

From Newton's law we get the two equations

$$\begin{aligned} m_1x_1'' &= -k_1x_1 + k_2(x_2 - x_1) = -(k_1 + k_2)x_1 + k_2x_2 \\ m_2x_2'' &= -k_3x_2 - k_2(x_2 - x_1) = k_2x_1 - (k_3 + k_2)x_2. \end{aligned}$$

This is an example of a homogeneous system of second-order linear differential equations. Examples like this, and those in your homework, motivate us to learn how to solve linear systems.

In fact, we will develop methods for solving *every* first-order linear system of ordinary differential equations. It is important to see how we can reduce higher-order systems, like what we just derived, to first-order systems.

Example. The simplest system of equations are just single equations. Consider

$$x_1'' + 2x_1' + 2x_1 = 0.$$

To rewrite this as a first-order equation, we introduce the variable $x_2 := x_1'$. By differentiating, we have $x_2' := x_1''$, so that

$$0 = x_1'' + 2x_1' + 2x_1 = x_2' + 2x_2 + 2x_1.$$

If we solve this equation for x_2', we obtain

$$x_2' = -2x_1 - 2x_2.$$

This is *not* yet a first-order linear system, since x_2' depends on the variable x_1, and yet we have not explicitly stated how x_1' depends on x_1 and x_2.

From the definition of x_2, we know that $x_1' = x_2$. So the complete first-order system is

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= -2x_1 - 2x_2. \end{aligned}$$

□

It is important to understand that the previous example establishes a one-to-one correspondence between solutions $x_1(t)$ to the second-order equation $x_1'' + 2x_1' + 2x_1 = 0$, and solutions $(x_1(t), x_2(t))$ to the system of equations

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= -2x_1 - 2x_2. \end{aligned}$$

In detail, a solution x_1 to the second-order equation determines the solution $(x_1(t), x_2(t)) = (x_1, x_1')$ to the first-order system; conversely, any solution $(x_1(t), x_2(t))$ to the first order system determines the solution $x_1(t)$ to the second-order equation.

The two problems have corresponding solutions and should be thought of as "the same."

Example. Consider the system of second-order differential equations

$$\begin{aligned} x_1'' &= -2x_1 + x_2 \\ x_2'' &= x_1 - 2x_2. \end{aligned}$$

We will rewrite this as a system of first-order equations. As before, introduce new variables

$$x_3 := x_1' \qquad x_4 := x_2'.$$

Since $x_3' = x_1''$, and $x_4' = x_2''$, we get the partial system

$$\begin{aligned} x_3 &= -2x_1 + x_2 \\ x_4 &= x_1 - 2x_2. \end{aligned}$$

To complete it, we need to know x_1' and x_2' in terms of the variables $\{x_1, x_2, x_3, x_4\}$. From the definitions of x_3 and x_4, we know that $x_1' = x_3$ and $x_2' = x_4$. So the complete system is

$$\begin{aligned} x_1' &= x_3 \\ x_2' &= x_4 \\ x_3' &= -2x_1 + x_2 \\ x_4' &= x_1 - 2x_2. \end{aligned}$$

□

Example. Consider a higher-order equation, say

$$x_1^{(4)} - x_1 = 0.$$

To rewrite this as a linear system of first-order equations, we introduce the variables

$$x_2 := x_1' \quad x_3 := x_2' = x_1'' \quad x_4 := x_3' = x_2'' = x_1'''.$$

Then $x_1^{(4)} = x_4'$, so $0 = x_1^{(4)} - x_1 = x_4' - x_1$, i.e.

$$x_4' = x_1.$$

To complete this system, we need to know first how x_1', x_2', and x_3' depend on the variables $x_1, \dots, x_4$. This is given by the definitions above, so we obtain the completed system:

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= x_3 \\ x_3' &= x_4 \\ x_4' &= x_1. \end{aligned}$$

□

1.1 Definition. A *system of n first-order linear equations* has the form

$$\begin{aligned} x_1' &= p_{11}\,x_1 + p_{12}\,x_2 + \cdots + p_{1n}\,x_n + g_1 \\ x_2' &= p_{21}\,x_1 + p_{22}\,x_2 + \cdots + p_{2n}\,x_n + g_2 \\ &\vdots \\ x_n' &= p_{n1}\,x_1 + p_{n2}\,x_2 + \cdots + p_{nn}\,x_n + g_n \end{aligned}$$

where all of the p_{ij}'s and g_i's are continuous functions of time in some interval (a, b). A *solution* consists of an n-tuple of differentiable functions $(x_1, \dots, x_n)$ defined in (a, b) which satisfy the equation.

The system is called *homogeneous* when every $g_i = 0$. The system has *constant coefficients* when every function p_{ij} and g_i are constant functions.

All of the examples that we have seen so far are linear homogeneous equations with constant coefficients. In this case, it is often better to use the following matrix notation to represent the system:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}' = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix};$$

or, defining $\underline{x} = (x_1, \ldots, x_n)^T$,

$$\underline{x}' = P\underline{x},$$

where $P = (p_{ij})$ is the $n \times n$ matrix with ijth entry p_{ij}.

Example. The previous system

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= x_3 \\ x_3' &= x_4 \\ x_4' &= x_1. \end{aligned}$$

can be re-written as

$$\underline{x}' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \underline{x}.$$

□

The simplest example of a non-homogeneous linear system of n-first order differential equations is the case where $n = 1$. A system of this form is just a linear equation

$$x_1' = p_{11}x_1 + g_1.$$

We solve these using the technique of integrating factors from Math 307:

$$\text{define } \mu := e^{-\int p_{11}(t)dt}; \qquad \text{then } [\mu x_1]' = \mu \cdot g.$$

Integrate the second equation to obtain

$$\mu x_1 = \int \mu g \, dt + C \quad \Rightarrow \quad x_1 = \frac{1}{\mu} \int \mu g \, dt + \frac{C}{\mu}.$$

1.2 Theorem (Every linear system can be solved)**.** *Suppose that $p_{ij} : (a, b) \longrightarrow \mathbb{R}$ is continuous for all i, j. If t_0 is in (a, b), for every initial condition $\underline{x}(t_0) = \underline{x}_0$ there exists a unique solution $\underline{\phi} = (\phi_1, \ldots, \phi_n)$ defined on (a, b).*

2 Review of Matrices

A homogeneous linear system of first-order equations with constant coefficients has the form

$$\underline{x}' = P\underline{x}$$

for some P in $M_n(\mathbb{C})$. We will see that solving these systems requires quite a familiarity with matrices, eigenvalues, eigenvectors and diagonalization, etc. So with that as motivation, we begin by reviewing matrices.

An $m \times n$ matrix M is a rectangular array of $m \cdot n$ elements arranged in m rows and n columns. The element in the ith row and jth column is traditionally labeled m_{ij} or $M[i, j]$. If the elements are contained in some set where we have an addition and multiplication defined, then we can define the addition and multiplication of matrices. Indeed, if M and N are $m \times n$ matrices,

$$(M + N)[i, j] := M[i, j] + N[i, j];$$

this just means that we sum the matrices by summing corresponding entries. If M is an $m \times n$ matrix, and if N is an $n \times k$ matrix, then we define the product by

$$(M \cdot N)[i, j] := \sum_{k=1}^{n} M[i, k] \cdot M[k, j];$$

this product makes sense since the number of columns of M is n, which is the same as the number of rows of N. The product $M \cdot N$ is a matrix of size $m \times k$.

Notice that $M \cdot N \neq N \cdot M$ in general; indeed, often only one of these expressions is even defined. Even for square $n \times n$ matrices where both expressions make sense, $M \cdot N \neq N \cdot M$ in general.

Left multiplication by a square matrix

Suppose that $A = (a_{ij})$ is a square matrix. Then A consists of n row-vectors $r_1, \ldots, r_n$, where each r_i is in $\mathbb{R}^n$. Suppose that c is an $n \times 1$ matrix, i.e. a column vector of length n. Then

$$A \cdot c = \begin{pmatrix} r_1 \cdot c^T \\ r_2 \cdot c^T \\ \vdots \\ r_n \cdot c^T \end{pmatrix},$$

where $r_1 \cdot c^T$ is the usual dot product from calculus. Notice that the resulting matrix is again a column vector of length n. Thus, left multiplication by a square matrix is a special kind of function (a *linear* function) sending column vectors to column vectors.

Let B be another square $n \times n$ matrix, consisting of n columns $c_1, \ldots, c_n$. Then we have

$$A \cdot B = \begin{pmatrix} A \cdot c_1 & A \cdot c_2 & \cdots & A \cdot c_n \end{pmatrix};$$

or in words, the jth column of the product $A \cdot B$ is the product of A with the jth column of B. This is a very useful way to understand matrix multiplication. We will make frequent use of this in the future.

2.1 Definition. Suppose $A = (a_{ij})$ is an $m \times n$ matrix with entries a_{ij} in $\mathbb{C}$.

- The *transpose* of A is the $n \times m$ matrix
$$A^T := (a_{ji});$$
- The *conjugate* of A is the matrix
$$\overline{A} := (\overline{a}_{ij});$$
- The *adjoint* of A is the matrix
$$A^* := (\overline{a}_{ji}) = (\overline{A})^T = \overline{(A^T)}.$$

2.2 Remarks.

- The transpose exchanges the ith row with the ith column.
- $(A^T)^T = A$, $\overline{(\overline{A})} = A$, and $(A^*)^* = A$;
- $(AB)^T = B^T A^T$, and $(AB)^* = B^* A^*$. Remember to change the order!

Example.

1.
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

2.
$$\overline{\begin{pmatrix} 1-i & 2+2i \\ 3+i & 4-5i \end{pmatrix}} = \begin{pmatrix} 1+i & 2-2i \\ 3-i & 4+5i \end{pmatrix}$$ □

2.3 Definition. Suppose that $A = (a_{ij})$ is a *square* $n \times n$ matrix with entries a_{ij} in $\mathbb{C}$.

- The *ijth minor of A*, denoted A_{ij} is the determinant of the square $(n-1) \times (n-1)$ submatrix obtained by deleting the ith row and jth column of A.
- The *adjugate* of A is the matrix
$$adj(A) = ((-1)^{i+j} A_{ji}).$$

EXAMPLE.

1. If $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, then $A_{12} = |3| = 3$.

2. If $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, then

$$A_{13} = \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 4 \cdot 8 - 5 \cdot 7 = -3.$$

3. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

$$adj(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Notice that

$$A \cdot adj(A) = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = \det(A) \cdot I_2.$$ □

2.4 Proposition. *If A is a (square) $n \times n$ matrix, then*

$$A \cdot adj(A) = adj(A) \cdot A = \det(A) \cdot I_n.$$

From this proposition, we see that A^{-1} exists and

$$A^{-1} = \frac{1}{\det(A)} adj(A) \qquad \text{when } \det(A) \neq 0.$$

2.5 Definition. If $\det(A) = 0$, then we say that A is *singular*. Otherwise, it is nonsingular.

2.6 Remark. If A and B are non-singular $n \times n$ matrices, then the product $A \cdot B$ is non-singular, and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Remember to change the order!

Using the adjugate, we have an important formula for the inverse of a non-singular matrix, at least for theoretical purposes. For instance, when $n = 2$ we get the closed formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

for the inverse of a non-singular matrix. However, the actual computation of A^{-1} for $n > 2$ should not usually be done this way. It is generally much(!) faster to perform elementary row operations on the augmented matrix:

$$[\, A \,|\, I_n \,] \xrightarrow{\text{rref}} [\, I_n \,|\, A^{-1} \,].$$

Once the matrix is in reduced row echelon form, the inverse matrix can be read off.

Example. Suppose that $A = \begin{pmatrix} 1 & -1 & 4 \\ 0 & 2 & 0 \\ 0 & 3 & 1 \end{pmatrix}$, so $\det(A) = 1 \cdot \left| \begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix} \right| = 2$. Since $\det(A) \neq 0$, A is non-singular and so A^{-1} exists. We want to perform elementary row operations on the augmented matrix $[\,A \,|\, I_3\,]$ to put it in reduced row-echelon form:

$$\begin{pmatrix} 1 & -1 & 4 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 3 & 1 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{2:\,\frac{r_2}{2}} \begin{pmatrix} 1 & -1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 3 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[3:\,r_3-3r_2]{1:\,r_1+r_2} \begin{pmatrix} 1 & 0 & 4 & 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & -\frac{3}{2} & 1 \end{pmatrix}$$

$$\xrightarrow{1:\,r_1-4r_3} \begin{pmatrix} 1 & 0 & 0 & 1 & \frac{13}{2} & -4 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & -\frac{3}{2} & 1 \end{pmatrix}.$$

Then

$$A^{-1} = \begin{pmatrix} 1 & \frac{13}{2} & -4 \\ 0 & \frac{1}{2} & 0 \\ 0 & -\frac{3}{2} & 1 \end{pmatrix}.$$

□

2.7 Theorem. *Suppose that $A = (a_{ij})$ is a square $n \times n$ matrix with entries in $\mathbb{C}$. Then the following are equivalent:*

i) A is not singular, i.e. $\det(A) \neq 0$;

ii) The matrix A^{-1} exists;

iii) For every column vector b in $\mathbb{C}^n$, there exists a column vector x in $\mathbb{C}^n$ such that

$$A \cdot x = b;$$

iv) For every two column vectors x_1, x_2 in $\mathbb{C}^n$,

$$\text{if} \quad Ax_1 = Ax_2 \quad \text{then} \quad x_1 = x_2.$$

v) The columns of A are linearly independent in $\mathbb{C}^n$;

vi) The rows of A are linearly independent in $\mathbb{C}^n$.

Proof. Exercise from Math 308. □

2.8 Definition. Let V be a vector space defined over $\mathbb{C}$. An *inner product* on V is a function $(-,-) : V \times V \longrightarrow \mathbb{C}$ that satisfies the following properties

i) Conjugate symmetry:

$$\overline{(x, y)} = (y, x).$$

ii) Complex linearity in the first argument: for every scalar c in $\mathbb{C}$ and vectors x_1, x_2, y in $\mathbb{C}^n$,

$$(c \cdot x_1 + x_2, y) := c(x_1, y) + (x_2, y).$$

iii) Positive definiteness: for every vector x in $\mathbb{C}^n$,

$$(x, x) \text{ is in } \mathbb{R}; \quad (x, x) \geq 0; \qquad \text{if } (x, x) = 0 \quad \text{then } x = 0.$$

2.9 Definition. The *standard inner product* on $\mathbb{C}^n$ is defined as follows: for x, y in $\mathbb{C}^n$:

$$(x, y) := y^* x = \overline{(y^T)} x = \sum_{i=1}^{n} \overline{y_i} x_i.$$

2.10 Proposition. *The standard inner product is an inner product, i.e. it satisfies the three parts of the definition.*

Proof. i):

$$\begin{aligned} \overline{(x, y)} &= \overline{y^* x} = (\overline{y^* x})^T \\ &= (y^* x)^* = x^* y^{**} \\ &= x^* y = (y, x). \end{aligned}$$

ii):

$$\begin{aligned} (cx_1 + x_2, y) &= y^*(cx_1 + x_2) \\ &= cy^* x_1 + y^* x_2 = c(x_1, y) + (x_2, y). \end{aligned}$$

iii):

$$(x, x) = \sum_{i=1}^{n} \overline{x_i} x_i = \sum_{i=1}^{n} |x_i|^2 \geq 0.$$

The only way for this sum to be 0 is if every $x_i = 0$, in which case $x = 0$. □

2.1 Matrix Functions

We can consider matrices whose entries are themselves functions. For example, define

$$A(t) := \begin{pmatrix} e^t & 2e^{-t} & e^{2t} \\ 2e^t & e^{-t} & -e^{2t} \\ -e^t & 3^{-t} & 2e^{2t} \end{pmatrix} \qquad B(t) := \begin{pmatrix} 2e^t & e^{-t} & 3e^{2t} \\ -e^t & 2e^{-t} & e^{2t} \\ 3e^t & -e^{-t} & -e^{2t} \end{pmatrix}.$$

For these particular matrices, we could also write, e.g.

$$A(t) = \left(\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} e^t, \quad \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} e^{-t}, \quad \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} e^{2t}, \right).$$

Define the addition $A+B$ and $A \cdot B$ as usual; for example

$$(A \cdot B)_{11} =$$

The entries of A are differentiable (integrable) functions, we can also make sense of

$$\frac{dA}{dt} \qquad\qquad (\int_a^b A(t)dt)$$

by just applying the operator $\frac{d}{dt}$ $(\int_a^b -dt)$ to every coordinate.

3 Linear Equations

A system of n linear equations in n unknowns has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

where the $a'_{ij}s$ and the b_i's are given complex numbers, and the variables $(x_1, \dots, x_n)$ are what we would like to solve for. Of course, it is much more natural to write this with matrix notation:

$$A\underline{x} = \underline{b}, \tag{1.1}$$

where $A = (a_{ij})$ is in $M_n(\mathbb{C})$, while $\underline{x} = (x_1, \dots, x_n)^T$ and $\underline{b} = (b_1, \dots, b_n)^T$ are column vectors in $\mathbb{C}^n$.

If A is non-singular, then $A\underline{x} = \underline{b}$ implies that $\underline{x} = A^{-1}\underline{b}$; so the system has a unique solution. Otherwise, $\det(A) = 0$. Then there must be a linear dependence relation among the columns of A; so the nullspace,

$$N(A) := \{x \mid Ax = 0\} \neq \{0\}.$$

Recall that the nullspace $N(A)$ is always a vector subspace of $\mathbb{C}^n$. Thus, we can find a basis $\{\underline{h}_1, \dots, \underline{h}_k\}$ for $N(A)$.

3.1 Proposition. *If A is a singular matrix, then equation 1.1 either has no solutions, or it has infinitely many solutions. In the latter case, every solution may be uniquely written as*

$$\underline{x} = \underline{x}_p + c_1\underline{h}_1 + \cdots + c_k\underline{h}_k,$$

where $\underline{x}_p$ is some fixed particular solution, and $\{\underline{h}_1, \dots, \underline{h}_k\}$ is a fixed basis for the nullspace $N(A)$.

Proof. Perhaps the equation has no solutions. Otherwise, there is at least one solution, $\underline{x}_p$ that we fix. It follows that $\underline{x}_p + \underline{h}$ is a solution to 1.1 for every $\underline{h}$ in $N(A)$, since

$$A(\underline{x}_p + \underline{h}) = A\underline{x}_p + A\underline{h} = \underline{b} + 0 = \underline{b}.$$

Conversely, assume that $\underline{x}$ is a solution to 1.1, i.e. $A\underline{x} = \underline{b}$. Then

$$A(\underline{x} - \underline{x}_p) = A\underline{x} - A\underline{x}_p = \underline{b} - \underline{b} = 0,$$

which shows that $\underline{x} - \underline{x}_p$ is in $N(A)$. Since we have specified a basis for $N(A)$, we may write $\underline{x} - \underline{x}_p = c_1\underline{h}_1 + \cdots + c_k\underline{h}_k$ uniquely for scalars c_i in $\mathbb{C}$. Thus,

$$\underline{x} = \underline{x}_p + c_1\underline{h}_1 + \cdots + c_k\underline{h}_k.$$ □

This is an important theoretical result for understanding the structure of solutions to linear equations. Now we review the algorithm for computing solutions:

Example. Consider linear system

$$\begin{aligned} x_1 - 2x_3 &= b_1 \\ 2x_2 + 2x_3 &= b_2 \\ x_1 + 4x_2 + 2x_3 &= b_3, \end{aligned}$$

for some fixed $\underline{b} := (b_1, b_2, b_3)^T$ in $\mathbb{C}^3$. We solve this system of equations by putting the augmented matrix in reduced row echelon form:

$$\begin{pmatrix} 1 & 0 & -2 & b_1 \\ 0 & 2 & 2 & b_2 \\ 1 & 4 & 2 & b_3 \end{pmatrix} \xrightarrow[3:\, r_3 - r_1]{2:\, r_2/2} \begin{pmatrix} 1 & 0 & -2 & b_1 \\ 0 & 1 & 1 & b_2/2 \\ 0 & 4 & 4 & b_3 - b_1 \end{pmatrix}$$

$$\xrightarrow{3:\, r_3 - 4r_2} \begin{pmatrix} 1 & 0 & -2 & b_1 \\ 0 & 1 & 1 & b_2/2 \\ 0 & 0 & 0 & b_3 - b_1 - 2b_2 \end{pmatrix}$$

Then there are two-cases. 1. If $b_3 \neq b_1 + 2b_2$, then the system is inconsistent, i.e. there are no solutions. 2. Otherwise, $b_3 = b_1 + 2b_2$, so the last row is $\underline{0}$. Then the first and second rows are the equations:

$$\begin{aligned} x_1 &= b_1 + 2x_3 \\ x_2 &= b_2/2 - x_3 \end{aligned}$$

Hence, every $\underline{x} = (x_1, x_2, x_3)^T$ can be written

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2/2 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix},$$

for some choice of scalar x_3 in $\mathbb{C}$. □

3.2 Remark. Left multiplication by A in the previous example is a linear function from $\mathbb{C}^3$ to itself (symbolically, the function sends $\underline{x} \mapsto A\underline{x}$). Solving the equation $A\underline{x} = \underline{b}$ is the same thing as finding a pre-image of $\underline{b}$. The condition that we derived in the last example, namely that $b_3 = b_1 + 2b_2$ is the condition that $\underline{b}$ is in the image of A.

4 Eigenvalues and Eigenvectors; Diagonalization

4.1 Proposition. *For every linear transformation $\phi : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, there is a matrix A in $M_n(\mathbb{C})$ such that*

$$\phi(\underline{x}) = A \cdot \underline{x} \qquad \text{for every } \underline{x} \text{ in } \mathbb{C}^n.$$

Proof. Let $\underline{e}_i$ denote the standard basis vector whose components are all zero except for a 1 in the ith position. Define $\underline{c}_i := \phi(\underline{e}_i)$, and define

$$A := \begin{pmatrix} \underline{c}_1 & \underline{c}_2 & \cdots & \underline{c}_n \end{pmatrix},$$

the matrix whose ith column is $\underline{c}_i$. It is easy to verify that $\phi(\underline{e}_i) = \underline{c}_i = A\underline{e}_i$. To show that $\phi(\underline{x}) = A \cdot \underline{x}$, just write $\underline{x}$ as a linear combination of the $\underline{e}_i$'s and use linearity of each map. □

Perhaps the simplest kind of linear transformations are the functions defined by

$$\underline{x} \mapsto \lambda\underline{x},$$

where λ is some scalar in $\mathbb{C}$, $\underline{x}$ in $\mathbb{C}^n$. We want to understand an arbitrary linear transformation by approximating it with transformations of this type. More precisely, we ask the question: for which vectors $\underline{x}$ is there a scalar λ such that

$$A\underline{x} = \lambda\underline{x}\,?$$

Notice that if $\underline{x}$ is such a vector, then every scalar multiple of $\underline{x}$ also satisfies the equation. Geometrically, the question is asking about which *directions* does the matrix A act as just scalar multiplication. Of course, we will exclude $\underline{x} = \underline{0}$, since this does not represent any direction and the equation would hold trivially for every value of λ.

4.2 Definition. Consider a square $n \times n$ matrix A.

i) A non-zero vector $\underline{x}$ is an *eigenvector* of A if there exists a scalar λ in $\mathbb{C}$ such that

$$A\underline{x} = \lambda\underline{x}.$$

ii) In this case, we say that λ is an *eigenvalue* of A.

iii) If λ is an eigenvalue, then the set

$$E_\lambda := \{\underline{x} \mid A\underline{x} = \lambda\underline{x}\}$$

is called the corresponding *eigenspace*.

4.3 Proposition. *Consider a square $n \times n$ matrix A. If λ is an eigenvalue, then the eigenspace E_λ is a non-zero vector subspace of $\mathbb{C}^n$.*

Proof. The eigenspace $E_\lambda = N(A - \lambda I_n)$. The nullspace is non-zero since it contains at least one (non-zero) eigenvector, by definition. □

Observe that $N(A - \lambda I_n) \neq \{0\}$ if and only if the matrix $A - \lambda I_n$ is singular. So the eigenvalues of A are precisely those numbers λ such that

$$\det(A - \lambda I_n) = 0. \tag{1.2}$$

This is called the *characteristic equation* of A. It is always a polynomial of degree n in the variable λ.

Example. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a 2×2 matrix. What is the characteristic equation of A?

$$\begin{aligned} 0 = \det(A - \lambda I_2) &= \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} \\ &= (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + ad - bc \\ &= \lambda^2 - tr(A)\lambda + \det(A), \end{aligned}$$

where $tr(A)$ is the *trace* of A, which is the sum of its diagonal entries. □

Example. Let $A = \begin{pmatrix} 2 & 3 \\ 0 & -1 \end{pmatrix}$. We will find all of the eigenvalues and eigenvectors for A.

First, observe that

$$A - \lambda I_2 = \begin{pmatrix} 2-\lambda & 3 \\ 0 & -1-\lambda \end{pmatrix} \qquad \text{and } \det(A - \lambda I_2) = (2-\lambda)(-1-\lambda).$$

So the eigenvalues are exactly the roots of the polynomial, $(2-\lambda)(-1-\lambda)$, i.e. $\lambda = 2, -1$.

Second, we determine the eigenspaces:

$$\begin{aligned} E_2 = N(A - 2I_2) &= N\begin{pmatrix} 0 & 3 \\ 0 & -3 \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\} \\ E_{-1} = N(A - 2I_2) &= N\begin{pmatrix} 3 & 3 \\ 0 & 0 \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}. \end{aligned}$$ □

Let $\underline{b}_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and $\underline{b}_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Observe that every vector in $\mathbb{C}^2$ is a linear combination of eigenvectors, because

$$\beta := \{b_1, b_2\}$$

is a linearly independent set, and thus a basis for $\mathbb{C}^2$. In otherwords, every vector

$$\underline{x} = c_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

for some choice of scalars c_1, c_2. We denote this equality by $\underline{x} = [c_1, c_2]^T_\beta$. Describing vectors this way allows us to write the the linear transformation $x \mapsto Ax$ very simply. Indeed, since $A\underline{x} = c_1 A\underline{b}_1 + c_2 A\underline{b}_2 = 2c_1\underline{b}_1 - c_2\underline{b}_2$, we see that

$$A \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}_\beta = \begin{pmatrix} 2c_1 \\ -c_2 \end{pmatrix}_\beta .$$

Example. Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. We will compute the eigenvalues and eigenvectors of A.

Check that $\det(A - \lambda I_3) = (1-\lambda)^3$. Then $\lambda = 1$ is the only eigenvalue. The eigenspace

$$E_1 = N \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = span\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \}.$$

Not every vector in $\mathbb{C}^3$ can be written as a linear combination of eigenvectors. □

4.4 Definition. Consider a square $n \times n$ matrix A with entries in $\mathbb{C}$, and fix an eigenvalue λ_i of A.

- The *algebraic multiplicity* of λ_i is the highest power m_i such that $(\lambda - \lambda_i)^{m_i}$ divides $\det(A - \lambda I_n)$.
- The *geometric multiplicity* of λ_i is

  $$q_i := \dim E_{\lambda_i},$$

  the dimension of the corresponding eigenspace.

What we saw in the previous example was that the algebraic multiplicity $m_1 = 3$, while the geometric multiplicity $q_1 = 1$.

Fact. $1 \le q_i \le m_i$. □

4.5 Lemma. *Let A be an $n \times n$ matrix over $\mathbb{C}$, and let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues. Choose non-zero vectors x_1 in E_{λ_1}, x_2 in E_{λ_2}, etc. Then $\{x_1, \dots, x_k\}$ is a linearly independent set.*

Proof. Here we employ the technique of mathematical induction: first, we prove the case where $k = 1$; then we prove that if the statement holds for some positive integer $k = m$, then it also holds for $m + 1$. It follows that the statement must hold for every k.

If $k = 1$, then of course $\{x_1\}$ is linearly independent since $x_1 \neq 0$ by assumption. Now suppose the lemma is true for some positive integer $k = m$; we will prove that it holds for

$m+1$. Consider a dependence relation $\sum_{i=1}^{m+1} c_i x_i = 0$; we need to prove that every $c_i = 0$. We multiply this equation on the left by the matrix $A - \lambda_{m+1} I_n$:

$$\begin{aligned} 0 &= (A - \lambda_{m+1} I_n) \cdot \sum_{i=1}^{m+1} c_i x_i = \sum_{i=1}^{m+1} c_i (A - \lambda_{m+1} I_n) x_i \\ &= \sum_{i=1}^{m+1} c_i \left(A x_i - \lambda_{m+1} x_i \right) = \sum_{i=1}^{m+1} c_i (\lambda_i x_i - \lambda_{m+1} x_i) \\ &= \sum_{i=1}^{m} c_i (\lambda_i - \lambda_{m+1}) x_i. \end{aligned}$$

Notice that in the last summation, the upper limit changed since $\lambda_{m+1} - \lambda_{m+1} = 0$. Then this is a linear dependence relation between the m-vectors, $\{x_1, \ldots, x_m\}$, which by induction we are assuming to be linearly independent (since the statement is assumed to be true for $k = m$). It follows that the coefficients $c_i(\lambda_i - \lambda_{m+1}) = 0$ for each $1 \leq i \leq m$. By assumption, $\lambda_i \neq \lambda_{m+1}$; so we see that $c_i = 0$ for each $1 \leq i \leq m$. Then the original dependence relation is

$$0 = \sum_{i=1}^{m+1} c_i x_i = c_{m+1} x_{m+1}.$$

Since $x_{m+1} \neq 0$, we conclude that $c_{m+1} = 0$ as well. This proves that $\{x_1, \ldots, x_{m+1}\}$ is linearly independent, as desired. The result follows by induction. □

4.6 Proposition. *Consider a matrix A in $M_n(\mathbb{C})$. If every eigenvalue of A has algebraic multiplicity 1, then there is a basis of eigenvectors of A.*

Proof. Since every polynomial can be factored as a product of linear factors over $\mathbb{C}$, the characteristic equation of A is

$$0 = \det(A - \lambda I_n) = (\lambda_1 - \lambda) \cdot (\lambda_2 - \lambda) \cdots (\lambda_n - \lambda);$$

each λ_i is distinct since every eigenvalue has multiplicity only 1. So pick $0 \neq x_i$ in E_{λ_i}, and apply the lemma to $\beta = \{x_1, \ldots, x_n\}$. This shows that β is linearly independent. Since it has n elements it must be a basis of $\mathbb{C}^n$. □

4.7 Definition. If A is a matrix such that there is a basis of eigenvectors of A, we say that A is *diagonalizable*. Of course, this is equivalent to the property that every vector can be written as a linear combination of eigenvectors.

4.8 Remark. The proposition is a sufficient, but not a necessary condition for a matrix to be diagonalizable. For example, the identity matrix I_n has only 1 eigenvalue $\lambda = 1$ with algebraic multiplicity n, but it is diagonalizable. Indeed, every vector is an eigenvector, so every basis works.

4.9 Proposition (Diagonalization)**.** *If A is diagonalizable, choose a basis $\beta = \{\underline{x}_1, \dots, \underline{x}_n\}$ of eigenvectors of A. Let λ_i be the corresponding eigenvalue for $\underline{x}_i$ (these eigenvalues may not be distinct). Let Q be the matrix whose jth column is $\underline{x}_j$. Then Q is invertible and*

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

is the diagonal matrix D consisting of the eigenvalues of A.

Proof. The matrix Q is invertible because the set of its columns is β, which is linearly independent. Then $Q^{-1}AQ$ is well-defined, and the equation $Q^{-1}AQ = D$ is equivalent to $AQ = QD$. Observe that

$$AQ = \begin{pmatrix} A\underline{x}_1 & A\underline{x}_2 & \cdots & A\underline{x}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\underline{x}_1 & \lambda_2\underline{x}_2 & \cdots & \lambda_n\underline{x}_n \end{pmatrix}.$$

Denote the standard basis vector $\underline{e}_i$ in $\mathbb{C}^n$, that column vector which is 0 except for 1 in the ith coordinate. Then recall that $Q\underline{e}_i$ is the ith column of Q for *any* matrix Q (in particular, for our choice of Q). It follows that

$$QD = \begin{pmatrix} Q(\lambda_1\underline{e}_1) & Q(\lambda_2\underline{e}_2) & \cdots & Q(\lambda_n\underline{e}_n) \end{pmatrix} = \begin{pmatrix} \lambda_1\underline{x}_1 & \lambda_2\underline{x}_2 & \cdots & \lambda_n\underline{x}_n \end{pmatrix},$$

the same as AQ. □

5 Homogeneous Systems

Consider a linear system of n first-order differential equations with constant coefficients:

$$\underline{x}' = P \cdot \underline{x}, \qquad \underline{x} = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix},$$

and $P = (p_{ij})$ is a matrix in $M_n(\mathbb{R})$.

A solution to this equation has the form

$$\underline{\phi} = \begin{pmatrix} \phi_1(t) \\ \vdots \\ \phi_n(t) \end{pmatrix},$$

where each component function $\phi_i(t)$ is a differentiable function such that

$$\underline{\phi}' = \begin{pmatrix} \phi_1'(t) \\ \vdots \\ \phi_n'(t) \end{pmatrix} = P \begin{pmatrix} \phi_1(t) \\ \vdots \\ \phi_n(t) \end{pmatrix}.$$

Our goal is to find *all* solutions.

Fact. The set of solutions to a homogeneous system of n first-order differential equations is a vector space of dimension n. □

Goal. Find a basis for the solution space, viz. find a set of n solutions, $\{\underline{x}^{(1)}, \dots, \underline{x}^{(n)}\}$ that is linearly independent. □

The point is that if we had such a basis then we would know every solution. Indeed, every solution would be of the form

$$\underline{\phi} = c_1\underline{x}^{(1)} + \cdots c_n\underline{x}^{(n)},$$

for a unique choice of scalars $c_1, \dots, c_n$. We call the set $\{\underline{x}^{(1)}, \dots, \underline{x}^{(n)}\}$ a *fundamental set of solutions* to the system.

5.1 Definition. Consider the linear homogeneous system $\underline{x}' = P\underline{x}$.

- A *fundamental matrix* for the system is a matrix

  $$\Psi(t) = \begin{pmatrix} \underline{x}^{(1)}(t) & \cdots & \underline{x}^{(n)}(t) \end{pmatrix}$$

  whose set of columns $\{\underline{x}^{(1)}, \dots, \underline{x}^{(n)}\}$ is a fundamental set of solutions.

- The *Wronskian* of an (ordered) set of solutions, $\{\underline{x}^{(1)}(t), \dots, \underline{x}^{(n)}(t)\}$ is

  $$W(\underline{x}^{(1)}, \dots, \underline{x}^{(n)}) := \det \begin{pmatrix} \underline{x}^{(1)}(t) & \cdots & \underline{x}^{(n)}(t) \end{pmatrix}.$$

We use the Wronskian to *test* if a given set of n-solutions is a fundamental set of solutions. If the Wronskian is non-zero (for every value of t), then the solutions are linearly independent. But taking the determinant for every value of t is more work than is necesssary–

5.2 Proposition. *If $\{\underline{x}^{(1)}, \dots, \underline{x}^{(n)}\}$ is a set of n solutions to $\underline{x}' = P\underline{x}$, then either*

i) $W(\underline{x}^{(1)}, \dots, \underline{x}^{(n)}) = 0$ for every value of t, or

ii) $W(\underline{x}^{(1)}, \dots, \underline{x}^{(n)}) \neq 0$ for every value of t.

Proof. Perhaps $W(\underline{x}^{(1)}, \dots, \underline{x}^{(n)}) \neq 0$ for every value of t; this is case (ii). Otherwise, there is some t_0 such that

$$W(\underline{x}^{(1)}, \dots, \underline{x}^{(n)})(t_0) = \det \begin{pmatrix} \underline{x}^{(1)}(t_0) & \cdots & \underline{x}^{(n)}(t_0) \end{pmatrix} = 0.$$

The determinant is zero implies that the columns are linearly dependent; choose some dependence relation $c_1\underline{x}^{(1)}(t_0) + \cdots + c_n\underline{x}^{(n)}(t_0) = 0$, where at least one of the c_i's is not zero, and define

$$\underline{y}(t) := c_1\underline{x}^{(1)}(t) + \cdots + c_n\underline{x}^{(n)}(t).$$

Then $\underline{y}(t)$ is a solution to the linear system because it is a linear combination of solutions. Also, $\underline{y}(t_0) = 0$. The zero solution $\underline{0}(t) = (0, \dots, 0)^T$ is also a solution to the system of equations, and satisfies the same initial condition $\underline{0}(t_0) = 0$. By uniqueness of solutions to differential equations, we conclude that $\underline{y}(t)$ is the zero solution, viz.

$$\underline{y}(t) = c_1\underline{x}^{(1)}(t) + \cdots + c_n\underline{x}^{(n)}(t) = 0$$

for *every* value of t (not just t_0). This dependence relation proves that

$$W(\underline{x}^{(1)}, \dots, \underline{x}^{(n)}) = \det \begin{pmatrix} \underline{x}^{(1)}(t) & \cdots & \underline{x}^{(n)}(t) \end{pmatrix} = 0.$$ □

The case $n = 1$

If $n = 1$, a linear homogeneous system with constant coefficients is just the following differential equation:

$$x_1'(t) = p_{11}x(t).$$

This equation is *separable*; when you separate variables and solve you find the general solution

$$x_1(t) = Ce^{p_{11}t}.$$

The case $n > 1$

Now if $n > 1$, consider the system $\underline{x}' = P\underline{x}$, P in $M_n(\mathbb{C})$. We look for solutions that have the same form as the $n = 1$ case, namely

$$\underline{x}(t) = Ce^{\lambda t},$$

where C is a constant column in $\mathbb{C}^n$ and λ is some scalar in $\mathbb{C}$. Compute:

$$\underline{x}' = \lambda Ce^{\lambda t}$$
$$P\underline{x} = (PC)e^{\lambda t}$$

So $\underline{x}$ is a solution if the difference

$$P\underline{x} - \underline{x}' = (PC - \lambda C)\, e^{\lambda t} = 0,$$

for every value of t. Since $e^{\lambda t} \neq 0$, we divide and get

$$(P - \lambda I_n)C = 0, \qquad \Leftrightarrow \qquad PC = \lambda C.$$

Thus, our guess $\underline{x}(t)$ *is* a solution if and only if C is an eigenvector of P with eigenvalue λ.

Example. We solve the homogeneous linear system

$$\begin{aligned} x_1' &= x_1 - bx_2 \\ x_2' &= 2x_2 \end{aligned}$$

where b is some scalar in $\mathbb{C}$.

First, re-write in matrix form: $\underline{x}' = P\underline{x}$, where

$$P = \begin{pmatrix} 1 & -b \\ 0 & 2 \end{pmatrix}.$$

We know that functions of the form $\underline{x} = Ce^{\lambda t}$ are solutions when C is an eigenvalue of P with eigenvalue λ. The eigenvalues of P are the solutions to the characteristic equation

$$0 = \det(P - \lambda I_2) = (1 - \lambda)(2 - \lambda);$$

so the eigenvalues are $\lambda = 1, 2$. Next, we compute the eigenspaces:

$$\begin{aligned} E_1 = N(P - I_2) &= N\begin{pmatrix} 0 & -b \\ 0 & 1 \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\} \\ E_2 = N(P - 2I_2) &= N\begin{pmatrix} -1 & b \\ 0 & 0 \end{pmatrix} = span\left\{\begin{pmatrix} b \\ 1 \end{pmatrix}\right\}. \end{aligned}$$

So we have two explicit solutions

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t = \begin{pmatrix} e^t \\ 0 \end{pmatrix} \qquad \underline{x}_2 = \begin{pmatrix} b \\ 1 \end{pmatrix} e^{2t} = \begin{pmatrix} be^{2t} \\ e^{2t} \end{pmatrix}.$$

We check if these two solutions are linearly independent by computing the Wronskian

$$W(\underline{x}_1, \underline{x}_2) = \det\begin{pmatrix} e^t & be^{2t} \\ 0 & e^{2t} \end{pmatrix} = e^{3t};$$

since their Wronskian is not zero, these two solutions are independent. Thus, every solution has the form

$$\underline{x} = c_1\underline{x}^{(1)} + c_2\underline{x}^{(2)}.$$

□

5.3 Remark. We did not need to compute the Wronskian for every value of t. From Proposition 5.2, we know that it is sufficient to only check the value at $t = 0$ (or any other time), which means computing the determinant of the two eigenvectors

$$\det \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = 1.$$

From Lemma 4.5, we know that these two eigenvectors are linearly independent because they are chosen from distinct eigenspaces. For these "theoretical" reasons, the computation was somewhat unnecessary.

Recall: Reduction of Order

Recall from Math 307 the technique of *reduction of order*: consider the linear 2nd order equation

$$y''(t) + p(t)y'(t) + q(t)y(t) = 0;$$

and suppose that $y_1(t) \neq 0$ is a solution. The method of finding another independent solution is to guess:

$$y_2(t) = w(t) \cdot y_1(t);$$

then "plug-in" this guess and reduce to an equation that you can solve for $w(t)$.

5.4 Example. Solve $y'' - 2y' + y = 0$, given that $y_1(t) = e^t$ is a solution.

We guess that $y_2(t) = w(t)e^t$; it follows that $y_2'(t) = (w' + w)e^t$ and $y_2'' = (w + 2w' + w'')e^t$. So,

$$\begin{aligned} 0 &= y_2'' - 2y_2' + y_2 \\ &= ((w + 2w' + w'') - 2(w + w') + w)\, e^{2t} \\ &= w'' e^{2t}. \end{aligned}$$

Since $e^{2t} \neq 0$ for every t, we get $w''(t) = 0$. It follows that $w(t) = at + b$ is (any) polynomial of order 1. So

$$y_2(t) = (at + b)e^t,$$

is a solution, and $\{y_1, y_2\}$ is a fundamental set of solutions.

In the next example, we apply an analogue of this to a linear system.

Example. We solve the system

$$\begin{aligned} x_1' &= x_1 + x_2 \\ x_2' &= x_2. \end{aligned}$$

Re-writing in matrix form, we have $\underline{x}' = P\underline{x}$, for $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Since P is upper-triangular, we know its eigenvalues are its diagonal entries. Thus $\lambda = 1$ is the only eigenvalue. The corresponding eigenspace is

$$E_1 = N(P - I_2) = N\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}.$$

This implies that

$$\underline{x}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t$$

is a solution to the system of equations. But how to find another solution?

We make the bold guess of copying the technique of reduction of order. Namely, we guess

$$\underline{x}^{(2)} = (\underline{a}t + \underline{b})e^t,$$

where $\underline{a}$ and $\underline{b}$ are some vectors to be determined. We check our guess with the equation $\left(\underline{x}^{(2)}\right)' = P\underline{x}^{(2)}$:

$$\begin{aligned} P\underline{x}^{(2)} &= P(\underline{a}t + \underline{b})e^t = (P\underline{a})(te^t) + (P\underline{b})e^t \\ \left(\underline{x}^{(2)}\right)' &= \underline{a}(te^t)' + \underline{b}(e^t)' = \underline{a}te^t + \underline{a}e^t + \underline{b}e^t \\ &= \underline{a}te^t + (\underline{a} + \underline{b})e^t. \end{aligned}$$

We want $P\underline{x}^{(2)} - \left(\underline{x}^{(2)}\right)' = 0$, i.e.

$$[(P - I)\underline{a}]\,(te^t) + [(P - I)\underline{b} - \underline{a}]\,e^t = 0.$$

But $\{te^t, e^t\}$ are linearly independent since they form a fundamental set of solutions to the differential equation

$$y'' - 2y' + y = 0.$$

So we know that

$$\begin{aligned} (P - I_2)\,\underline{a} &= 0 \\ (P - I_2)\underline{b} &= \underline{a}. \end{aligned}$$

The first equation says that $\underline{a}$ is an eigenvector of P with eigenvalue 1; we might as well choose $\underline{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. We then just solve for $\underline{b}$ as usual; the augmented matrix is

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

which is already in reduced row echelon form. One solution is $\underline{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then we want to check that the two solutions

$$\underline{x}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t \qquad \underline{x}^{(2)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} te^t + \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^t$$

form a fundamental set of solutions. So compute the Wronskian:

$$W(\underline{x}^{(1)}, \underline{x}^{(2)}) = \det \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix} = e^{2t} \neq 0.$$

So every solution is a linear combination of these two. □

Exercises. Solve each of these linear systems:

1.

$$\begin{aligned} x_1' &= -x_1 + 4x_2 \\ x_2' &= -x_4 + 3x_2. \end{aligned}$$

2.

$$\begin{aligned} x_1' &= 4x_1 - 2x_2 \\ x_2' &= 8x_1 - 4x_2. \end{aligned}$$

□

6 Phase Planes

Example. Solve

$$\underline{x}' = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \underline{x}.$$

The characteristic equation of the coefficient matrix is $0 = \lambda^2 - 2\lambda - 8 = (\lambda + 2)(\lambda - 4)$; so $\lambda = -2, 4$ are the two eigenvalues of the coefficient matrix. Then compute

$$\begin{aligned} E_{-2} &= N \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} = span \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \\ E_4 &= N \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} = span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

So $\underline{x}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-2t}$ and $\underline{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{4t}$ are both solutions. These are linearly independent (check the Wronskian), so every solution can be written

$$\underline{x} = c_1\underline{x}_1 + c_2\underline{x}_2,$$

for some choice of scalars c_1 and c_2. □

One often wants to understand what the general solution "looks like" qualitatively. We draw the *phase plane*: for each point (x_1, x_2) in $\mathbb{R}^2$, we can draw the tangent vector to the solution at that point. We can imagine a particle being placed at the initial position, and then moving along the path which is given by the unique solution to the system of equations.

In the previous example, if the starting point is in one of the eigenspaces, E_{-2} or E_4, then the particle remains in the eigenspace for all time. This is because the direction of motion, i.e. the tangent vector to the solutions at that point, are vectors *in that space*.

6.1 Definition.

- A *trajectory* is a graph of a solution to the system.
- A phase plane that includes a few trajectories is called a *phase portrait*.
- An *equilibrium solution* to a differential equation is a constant solution.

Notice that the zero solution $\underline{x} = \underline{0}$ in $\mathbb{C}^n$ is an equilibrium solution to every homogeneous linear system of n variables.

Consider the previous example. If the starting point is near $\underline{0}$, but not in the eigenspace E_{-2}, then as time moves forward the particle moves away from $\underline{0}$ (and actually diverges "to ∞"). However, if the particle starts in the eigenspace E_{-2}, then it does move towards $\underline{0}$. In this case we say that the equilibrium solution $\underline{0}$ is a saddle point, since it "attracts" certain directions and "repels" others. More formally:

6.2 Definition. Consider the differential equation $\underline{x}' = P\underline{x}$, where the coefficient matrix P is in $M_2(\mathbb{R})$. If P has real non-zero eigenvalues of opposite signs, say $\lambda_1 < 0 < \lambda_2$, then $\underline{0}$ is a *saddle point* and is an *unstable equilibrium*.

Example. Consider the system $\underline{x}' = P\underline{x}$, where $P = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix}$. You can check that the eigenvalues are $-2, -4$, and

$$E_{-2} = N\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$
$$E_{-4} = N\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}.$$

If you sketch the phase plane, you see that all tangent lines point towards $\underline{0}$; so we say that $\underline{0}$ is *asymptotically stable*, and a *node* for the system of equations. □

6.3 Definition. Consider the differential equation $\underline{x}' = P\underline{x}$, where the coefficient matrix P is in $M_2(\mathbb{R})$. If P has real non-zero eigenvalues of the same sign, then $\underline{0}$ is a *node*. If both eigenvalues are negative, then we say that $\underline{0}$ is *asymptotically stable*; if both eigenvalues are positive, we say that $\underline{0}$ is *unstable*.

EXAMPLE. Consider $\underline{x}' = P\underline{x}$ for $P = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then P has eigenvalues $\lambda = \pm i$; compute that

$$E_{+i} = N\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ i \end{pmatrix}\right\}$$
$$E_{-i} = N\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ -i \end{pmatrix}\right\}. \qquad \square$$

How do we draw the phase plane for the previous example? It is not possible to draw the eigenspaces, since they only intersect $\mathbb{R}^2$ at $\underline{0}$ (but they are non-zero subspaces of $\mathbb{C}^2$. Instead, we have to sketch many tangent vectors. If you do this, you will see that the trajectories are *circles*! In this case, $\underline{0}$ is called a *center*; it is a *stable equilibrium.*

EXAMPLE. Consider $\underline{x}' = P\underline{x}$ for $P = \begin{pmatrix} -1/2 & 1 \\ -1 & -1/2 \end{pmatrix}\underline{x}$. The eigenvalues of P are $\lambda = -1/2 \pm i$; compute that

$$E_{-1/2+i} = N\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ i \end{pmatrix}\right\}$$
$$E_{-1/2-i} = N\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} = span\left\{\begin{pmatrix} 1 \\ -i \end{pmatrix}\right\}.$$

Then two solutions are $\underline{x}_1 = \begin{pmatrix} 1 \\ i \end{pmatrix} e^{(-1/2+i)t}$ and $\underline{x}_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{(-1/2-i)t}$. Observe that

$$\frac{1}{2}(\underline{x}_1 + \underline{x}_2) = e^{-t/2}\begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix}$$
$$\frac{1}{2i}(\underline{x}_1 + \underline{x}_2) = e^{-t/2}\begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}.$$

Check (by computing the Wronskian) that these particular linear combinations are still linearly independent; they have the advantage of being real-valued. We can easily understand the behavior of these solutions – a particle following either solution moves clock-wise around the origin, and its distance from the origin is exactly $e^{-t/2}$; thus the particle is "spiraling" inwards. We say that $\underline{0}$ is a *spiral point*, and is *asymptotically stable.* $\square$

If the real part of the eigenvalue in the previous example were positive, instead of negative, then the solutions would "spiral out" instead of spiraling inwards. In that case, we say that $\underline{0}$ is an unstable spiral point.

6.1 Improper Nodes / Repeated Eigenvalues

6.4 Definition. A *defective matrix* is a square matrix that is not diagonalizable.

EXAMPLE. Consider the equation $\underline{x}' = P\underline{x}$, for $P = \begin{pmatrix} -3 & 1 \\ -1 & -1 \end{pmatrix}$.

The eigenvalues of P are the solutions of the characteristic equation

$$0 = \lambda^2 + 4\lambda + 4 = (\lambda + 2)^2.$$

Thus, $\lambda = -2$ is the only eigenvalue. THen

$$E_{-2} = N \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} = span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

Since the geometric multiplicity is 1, and the algebraic multiplicity is 2, we see that P is defective! Of course,

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-2t},$$

is one solution. We have to find a second linearly independent solution. First, *guess* $x_2(t) = (\underline{a}t + \underline{b})e^{-2t}$; then check (compute)

$$\begin{aligned} \underline{x}_2' &= (\underline{a})e^{-2t} + (\underline{a}t + \underline{b})e^{-2t} \cdot (-2) \\ &= (-2\underline{a})te^{-2t} + (\underline{a} - 2\underline{b})e^{-2t} \\ P\underline{x}_2 &= (P\underline{a})te^{-2t} + (P\underline{b})e^{-2t}. \end{aligned}$$

Then $\underline{x}_2' = P\underline{x}_2$ if and only if the coefficients above are equal, viz. $P\underline{a} = -2\underline{a}$ and $P\underline{b} = \underline{a} - 2\underline{b}$. Check that this is equivalent to the system

$$\begin{aligned} (P + 2I_2)\underline{a} &= 0 \\ (P + 2I_2)\underline{b} &= \underline{a}. \end{aligned}$$

The first equation just says that $\underline{a}$ is an eigenvector of P with eigenvalue -2, i.e. $\underline{x} = c \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. We will just pick $c = 1$. Then the second equation is just a linear equation for $\underline{b}$ (since we have already fixed $\underline{a}$). If $(P + 2I_2)\underline{b} = \underline{a} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, find $\underline{b}$ by

$$\begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \overset{rref}{\rightarrow} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Setting the free variable $b_2 = 1$, we see that $b_1 = -1 + b_2 = 0$, i.e. $\underline{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is a solution. Thus,

$$\underline{x}_2 = \underline{a}te^{-2t} + \underline{b}e^{-2t} = \underline{a} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} te^{-2t} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-2t}$$

is a solution. Finally, we check that these two are linearly independent. We compute the Wronskian

$$W(\underline{x}_1, \underline{x}_2) = \det \begin{pmatrix} e^{-2t} & te^{-2t} \\ e^{-2t} & te^{-2t} + e^{-2t} \end{pmatrix} = e^{-4t} \neq 0.$$

Since the Wronskian is not zero, the two solutions are independent and thus every solution $\underline{x}$ to the system is a linear combination

$$\underline{x} = c_1\underline{x}_1 + c_2\underline{x}_2$$

for some choice of scalars c_1 and c_2. □

6.5 Remark. In the previous example, notice that

$$(P + 2I_2)^2\underline{b} = (P + 2I_2) \cdot (P + 2I_2)\underline{b} = (P + 2I_2)\underline{a} = 0.$$

This partially motivates the following definition.

6.6 Definition. Let P be a square $n \times n$ matrix with entries in $\mathbb{C}$, and let λ be an eigenvalue of P.

i) If $\underline{b}$ is a non-zero vector in $\mathbb{C}^n$ such that

$$(P - \lambda I_n)^k\underline{b} = 0,$$

for some integer $k \geq 1$, then $\underline{b}$ is called a *generalized eigenvector* of P.

ii) Define $\mathcal{E}_\lambda := \sum_{k\geq 1} N(P - \lambda I_n)^k$; this is the linear span of all of the generalized eigenvectors with eigenvalue λ. This is called the *generalized eigenspace* corresponding to eigenvalue λ.

What about the equilibrium solution? The equilibrium $\underline{x} = \underline{0}$ in the previous example is (informally) half-way between a node and a spiral point. There is one eigendirection that flows directly towards the origin, and then there is a generalized eigendirection that "spirals in" – in this case we say that $\underline{0}$ is an *improper node*. Since all paths go towards the origin, we call it *asymptotically stable.*

Exercises. Solve the following systems of equations, graph the phase plane and classify the equilibrium solution, $\underline{0}$:

i)

$$\begin{aligned} x_1' &= -x_1 + 4x_2 \\ x_2' &= -x_1 + 3x_2 \end{aligned}$$

ii)

$$\begin{aligned} x_1' &= 4x_1 - 2x_2 \\ x_2' &= 8x_1 - 4x_2 \end{aligned}$$

□

7 Functional Calculus

Consider the linear homogeneous equation with constant coefficients:

$$x'(t) = Px(t), \qquad \text{for } P \text{ in } \mathbb{C}.$$

We know a solution to this is given by $x(t) = e^{Pt}$.

Naive Question: What if we replace P in $\mathbb{C}$ with a matrix P in $M_n(\mathbb{C})$?

The reason this might work is that the (formal) Taylor series

$$e^{Pt} = \sum_{k=0}^{\infty} \frac{1}{k!}(Pt)^k$$

satisfies the differential equation $x'(t) = Px(t)$ formally as a Taylor series, and all of the expressions $\frac{1}{k!}(Pt)^k$ still make sense when P is a matrix.

7.1 Definition. Let A be an $n \times n$ matrix, and let $f(t) = \sum_{k=0}^{\infty} a_k t^k$ be a convergent power series. Define

$$f(A) := \lim_{K \longrightarrow \infty} \sum_{k=0}^{K} a_k A^k,$$

whenever this limit exists.

Problem. How can we effectively compute $f(A)$?

First assume that A is diagonal; we will compute $f(A)$. If $f(t) = \sum_{k=0}^{\infty} a_k t^k$;

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad \Rightarrow \quad a_k A^k = \begin{pmatrix} a_k\lambda_1^k & 0 & \cdots & 0 \\ 0 & a_k\lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_k\lambda_n^k \end{pmatrix},$$

so that

$$f(A) = \lim_{K \longrightarrow \infty} \begin{pmatrix} \sum_{k=0}^{K} a_k\lambda_1^k & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{K} a_k\lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{k=0}^{K} a_k\lambda_n^k \end{pmatrix} = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}.$$

Now assume that A is not necessarily diagonal, but is diagonalizable. Then there is an invertible matrix Q such that $Q^{-1}AQ = D$ is diagonal. It is more convenient to write $A = QDQ^{-1}$, since then

$$\begin{aligned}
A^2 &= (QDQ^{-1})(QDQ^{-1}) = QD^2Q^{-1} \\
A^3 &= (QDQ^{-1})(QD^2Q^{-1}) = QD^3Q^{-1} \\
&\vdots \\
A^k &= (QDQ^{-1})(QD^{k-1}Q^{-1}) = QD^kQ^{-1}.
\end{aligned}$$

It follows that

$$f(A) = \sum_{k=0}^{\infty} a_k A^k = \sum_{k=0}^{\infty} a_k QD^kQ^{-1} = Q\left(\sum_{k=0}^{\infty} a_k D^k\right)Q^{-1} = Qf(D)Q^{-1}.$$

Since we already know how to compute $f(D)$, we know how to compute $f(A)$.

7.2 Corollary. *If A is diagonalizable, then $f(A)$ exists provided that $f(t)$ is a convergent Taylor series in an interval containing all of the eigenvalues of the matrix A.*

The next theorem shows how this is useful to us.

7.3 Theorem. *Let $f(t) = \sum_{k=0}^{\infty} \frac{1}{k!}t^k = e^t$, and consider the homogeneous linear system $\underline{x}' = P\underline{x}$. If $f(Pt)$ exists, then $\Phi(t) := f(Pt)$ is a fundamental matrix for the system. Moreover, if a solution $\underline{x}$ is written as*

$$\underline{x} = \Phi(t) \cdot \underline{c},$$

then $\underline{c} = \underline{x}(0)$ is the initial condition.

Proof. First, we show that $\Phi'(t) = P \cdot \Phi(t)$, by looking at the formal Taylor series:

$$\Phi'(t) = \left(\sum_{k=0}^{\infty} \frac{1}{k!}P^k t^k\right)' = \sum_{k=1}^{\infty} \frac{1}{(k-1)!}P^k t^{k-1} = P \cdot \Phi(t).$$

What was really important here is that P and $I_n \cdot t$ commute so that we could write $(Pt)^k$ as $P^k \cdot t^k$.

Since $\Phi'(t) = P \cdot \Phi(t)$, and left multiplication by P is done column-by-column, we see that each column of $\Phi(t)$ is a solution to the equation $\underline{x}' = P\underline{x}$. Furthermore, at $t = 0$ these solutions are independent since $\Phi(0) = I_n$; since they are independent at one point in time, they must be independent for all t. Thus $\Phi(t)$ is a fundamental matrix for the system (its columns are a fundamental set of solutions).

Since $\underline{x} = \Phi(t) \cdot \underline{c}$ is a linear combination of solutions, it is a solution; evidently,

$$\underline{x}(0) = \Phi(0) \cdot \underline{c} = I_n \cdot \underline{c} = \underline{c};$$

thus $\underline{c}$ is precisely the initial condition $\underline{x}(0)$. $\square$

EXAMPLE. Solve the system of equations $\underline{x}' = \begin{pmatrix} 1 & -2 \\ 0 & 2 \end{pmatrix} \underline{x}$.

Let P be the coefficient matrix; we have already diagonalized P in a worksheet: the eigenvalues are 1 and 2, and corresponding eigenvectors are $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$; then P is diagonalizable by $Q = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$, and we know that

$$Q^{-1}PQ = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = D.$$

It follows that $P = QDQ^{-1}$, and so for every value of t, $Pt = Q(Dt)Q^{-1}$. It follows that

$$\begin{aligned} e^{Pt} &= Q \cdot e^{Dt} \cdot Q^{-1} \\ &= Q \cdot \begin{pmatrix} e^t & 0 \\ 0 & e^{2t} \end{pmatrix} \cdot Q^{-1} \\ &= \begin{pmatrix} e^t & 2e^t - 2e^{2t} \\ 0 & e^{2t} \end{pmatrix}. \end{aligned}$$

So the general solution is

$$\underline{x} = \begin{pmatrix} e^t & 2e^t - 2e^{2t} \\ 0 & e^{2t} \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}.$$

□

7.1 Defective Matrices and Jordan Blocks

Unfortunately, not every coefficient matrix A is diagonalizable! Even when A is defective, $\Phi(t) = e^{At}$ is a fundamental matrix for the system of equations

$$\underline{x}' = A\underline{x},$$

but it is less clear how to actually compute $\Phi(t)$. This is one of the uses of eigendecomposition (Jordan decomposition) of the matrix, A. I will not prove the following fact here, but it is true that we can always find an invertible matrix Q such that

$$J := Q^{-1}AQ$$

is a matrix in Jordan form. This means that J is a "block-diagonal" matrix

$$J = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & B_\ell. \end{pmatrix}, \qquad B_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i. \end{pmatrix}.$$

Of course, it is possible to have a block B_i of size 1×1, in which case it is just the 1×1 matrix (λ_i).

Just as before, $A^n = QJ^nQ^{-1}$, and so $f(At) = Qf(Jt)Q^{-1}$. Since Jt is a block-diagonal matrix, we only need to know how to compute each block $f(B_it)$, since

$$f(Jt) = \begin{pmatrix} f(B_1t) & 0 & \cdots & 0 \\ 0 & f(B_2t) & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & f(B_\ell t) \end{pmatrix}.$$

Extra Credit. (2 points each) Fix a complex number λ in $\mathbb{C}$, and let A be the $m \times m$ Jordan block

$$A = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

In particular, λ is the only eigenvalue of A; the geometric multiplicity $q = \dim(E_\lambda) = 1$, while the algebraic multiplicity of λ is m.

Fix a function $f(t) = \sum_{k=0}^{\infty} a_k t^k$ which converges in some interval containing λ.

(I) Prove that A^k is upper-triangular, and that it is constant along its "super-diagonals."

(II) Prove by induction on k that

$$(A^k)_{1,j} = \binom{k}{j-1} \lambda^{k-(j-1)};$$

here $\binom{k}{j-1} = \frac{k!}{(j-1)!(k-(j-1))!}$ is the binomial coefficient (think Pascal's triangle).

Hint: Use the recursive identity from Pascal's triangle that

$$\binom{k-1}{j-1} + \binom{k-1}{j} = \binom{k}{j}.$$

(III) By differentiating $f(t)$, show that

$$\frac{f^{(n)}(\lambda)}{n!} = \sum_{k=n}^{\infty} \binom{k}{n} a_k \lambda^{k-n}.$$

(IV) Use all of these facts to conclude that the ijth entry of $f(A)$ is given by

$$f(A)_{i,j} = \begin{cases} \frac{f^{(j-i)}(\lambda)}{(j-i)!} & \text{if } j \geq i \\ 0 & \text{otherwise.} \end{cases}$$

It is not much more difficult to compute $f(At)$:

$$f(At) = \begin{pmatrix} f(\lambda t) & t \cdot f'(\lambda t) & t^2 \frac{f''(\lambda t)}{2!} & \cdots & t^{m-2} \frac{f^{(m-2)}(\lambda t)}{(m-2)!} & t^{m-1} \frac{f^{(n-1)}(\lambda t)}{(n-1)!} \\ 0 & f(\lambda t) & t \cdot f'(\lambda t) & \cdots & t^{m-3} \frac{f^{(m-3)}(\lambda t)}{(m-3)!} & t^{m-2} \frac{f^{(m-2)}(\lambda t)}{(n-2)!} \\ 0 & 0 & f(\lambda t) & \cdots & t^{m-4} \frac{f^{(m-4)}(\lambda t)}{(m-4)!} & t^{m-3} \frac{f^{(m-3)}(\lambda t)}{(m-3)!} \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & f(\lambda t) & t \cdot f'(\lambda t) \\ 0 & 0 & 0 & \cdots & 0 & f(\lambda t) \end{pmatrix}$$

Indeed, the ijth entry of $f(At)$ is

$$f(At)_{i,j} = \begin{cases} t^{j-i} \cdot \frac{f^{(j-i)}(\lambda)}{(j-i)!} & \text{if } j \geq i \\ 0 & \text{otherwise.} \end{cases}$$

7.2 Applications

Example. Consider the system:

$$\underline{x}' = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \cdot \underline{x}.$$

If A is the coefficient matrix, then a fundamental matrix is given by $\Phi(t) = e^{At}$; we can quickly compute this using the previous formula. Indeed, apply the formula for $f(t) = e^t$, $m = 3$, $\lambda = 2$, and obtain

$$e^{At} = \begin{pmatrix} e^{2t} & te^{2t} & t^2 e^{2t}/2! \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix}.$$

Thus, the general solution is given by

$$\underline{x} = \begin{pmatrix} e^{2t} & te^{2t} & t^2 e^{2t}/2! \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix} \cdot \underline{c}.$$

Just as in the diagonalizable case, $\underline{c} = \underline{x}(0)$ is the initial condition. □

Now we consider a slightly more interesting example, where the Taylor series that we use is not the exponential function. Recall that the general solution to the second-order linear equation

$$y'' + a^2 y = 0$$

is $y = c_1 \cos(at) + c_2 \sin(at)$.

Example. Consider the *second-order* linear system:

$$\underline{x}'' = \begin{pmatrix} -4 & -4 \\ 0 & -4 \end{pmatrix} \underline{x}.$$

We can rewrite this equation as

$$\underline{x}'' + \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}^2 \underline{x} = 0.$$

Let $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$; it may be surprising to see that the general solution is given by

$$\underline{x} = \cos(At) \cdot \underline{c}^{(1)} + \sin(At) \cdot \underline{c}^{(2)},$$

in perfect analogy with the problem in 1-dimension. The point is that the matrix functions $\sin(At)$ and $\cos(At)$ satisfy the same identities as the usual functions $\sin(at)$ and $\cos(at)$, namely:

$$[\sin(At)]' = A\cos(At) \qquad [\cos(At)]' = -A\sin(At).$$

It remains to compute the matrices $\cos(At)$, and $\sin(At)$. Again, we use the formula from the previous section with $\lambda = 2$, $m = 2$, and $f(t)$ either $\cos(t)$ or $\sin(t)$:

$$\cos(At) = \begin{pmatrix} \cos(2t) & -t\sin(2t) \\ 0 & \cos(2t) \end{pmatrix} \qquad \sin(At) = \begin{pmatrix} \sin(2t) & t\cos(2t) \\ 0 & \sin(2t) \end{pmatrix}.$$

Explicitly, the general solution is then

$$\underline{x} = \begin{pmatrix} \cos(2t) & -t\sin(2t) \\ 0 & \cos(2t) \end{pmatrix} \begin{pmatrix} c_1^{(1)} \\ c_2^{(1)} \end{pmatrix} + \begin{pmatrix} \sin(2t) & t\cos(2t) \\ 0 & \sin(2t) \end{pmatrix} \begin{pmatrix} c_1^{(2)} \\ c_2^{(2)} \end{pmatrix}.$$ □

8 Nonhomogeneous Systems

The general form of a non-homogeneous system is

$$\underline{x}' = P\underline{x} + \underline{g},$$

where P is an $n \times n$-matrix of continuous functions, $\underline{g}$ is a column of length n of continuous functions and $\underline{x}$ is to be determined. Just like with solving linear equations, for a fixed particular solution $\underline{x}_p$ to the system, every solution can be written as

$$\underline{x} = \underline{x}_p + \underline{x}_h,$$

where $\underline{x}_h$ satisfies the *corresponding homogeneous equation*, namely $\underline{x}' = P\underline{x}$. In this class, we will almost always restrict our attention to the case where the coefficient matrix P is constant.

Diagonalization/Substitution

Consider the linear equation $x' = px + g$, where p is a constant and $g(t)$ is a continuous function. In Math 307, we learn how to solve this equation using integrating factors. Indeed, if $\mu = e^{-pt}$, then the equation is equivalent to

$$\frac{d}{dt}\left[\mu \cdot x\right] = \mu g \Rightarrow x(t) = \frac{1}{\mu}\int_0^t \mu(s) \cdot g(s)ds + \frac{C}{\mu}.$$

The particular solution is $\underline{x}_p(t) = \frac{1}{\mu(t)}\int_0^t \mu(s) \cdot g(s)ds$, while any multiple of $\underline{x}_h = \frac{1}{\mu} = e^{pt}$ is a solution to the associated homogeneous equation $x' = px$.

Now suppose that we have instead the system $\underline{x}' = P\underline{x} + \underline{g}$, where P is in $M_n(\mathbb{C})$. For this technique to work, we must assume that P is diagonalizable. If it is, then we can find a matrix Q such that

$$Q^{-1}PQ = D$$

is the diagonal matrix consisting of the eigenvalues of P. Now, substitute $\underline{y} := Q^{-1}\underline{x}$. It follows that $\underline{y}' = Q^{-1}\underline{x}'$. Then we have

$$\begin{aligned} \underline{x}' = P\underline{x} + \underline{g} &\Leftrightarrow Q^{-1}\underline{x}' = Q^{-1}P\underline{x} + Q^{-1}\underline{g} \\ &\Leftrightarrow Q^{-1}\underline{x}' = Q^{-1}P(QQ^{-1})\underline{x} + Q^{-1}\underline{g} \\ &\Leftrightarrow Q^{-1}\underline{x}' = Q^{-1}PQ(Q^{-1}\underline{x}) + Q^{-1}\underline{g} \\ &\Leftrightarrow \underline{y}' = D\underline{y} + Q^{-1}\underline{g}. \end{aligned}$$

Since D is a diagonal matrix, these equations are *uncoupled*: to solve them we can just apply the method of integrating factors above to each of the n equations of the form

$$\underline{y}_i' = \lambda_i \underline{y}_i + \left(Q^{-1}\underline{g}\right)_i.$$

Example. Solve the linear system

$$\underline{x}' = \begin{pmatrix} 5 & 4 \\ -6 & -5 \end{pmatrix}\underline{x} + \begin{pmatrix} e^{2t} \\ 2e^{2t} + t \end{pmatrix}.$$

We check to see if the coefficient matrix is diagonalizable. Let $P = \begin{pmatrix} 5 & 4 \\ -6 & -5 \end{pmatrix}$, then the eigenvalues of P are $\lambda = \pm 1$; check that

$$E_{+1} = span\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}, \qquad E_{-1} = span\left\{\begin{pmatrix} 2 \\ -3 \end{pmatrix}\right\}.$$

If $Q = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}$, we know that

$$Q^{-1}PQ = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} =: D$$

is diagonal. Then substitute $\underline{y} = Q^{-1}\underline{x}$; by the previous calculation we get the equivalent system

$$\begin{aligned} \underline{y}' &= D\underline{y} + Q^{-1} \begin{pmatrix} e^{2t} \\ 2e^{2t} + t \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \underline{y} + \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} e^{2t} \\ 2e^{2t} + t \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \underline{y} + \begin{pmatrix} 7e^{2t} + 2t \\ -3e^{2t} - t \end{pmatrix}. \end{aligned}$$

In otherwords, we want to solve the two (uncoupled) equations

$$\begin{aligned} y_1' &= y_1 + (7e^{2t} + 2t) \\ y_2' &= -y_2 + (-3e^{2t} - t). \end{aligned}$$

Using the method of integrating factors above, we get that

$$\begin{aligned} y_1 &= e^t \left(\int_0^t e^{-s}(7e^{2s} + 2s)ds + C_1 \right) \\ y_2 &= e^{-t} \left(\int_0^t e^{s}(-3e^{2s} - s)ds + C_2 \right). \end{aligned}$$

By integrating, we get

$$\begin{aligned} y_1 &= 7e^{2t} - 2(t+1) + C_1 e^t \\ y_2 &= -e^{2t} - (t-1) + C_2 e^{-t}. \end{aligned}$$

This is the general solution for y. Since $\underline{x} = Q\underline{y}$, we compute

$$\begin{aligned} \underline{x} &= \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix} \left(\begin{pmatrix} 7 \\ -1 \end{pmatrix} e^{2t} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} t + \begin{pmatrix} -2 \\ 1 \end{pmatrix} + C_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t + C_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-t} \right) \\ &= \begin{pmatrix} 5 \\ -4 \end{pmatrix} e^{2t} + \begin{pmatrix} -4 \\ 5 \end{pmatrix} t + \begin{pmatrix} 0 \\ -1 \end{pmatrix} + C_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^t + C_2 \begin{pmatrix} 2 \\ -3 \end{pmatrix} e^{-t}. \end{aligned}$$

□

Variation of Parameters

In Math 307, there is a technique for solving

$$y'' + p(t)y' + q(t)y = g(t)$$

called variation of parameters. It involves first solving the associated homogeneous equation

$$y'' + p(t)y' + q(t)y = 0;$$

if $\{y_1, y_2\}$ are a fundamental set of solutions to the associated homogeneous equation, then there is a particular solution of the form

$$y_p = u_1 y_1 + u_2 y_2.$$

Furthermore, there is an integral formula for u_1 and u_2.

Now consider a first order system of equations $\underline{x}' = P\underline{x} + \underline{g}$, for P in $M_n(\mathbb{C})$. Suppose that we have already solved the associated homogeneous equation $\underline{x}' = P\underline{x}$, and let $\{\underline{x}_1, \ldots, \underline{x}_n\}$ be a fundamental set of (homogeneous) solutions. Fix the corresponding fundamental matrix $\Psi(t)$. As we know, a general homogeneous solution has the form

$$\underline{x}_h = \Psi(t) \cdot \underline{c}, \qquad \underline{c} \text{ in } \mathbb{C}^n.$$

Analogously to the second-order equation above, we *guess* that a particular solution exists of the form

$$\underline{x}_p = \Psi \cdot \underline{u}(t).$$

We check (compute):

$$\begin{aligned} \underline{x}_p' &= \Psi' \underline{u} + \Psi \underline{u}' \\ P\underline{x}_p + \underline{g} &= P\Psi\underline{u} + \underline{g}. \end{aligned}$$

Then $\underline{x}_p$ is a particular solution if and only if $\Psi'\underline{u} + \Psi\underline{u}' = P\Psi\underline{u} + \underline{g}$. You can check that $\Psi' = P\Psi$ precisely because each of its columns is a solution to $\underline{x}' = P\underline{x}$. Then $\underline{x}_p$ is a particular solution if and only if $\underline{u}$ satisfies

$$\Psi\underline{u}' = \underline{g} \qquad \Leftrightarrow \qquad \underline{u}(t) = \int \Psi^{-1}\underline{g}dt;$$

Since $\underline{x}_p = \Psi\underline{u}$, we see that

$$\underline{x}_p = \Psi \cdot \int \Psi^{-1}\underline{g}dt$$

is a particular solution to the equation.

Extra Credit. (2 points) By converting the second-order equation

$$y'' + p(t)y' + q(t)y = g(t),$$

into a non-homogeneous first-order system in two variables, show that the new variation of parameters formula above produces the old variation of parameters formulas from math 307.

Example. We will solve the same example

$$\underline{x}' = \begin{pmatrix} 5 & 4 \\ -6 & -5 \end{pmatrix} \underline{x} + \begin{pmatrix} e^{2t} \\ 2e^{2t} + t \end{pmatrix}$$

using the variation of parameters formula.

We first find a fundamental matrix for the associated homogeneous equation. Since the coefficient matrix $P = \begin{pmatrix} 5 & 4 \\ -6 & -5 \end{pmatrix}$ has two distinct eigenvalues $\lambda = \pm 1$, with eigenspaces

$$E_{+1} = span\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}, \qquad E_{-1} = span\left\{ \begin{pmatrix} 2 \\ -3 \end{pmatrix} \right\},$$

we know two solutions $\underline{x}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^t$ and $\underline{x}_2 = \begin{pmatrix} 2 \\ -3 \end{pmatrix} e^{-t}$. So a fundamental matrix for the system is given by

$$\Psi(t) = \begin{pmatrix} e^t & 2e^{-t} \\ -e^t & -3e^{-t} \end{pmatrix}.$$

Using the variation of parameters formula, a particular solution is given by

$$\underline{x}_p = \Psi \int \Psi^{-1} \underline{g} dt.$$

We first compute that $\Psi^{-1} = \begin{pmatrix} 3e^{-t} & 2e^{-t} \\ -e^t & -e^t \end{pmatrix}$, so that

$$\Psi^{-1}\underline{g} = \begin{pmatrix} 3e^{-t} & 2e^{-t} \\ -e^t & -e^t \end{pmatrix} \begin{pmatrix} e^{2t} \\ 2e^{2t} + t \end{pmatrix} = \begin{pmatrix} 7e^t + 2te^{-t} \\ -3e^{3t} - te^t \end{pmatrix}.$$

So

$$\int \Psi^{-1} \underline{g} dt = \begin{pmatrix} 7e^t - 2(t+1)e^{-t} \\ -e^{3t} - (t-1)e^t \end{pmatrix}.$$

Finally,

$$\begin{aligned} \underline{x}_p = \Psi \int \Psi^{-1} \underline{g} dt &= \begin{pmatrix} e^t & 2e^{-t} \\ -e^t & -3e^{-t} \end{pmatrix} \begin{pmatrix} 7e^t - 2(t+1)e^{-t} \\ -e^{3t} - (t-1)e^t \end{pmatrix} \\ &= \begin{pmatrix} 7e^{2t} - 2(t+1) - 2(t-1) - 2e^{2t} \\ -7e^{2t} + 2(t+1) + 3e^{2t} + 3(t-1) \end{pmatrix} \\ &= \begin{pmatrix} 5 \\ -4 \end{pmatrix} e^{2t} + \begin{pmatrix} -4 \\ 5 \end{pmatrix} t + \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \end{aligned}$$

Finally, $\underline{x} = \Psi(t)\underline{c} + \underline{x}_p$ is the general solution. Observe that this is the same solution as given before. □

8.1 Remark. The variation of parameters method has the advantage that it works for every coefficient matrix P, even those that are defective. In fact, it even works for coefficient matrices which are not constant – provided that we can somehow solve the associated homogeneous equation first.

1.A Real Eigenvalues and the Spectral Theorem

In this section, we prove that a real symmetric matrix has real eigenvalues. We actually prove a more general theorem from which this fact will follow easily. This is an important result in many areas of mathematics, but it is also important for us. For example, when the eigenvalues of a coefficient matrix are real-valued (rather than being e.g. complex conjugates), the trajectories of solutions are much easier to understand.

1.A.1 Definition. Consider a matrix A in $M_n(\mathbb{C})$.

i) A is *self-adjoint* if $A = A^*$;

ii) A is *unitary* if $A^*A = I_n = AA^*$;

iii) A is *normal* if $A^*A = AA^*$.

Notice that if A is unitary, then $A^{-1} = A^*$. In this special case, computing the inverse matrix is easy! It follows immediately from the definition that A is unitary if and only if the columns of A form an orthonormal basis for $\mathbb{C}^n$.

For the rest of this document, suppose that we have fixed a matrix A in $M_n(\mathbb{C})$, and let $V = \mathbb{C}^n$ be the vector-space of columns of length n, with entries in $\mathbb{C}$.

1.A.2 Theorem (Schur-Decomposition)**.** *There exists a* unitary *matrix U such that*

$$T = U^{-1}AU = U^*AU$$

is an upper-triangular matrix over the complex numbers.

Remember that $T = (t_{ij})$ is upper-triangular means that $t_{ij} = 0$ whenever $i > j$. The matrix U in the theorem is not unique; indeed, in our proof we will need to make some arbitrary choices.

Proof of Theorem. The proof is by mathematical induction on the dimension n of V. Pick an eigenvalue λ of A with corresponding eigenspace E_λ, and geometric multiplicity $q_i = \dim(E_\lambda)$. Choose an orthonormal basis for E_λ, and extend this to an orthonormal basis $\beta = \{x_1, \ldots, x_{q_i}, y_{q_i+1}, \ldots, y_n\}$ for V. Define U_1 to be the matrix whose columns are the entries in β; then U_1 is unitary. Define $T_1 = U_1^*AU_1$; observe that

$$T_1 = \begin{pmatrix} \lambda \cdot I_{q_1} & T_{12} \\ T_{21} & T_{22} \end{pmatrix};$$

indeed, the ijth entry in the upper-left corner is $\lambda x_i^* x_j = (x_j, x_i) = \lambda\delta_{ij}$, as the x_i's are orthonormal. Similarly $T_{21} = 0$: the entries in T_{21} are given by $y_k^*(\lambda x_i) = \lambda(x_i, y_k)$; this is zero because distinct columns of U are orthogonal.

It follows that T_1 is "block upper-triangular." The square sub-matrix T_{22} is of dimension $(n - q_i) \times (n - q_i)$, so by induction we can find a unitary matrix U_2' such that $(U_2')^* T_{22} U_2'$ is upper-triangular. Then the matrix

$$U_2 := \begin{pmatrix} I_{q_1} & 0 \\ 0 & U_2' \end{pmatrix}$$

is also unitary, and

$$T := U_2^* T_1 U_2 = (U_1 U_2)^* A (U_1 U_2)$$

is upper-triangular, as desired. To complete the proof, we need only check that the product $U = U_1 U_2$ of two unitary matrices is again unitary. Indeed,

$$(U_1 U_2)^* (U_1 U_2) = U_2^* U_1^* U_1 U_2 = U_2^* I_n U_2 = U_2^* U_2 = I_n. \qquad \square$$

We now apply the Schur decomposition to get an easy form of the spectral theorem for a finite dimensional vector space.

1.A.3 Theorem (Spectral Theorem - Easy)**.** *Suppose that A is self-adjoint, i.e. $A = A^*$. Then*

i) There is a unitary matrix U which diagonalizes A.

ii) The eigenvalues of A are real.

iii) The columns of U are orthogonal eigenvectors of A of length 1.

Proof. Use the Schur decomposition to write

$$T = U^* A U,$$

where T is upper-triangular and U is unitary. By taking adjoints, we have

$$T^* = (U^* A U)^* = U^* A^* U^{**} = U^* A U = T.$$

Since T is upper-triangular, we know *a priori* that T^* is lower-triangular. Since $T^* = T$, T is both upper- and lower-triangular, and hence diagonal. This proves (i).

Since T is a diagonal matrix, and $T = T^*$, we conclude that the entries of T are all real. Since T is diagonal, its diagonal entries are precisely its eigenvalues. We prove that the eigenvalues of A are real by showing that T and A have the same eigenvalues. Indeed,

$$\begin{aligned} \det(T - \lambda I) &= \det(U^* U) \det(T - \lambda I) = \det(U^*) \det(T - \lambda I) \det(U) \\ &= \det(U^* (T - \lambda I) U) = \det(U^* T U - \lambda I) \\ &= \det(A - \lambda I), \end{aligned}$$

so T and A have the same characteristic equation and thus the same eigenvalues. This proves (ii).

To prove (iii), it is enough to show that the columns of U are eigenvectors - the other facts follow because U is unitary. We know that T is diagonal with diagonal entries consisting of the eigenvalues of A, say $\lambda_1, \ldots, \lambda_n$. Then if $U = (u_1, \ldots, u_n)$, we have

$$AU = UT \implies (Au_1, \ldots, Au_n) = (\lambda_1 u_1, \ldots, \lambda_n u_n),$$

which shows that each column u_i is an eigenvector with eigenvalue λ_i. □

1.A.4 Corollary. *If A is a real symmetric matrix, then A is diagonalizable and has real eigenvalues.*

Proof. If A is real and symmetric then $A^* = A^\top = A$ and the previous theorem applies. □

Extra Credit

Since every self-adjoint matrix is normal, the following exercise extends the previous theorem somewhat:

i) (2 points) Prove that if a matrix T is normal, and upper-triangular, then T is diagonal. *Hints*: T is normal if and only if $(r_i, r_j) = (c_i, c_j)$, where r_i is the ith row of T, and c_j is the jth column of T. What is (r_n, r_n)? What does this say about c_n?

ii) (2 points) Suppose A is a matrix and $T = U^*AU$ is the Schur decomposition of A. Prove that A is normal if and only if T is normal.

iii) (2 points) Use the previous parts to prove that A is diagonalizable by a unitary matrix U if and only if A is normal. This is a stronger form of the spectral theorem for finite dimensional vector spaces.

Chapter 2

Fourier Series and Partial Differential Equations

1 Boundary Value Problems

A boundary value problem is a differential equation (or possibly a system of them) together with certain values of the function specified. For example:

EXAMPLE. Solve the equation $y'' + y = 0$ subject to the boundary conditions

$$y(0) = 0, \qquad y\left(\frac{\pi}{2}\right) = 1.$$

Solution. The general solution to the differential equation is $y(t) = c_1 \cos(t) + c_2 \sin(t)$. The first equation, $y(0) = 0$, implies that

$$0 = y(0) = c_1 + 0 = c_1.$$

The second condition implies that

$$1 = c_2 \sin\left(\frac{\pi}{2}\right) = c_2,$$

so $c_2 = 1$. Thus, the *unique* solution to the differential equation and the two *boundary conditions* is $y(t) = \sin(t)$. □

Suppose that we replaced $y\left(\frac{\pi}{2}\right) = 1$ with $y(\pi) = 0$. Then the solution $y(t) = c_2 \sin(t)$ would have worked for every choice of coefficient c_2, and so the solution would *not* be unique! It's also possible that the solutions need not exist:

EXAMPLE. Show that there is no solution to the equation $y''+y = 0$ subject to the boundary conditions

$$y(\pi/4) = 0, \qquad y(3\pi/4) = 1.$$ □

We say that the boundary conditions are *homogeneous* if for distinct values of time, $t_0 < t_1 < \cdots < t_n$, they are

$$y(t_0) = 0, \ldots, y(t_n) = 0.$$

1.1 Proposition. *Consider a homogeneous differential equation with homogeneous boundary conditions. Then the set of solutions forms a vector space under the usual point-wise addition of functions.*

Example. Solve the equation $y'' + \pi^2 y = 0$ subject to the homogeneous boundary conditions

$$y(0) = 0, \qquad y(1) = 0.$$

Solution. A general solution is $y(t) = c_1 \cos(\pi t) + c_2 \sin(\pi t)$, while the first boundary condition implies that $c_1 = 0$. Then the second boundary condition implies that

$$0 = y(1) = c_2 \sin(\pi \cdot 1) = 0.$$

So every solution to the homogeneous boundary value problem is of the form

$$y(t) = c \cdot \sin(\pi t);$$

in particular, there is not a unique solution. □

1.2 Theorem. *Consider the family of boundary value problems with fixed homogeneous boundary conditions $y(0) = 0 = y(L)$, where $L > 0$, and differential equation*

$$y'' + \lambda y = 0,$$

where λ is any complex number. For every λ, $y(0) = 0$ is a solution. The only complex numbers λ which admit non-zero solutions are those of the form

$$\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2, \qquad n \geq 1,$$

and the general solution for λ_n is

$$y(t) = c \cdot \sin\left(\frac{\pi}{L} n\, t\right).$$

Proof. The complex number λ admits a square-root μ in $\mathbb{C}$. Then the general solution to

$$y'' + \lambda y = y'' + \mu^2 y = 0$$

is given by $y(t) = c_1 e^{i\mu t} + c_2 e^{i\mu t}$. (Here, we just solve the characteristic equation $r^2 + \mu^2 = 0$ for r, and then we know that the functions of the form e^{rt} are solutions.) These fundamental solutions are complex-valued, but since the differential equation has complex coefficients, this is the best we can hope for.

The first boundary condition $y(0) = 0$ implies that $0 = c_1 + c_2$, so replace c_2 with $-c_1$. Then the general solution may be written as

$$y(t) = c_1 \left(e^{i\mu t} - e^{-i\mu t}\right).$$

We plug in the second boundary condition, and get $0 = y(L) = c_1 \left(e^{i\mu L} - e^{-i\mu L}\right)$; we may assume that $c_1 \neq 0$, since otherwise we would only get the trivial solution. In that case, we have

$$e^{i\mu L} = e^{-i\mu L} \qquad \Leftrightarrow \qquad e^{2i\mu L} = 1.$$

I claim that the only complex numbers z such that $e^z = 1$ are those of the form $z = 2\pi i \cdot n$, where n is an integer. Indeed, if $z = a + ib$, then $e^z = e^a (\cos(b) + i\sin(b)) = 1$ implies that $a = 0$, and $b = 2\pi n$. Thus $z = 2\pi i n$, as required.

We conclude that $2i\mu L = 2\pi i n$, so that

$$\mu = \frac{\pi}{L} \cdot n \qquad \Rightarrow \qquad \lambda = \mu^2 = \left(\frac{\pi}{L}\right)^2 n^2.$$

If $n = 0$, then $\lambda = 0$. But $y'' = 0$ has general solution $y = a \cdot t + b$; then the homogeneous boundary conditions imply that $y = 0$. The case $n < 0$ is no different from $n > 0$ since λ_n is only a function of n^2.

Finally, fixing $\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2$, $n \geq 1$, we know the general solution to $y'' + \lambda_n y = 0$ and $y(0) = 0$ is just

$$y(t) = c_2 \sin\left(\frac{\pi}{L} n\, t\right),$$

and the additional condition $y(L) = 0$ is already satisfied for every choice of constant, c_2. □

In the previous theorem, the numbers λ that admit non-zero solutions to the homogeneous boundary value problem are called the *eigenvalues* of the problem. The theorem says that the only eigenvalues are those numbers λ_n, which turn out to be positive real-numbers. Fixing a particular eigenvalue λ_n, a non-zero solutions to that boundary value problem is called an *eigenfunction* of the problem. The theorem says that the eigenfunctions corresponding to the eigenvalue λ_n are the multiples of $y_n = \sin\left(\frac{\pi}{L} n\, t\right)$.

Extra Credit (3 points) Find all pairs of complex numbers $(a, b) \in \mathbb{C}^2$ such that there is a non-zero solution to the equation

$$y'' + ay' + by = 0,$$

subject to the homogeneous boundary conditions $y(0) = 0 = y(L)$, for $L > 0$.

Hint: Consider the substitution $u(t) = e^{\frac{a}{2}t} y(t)$.

2 Fourier Series

2.1 Definition. Consider a vector space V over $\mathbb{C}$, with an inner-product $\langle -, - \rangle : V \times V \longrightarrow \mathbb{C}$. A subset $S \subset V$ of vectors is called an *orthogonal family* if for *distinct* elements s and s' in S,

$$\langle s, s' \rangle = 0.$$

The subset S is called an *orthonormal family* if, in addition to being an orthogonal family, every s in S satisfies

$$\langle s, s \rangle = 1.$$

2.2 Proposition. *An orthogonal family S in V is linearly independent.*

Proof. Consider a linear dependence relation $c_1 s_1 + \cdots + c_n s_n = 0$, for distinct elements $\{s_1, \ldots, s_n\} \subset S$. By taking the inner product of $c_1 s_1 + \cdots + c_n s_n$ with s_j, we get

$$\begin{aligned} 0 &= \langle c_1 s_1 + \cdots + c_n s_n, s_j \rangle \\ &= c_1 \langle s_1, s_j \rangle + \cdots + c_j \langle s_j, s_j \rangle + \cdots + c_n \langle s_n, s_j \rangle \\ &= c_j \langle s_j, s_j \rangle. \end{aligned}$$

We used orthogonality to conclude that $\langle s_i, s_j \rangle = 0$ when $i \neq j$. Since $\langle s_j, s_j \rangle > 0$ (this is required for inner-products), we conclude that $c_j = 0$. Since j was arbitrary, we have that every coefficient is 0, so that $\{s_1, \ldots, s_n\}$ is linearly independent. □

When $V = \mathbb{C}^n$, there is an algorithm called Gram-Schmidt orthogonalization which takes an (ordered) basis for V and returns an orthogonal basis for V. You probably studied this in Math 308. The major advantage of picking an orthogonal basis is found when expressing a particular vector as a linear combination of those basis elements. Computing the coefficients is very easy:

2.3 Proposition. *Suppose V is a vector space, and $\mathcal{B}$ is an orthogonal family. Then every vector v in the span of $\mathcal{B}$ can be written as*

$$v = \sum_{b \text{ in } \mathcal{B}} c_b \cdot b,$$

where c_b is the scalar given explicitly by

$$c_b = \frac{\langle v, b \rangle}{\langle b, b \rangle}.$$

Proof. Since v is in the span of $\mathcal{B}$ it is possible to write

$$v = \sum_{b \text{ in } \mathcal{B}} c_b \cdot b,$$

for some scalars c_b. We want to prove that the scalars are given by the explicit formula above.

Compute the inner product of v with some particular basis element b' in $\mathcal{B}$:

$$\langle v, b' \rangle = \sum_{b \text{ in } \mathcal{B}} c_b \langle b, b' \rangle.$$

We have used linearity of the inner-product in the first variable. Since $\mathcal{B}$ is an orthogonal family, $\langle b, b' \rangle = 0$ when $b \neq b'$, so

$$\langle v, b' \rangle = c_b \langle b, b \rangle;$$

thus $c_b = \frac{\langle v, b' \rangle}{\langle b, b \rangle}$. □

The conceptual leap for us is the following: we consider the (infinite-dimensional) vector space $\mathcal{F}$ consisting of complex-valued functions $f(t)$ defined on $[-L, L]$, such that $f(t)$ and $f'(t)$ are piece-wise continuous. We define an inner-product on $\mathcal{F}$ by

$$\langle f(t), g(t) \rangle = \int_{-L}^{L} f(t)\overline{g(t)} dt.$$

It's an easy exercise to check that this satisfies the conditions of an inner-product, making $\mathcal{F}$ an inner-product space.

2.4 Theorem. *The subset*

$$\mathcal{B} = \left\{ 1, \cos\left(\frac{\pi}{L} m\, t\right), \sin\left(\frac{\pi}{L} n\, t\right) \mid m, n \textit{ are positive integers} \right\} \subset \mathcal{F}$$

is an orthogonal family. Furthermore,

$$\begin{aligned} \langle 1, 1 \rangle &= 2L \\ \langle \cos\left(\frac{\pi}{L} m\, t\right), \cos\left(\frac{\pi}{L} m\, t\right) \rangle &= L \\ \langle \sin\left(\frac{\pi}{L} n\, t\right), \sin\left(\frac{\pi}{L} n\, t\right) \rangle &= L. \end{aligned}$$

Proof. Exercise. Use the formulas:

$$\begin{aligned} \cos(A)\cos(B) &= \frac{\cos(A-B) + \cos(A+B)}{2} \\ \sin(A)\sin(B) &= \frac{\cos(A-B) - \cos(A+B)}{2} \\ \sin(A)\cos(B) &= \frac{\sin(A-B) + \sin(A+B)}{2}. \end{aligned}$$

□

We can consider a function in the span of $\mathcal{B}$: it is a function of the form

$$f(t) = c \cdot 1 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\, t\right) + b_n \sin\left(\frac{\pi}{L} n\, t\right),$$

for some choice of scalars c and a_n, b_n for $n \geq 1$. Since $f(t)$ is assumed to be in $\mathcal{F}$, this infinite sum converges. (However, if we just pick random scalars for the right-hand side, that infinite sum may not converge.) In particular, each coefficient is given by taking inner-products:

$$\begin{aligned} c &= \frac{\langle f, 1\rangle}{\langle 1, 1\rangle} = \frac{1}{2L}\int_{-L}^{L} f(t)dt \\ a_n &= \frac{\langle f, \cos\left(\frac{\pi}{L}n\,t\right)\rangle}{\langle \cos\left(\frac{\pi}{L}n\,t\right), \cos\left(\frac{\pi}{L}n\,t\right)\rangle} = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{\pi}{L}n\,t\right)dt \\ b_n &= \frac{\langle f, \sin\left(\frac{\pi}{L}n\,t\right)\rangle}{\langle \sin\left(\frac{\pi}{L}n\,t\right), \sin\left(\frac{\pi}{L}n\,t\right)\rangle} = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{\pi}{L}n\,t\right)dt. \end{aligned}$$

If we define $a_0 := 2c$, then we have

$$a_0 = 2c = \frac{1}{L}\int_{-L}^{L} f(t)dt = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{\pi}{L}\cdot 0\,t\right)dt,$$

which is consistent with the other coefficients a_n. We will do this going forward, but *remember* that this changes the expansion:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos\left(\frac{\pi}{L}n\,t\right) + b_n\sin\left(\frac{\pi}{L}n\,t\right).$$

2.5 Definition. Consider $f(t)$ in $\mathcal{F}$; if $f(t)$ is in the span of $\mathcal{B}$, then the (unique) expression

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos\left(\frac{\pi}{L}n\,t\right) + b_n\sin\left(\frac{\pi}{L}n\,t\right)$$

of $f(t)$ with respect to $\mathcal{B}$ is called the *Fourier series* of the function $f(t)$.

EXAMPLE. Assume that $f(t) = |t|$ is in the span of $\mathcal{B}$. Compute its Fourier series.

Solution: We need to determine all of the coefficients of the Fourier series:

$$
\begin{aligned}
a_0 &= \frac{1}{L}\int_{-L}^{L} |t| dt = L. \\
a_n &= \frac{1}{L}\int_{-L}^{L} |t| \cos\left(\frac{\pi}{L} n\, t\right) dt \\
&= \frac{2}{L}\int_{0}^{L} t \cos\left(\frac{\pi}{L} n\, t\right) dt \\
&= \frac{2}{L}\left(t \sin\left(\frac{\pi}{L} n\, t\right) \frac{L}{\pi n}\bigg|_{t=0}^{t=L} - \int_0^L \frac{L}{\pi n} \sin\left(\frac{\pi}{L} n\, t\right) dt \right) \\
&= \frac{2L}{(\pi n)^2}\left(\cos\left(\frac{\pi}{L} n\, L\right) - \cos\left(\frac{\pi}{L} n\, 0\right)\right) \\
&= \frac{2L}{(\pi n)^2}\left((-1)^n - 1\right) = \begin{cases} \frac{-4L}{(\pi n)^2} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}. \\
b_n &= \frac{1}{L}\int_{-L}^{L} |t| \sin\left(\frac{\pi}{L} n\, t\right) dt \\
&= 0
\end{aligned}
$$

So the Fourier series of $|t|$ is

$$
\begin{aligned}
|t| &= \frac{L}{2} - \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{4L}{(\pi n)^2} \cos\left(\frac{\pi}{L} n\, t\right) \\
&= \frac{L}{2} - \sum_{k=0}^{\infty} \frac{4L}{\pi^2 (2k+1)^2} \cos\left(\frac{\pi}{L}(2k+1)\, t\right).
\end{aligned}
$$

□

2.6 Definition. A function $f(t)$ defined on all of $\mathbb{R}$ is called *periodic* if there exists a number $T > 0$, called a *period* of $f(t)$, such that

$$f(t+T) = f(t) \qquad \text{for every } t \in \mathbb{R}.$$

If $f(t)$ is a periodic function with a smallest period, T_0, then T_0 is called the *fundamental period* of $f(t)$.

For example, the constant functions are periodic since any positive number is a period. However, there is no fundamental period.

Exercise. Suppose that $f(t)$ has a fundamental period T_0. Prove that every period T of $f(t)$ is of the form

$$T = n \cdot T_0, \qquad n \text{ an integer, } n \geq 1.$$

□

If $f(t)$ is periodic with period $2T$, we may restrict $f(t)$ to the interval $[-T, T]$ to get a function in $\mathcal{F}$ (letting $L = T$). Conversely, any function defined on $[-T, T]$ can be extended to an "almost periodic" function of period $2T$ on all of $\mathbb{R}$ by

$$f(t + 2T) = f(t).$$

An issue arises when $f(-T) \neq f(T)$. For example, if we extend $f(t) = t$ periodically from the interval $[-1, 1]$, we have

$$1 = f(1) = f(1 + 2) = f(3) = f(-1 + 2 + 2) = f(-1) = -1.$$

To avoid this problem, when we want to periodically extend a function $f(t)$ from $[-T, T]$ to all of $\mathbb{R}$, we do the following: redefine $f(T)$ and $f(-T)$ to be the average $\frac{f(-T)+f(T)}{2}$. This only changes the values of $f(t)$ at two points. In particular, it doesn't change the value of any of the coefficients in the Fourier series for $f(t)$, since the integrals

$$\frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{\pi}{L} n\, t\right) dt$$

have the same value if we change $f(t)$ on only finitely many points.

2.7 Proposition. *If f and g are periodic with period $2T$, then $f+g$, $f \cdot g$, f', are all periodic of period $2T$. If $\int_0^{2T} f(s)ds = 0$, then the function $h(t) = \int_0^t f(s)ds$ is also periodic of period $2T$.*

Proof. Exercise. □

We use the following notation for a function $f(t)$,

$$f(t_-) = \lim_{\substack{s \longrightarrow t \\ s<t}} f(s) \qquad f(t_+) = \lim_{\substack{s \longrightarrow t \\ s>t}} f(s);$$

this is the usual "limit from the left" and "limit from the right."

2.8 Definition. Consider $f(t)$ in $\mathcal{F}$. Define the function $f_{av}(t)$ by

$$f_{av}(t) := \frac{f(t_-) + f(t_+)}{2}, \qquad f_{av}(L) := \frac{f(-L) + f(L)}{2} =: f_{av}(-L).$$

It is important to notice that $f(t) = f_{av}(t)$ except at finitely many points.

2.9 Theorem (Fourier Convergence Theorem)**.** *Consider $f(t)$ in $\mathcal{F}$. Then $f_{av}(t)$ is in the span of $\mathcal{B}$, and the periodic extension of $f_{av}(t)$ to all of $\mathbb{R}$ agrees with the Fourier series at every point:*

$$f_{av}(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\, t\right) + b_n \sin\left(\frac{\pi}{L} n\, t\right), \qquad \textit{for every } t \in \mathbb{R}.$$

The Fourier coefficients for $f(t)$ are the same as those for $f_{av}(t)$.

The previous theorem shows that every function $f(t)$ in $\mathcal{F}$ has Fourier coefficients such that the resulting Fourier series is convergent for every t. We extend the definition of Fourier series to all functions $f(t)$ in $\mathcal{F}$ (not just those in the span of $\mathcal{B}$). The point is that $f(t)$ is not necessarily in the span of $\mathcal{B}$, because the Fourier series may not converge to $f(t)$ at points t where f is discontinuous.

Example. Consider the "square wave" function defined on $[-L, L]$ by

$$f(t) = \begin{cases} 0 & \text{if } -L \leq t < 0 \\ L & \text{if } 0 \leq t \leq L. \end{cases}$$

Compute $f_{av}(t)$, and then compute the Fourier series of f_{av}, and check that the Fourier series at $0, \pm L$ agrees with f_{av}.

Solution: The only points of discontinuity are the endpoints, and $t = 0$. At the endpoints, the average is $(0 + L)/2 = L/2$, which is the same as the average of the limit from the left and right at $t = 0$. So

$$f_{av}(t) = \begin{cases} 0 & \text{if } -L < t < 0, \\ L & \text{if } 0 < t < L, \\ \frac{L}{2} & \text{if } t = -L, 0, +L. \end{cases}$$

To compute the Fourier series, we integrate:

$$\begin{aligned} a_0 &= \frac{1}{L} \int_{-L}^{L} f(t) dt = \frac{1}{L} \int_0^L L dt = L; \\ a_n &= \frac{1}{L} \int_{-L}^{L} f(t) \cos\left(\frac{\pi}{L} n\, t\right) dt = \int_0^L \cos\left(\frac{\pi}{L} n\, t\right) dt = 0; \\ b_n &= \frac{1}{L} \int_{-L}^{L} f(t) \sin\left(\frac{\pi}{L} n\, t\right) dt = \int_0^L \sin\left(\frac{\pi}{L} n\, t\right) dt \\ &= \frac{-L}{\pi n} \cos\left(\frac{\pi}{L} n\, t\right)\Big|_{t=0}^{t=L} \\ &= \frac{L}{\pi n} (1 - (-1)^n) = \begin{cases} \frac{2L}{\pi n} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even.} \end{cases} \end{aligned}$$

Thus, the Fourier series is

$$f_{av}(t) = \frac{L}{2} + \frac{2L}{\pi} \sum_{\substack{n=1 \\ n \text{ is odd}}}^{\infty} \frac{1}{n} \sin\left(\frac{\pi}{L} n\, t\right) = \frac{L}{2} + \frac{2L}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \sin\left(\frac{\pi}{L}(2k+1)\, t\right).$$

If we evaluate the right-hand side at 0 or $\pm L$, we get $\frac{L}{2}$, as expected. □

3 Even and odd functions

3.1 Definition. Let $f : \mathbb{R} \longrightarrow \mathbb{C}$ be a complex-valued function on the real line.

i) The function f is *even* if $f(-t) = f(t)$ for every t in $\mathbb{R}$;

ii) The function f is *odd* if $f(-t) = -f(t)$ for every t in $\mathbb{R}$.

Some functions are neither even nor odd. For example, $f(t) = 1 + t$ is not even since $f(-1) = 0 \neq 2 = f(1)$, and it is not odd since $f(-1) = 0 \neq -2 = -f(1)$. However, we have the following proposition:

3.2 Proposition. *Every function $f(t)$ can be uniquely written as*

$$f(t) = f_{ev}(t) + f_{odd}(t),$$

the sum of an even and odd function.

Proof. Define

$$f_{ev}(t) := \frac{f(t) + f(-t)}{2}, \qquad \text{and} \qquad f_{odd}(t) := \frac{f(t) - f(-t)}{2}.$$

Then it is easy to check that $f_{ev}(t)$ is even and $f_{odd}(t)$ is odd. Furthermore,

$$f(t) = \frac{f(t) + f(-t)}{2} + \frac{f(t) - f(-t)}{2} = f_{ev}(t) + f_{odd}(t).$$

It remains to prove uniqueness. So suppose that we can write $f(t) = e(t) + o(t)$, for some (potentially different) even function $e(t)$ and odd function $o(t)$. Then

$$0 = f(t) - f(t) = (f_{ev}(t) + f_{odd}(t)) - (e(t) + o(t)) = (f_{ev}(t) - e(t)) + (f_{odd}(t) - o(t)).$$

This must hold for every t in $\mathbb{R}$. In particular, evaluating at t and $-t$ we get the two equations:

$$\begin{aligned} 0 &= (f_{ev}(t) - e(t)) + (f_{odd}(t) - o(t)) \\ 0 &= (f_{ev}(t) - e(t)) - (f_{odd}(t) - o(t)). \end{aligned}$$

By adding, we get that $2(f_{ev}(t) - e(t)) = 0$, so $e(t) = f_{ev}(t)$; by subtracting we get $2(f_{odd}(t) - o(t)) = 0$, so $f_{odd}(t) = o(t)$. This completes the proof of uniqueness. $\square$

Here are some easy facts to check: the set of even (resp. odd) functions forms a vector space; the product of even functions is even, and the product of odd functions is even; the product of an odd and an even function is odd. Quite importantly, for us, is how even and odd functions behave with respect to integration over a symmetric interval:

3.3 Proposition. *Let $f : \mathbb{R} \longrightarrow \mathbb{C}$ be a continuous function, and let*

$$I := \int_{-L}^{L} f(t)dt.$$

If $f(t)$ is even, then $I = 2\int_0^L f(t)dt$. If $f(t)$ is odd, then $I = 0$.

Proof. Exercise. □

3.4 Proposition. *Suppose that $f(t)$ is in $\mathcal{F}$, with Fourier series*

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L}n\,t\right) + b_n \sin\left(\frac{\pi}{L}n\,t\right).$$

i) If $f(t)$ is even, then $b_n = 0$ for every $n \geq 1$. Then

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L}n\,t\right)$$

is the Fourier series of $f(t)$, and it is called the Fourier cosine series *of $f(t)$.*

ii) If $f(t)$ is odd, then $a_n = 0$ for every $n \geq 0$. Then

$$f(t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,t\right)$$

is the Fourier series of $f(t)$, and is called the Fourier sine series *of $f(t)$.*

Proof. i) The formula for b_n is

$$b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{\pi}{L}n\,t\right)dt;$$

since $f(t)$ is even and $\sin(\left(\frac{\pi}{L}n\,t\right))$ is odd, their product is odd and thus the integral is 0 by the previous proposition.

ii) The formula for a_n, $n \geq 0$ is

$$a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{\pi}{L}n\,t\right)dt;$$

since $f(t)$ is odd, and $\cos\left(\frac{\pi}{L}n\,t\right)$ is even, their product is odd and thus the integral is 0. □

Evaluating some infinite sums*

It is possible to use Fourier series to evaluate tricky infinite sums. We present a method to compute the exact value of $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

3.5 Proposition. *Suppose that $f(t)$ in $\mathcal{F}$ is an odd function. Then $f(t)$ has Fourier series*

$$f(t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L} n\, t\right),$$

and the coefficients b_n satisfy

$$\sum_{n=1}^{\infty} b_n^2 = \frac{2}{L} \int_0^L (f(t))^2\, dt.$$

Proof. Observe that

$$\frac{2}{L} \int_0^L (f(t))^2\, dt = \frac{1}{L} \int_{-L}^{L} f(t) \cdot f(t) dt = \frac{1}{L} \langle f(t), f(t) \rangle.$$

By substituting $f(t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L} n\, t\right)$, and using orthogonality of $\mathcal{B}$, we get

$$\begin{aligned}
\frac{2}{L} \int_0^L (f(t))^2\, dt &= \frac{1}{L} \langle \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L} n\, t\right), \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L} n\, t\right) \rangle \\
&= \frac{1}{L} \sum_{n=1}^{\infty} b_n^2 \langle \sin\left(\frac{\pi}{L} n\, t\right), \sin\left(\frac{\pi}{L} n\, t\right) \rangle \\
&= \sum_{n=1}^{\infty} b_n^2.
\end{aligned}$$

□

3.6 Theorem.

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Proof. In a worksheet, we computed the Fourier series for $f(x) = x$. We got

$$f(x) = -\frac{2L}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{\pi}{L} n\, t\right).$$

We apply the previous proposition to this function, and obtain that

$$\left(\frac{2L}{\pi}\right)^2 \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{2}{L} \int_0^L x^2 dx.$$

Of course, the integral on the right is easy to do: $\frac{2}{L} \int_0^L x^2 dx = \frac{2}{L} \cdot \frac{L^3}{3} = \frac{2}{3} L^2$. It follows that

$$\frac{4L^2}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{2}{3} L^2 \qquad \Rightarrow \qquad \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

□

Even and Odd Extensions

Suppose that $f(t)$ is a piece-wise continuous function, with piece-wise continuous derivative, but $f(t)$ is only defined on the interval $(0, L]$. We can extend the function to the interval $[-L, L]$ so that the extension is either even or odd:

$$\begin{aligned} \text{even extension: } f(-t) &= f(t), \qquad t \text{ in } (0, L]; \\ \text{odd extension: } f(-t) &= -f(t), \qquad t \text{ in } (0, L]. \end{aligned}$$

This is not the only way to extend the function to $[-L, L]$, but these are two important ways. If we employ the even extension, then $f(t)$ becomes an *even* function in $\mathcal{F}$ and so it has a Fourier cosine series (i.e. its Fourier series only involves the cosine terms and the constant). If instead we empoy the odd extension, then $f(t)$ becomes an *odd* function in $\mathcal{F}$, and so it has a Fourier sine series (i.e. its Fourier series only involves the sine terms).

Example. Consider the function $f(x) = \begin{cases} x, & 0 \le x < 1, \\ 1, & 1 \le x < 2. \end{cases}$ Find the Fourier sine series of $f(t)$ with period 4.

Solution: Since we want a Fourier sine series with period $4 = 2L$, we must have that $L = 2$. We are given a function $f(t)$ defined only on the interval $[0, 2]$; we extend this function to be an *odd* function on $[-2, 2]$ since we are looking for the Fourier *sine* series. So the extended function is defined as

$$f(x) = \begin{cases} -1, & -2 \le x, \le -1 \\ x, & -1 < x \le 1 \\ 1, & 1 \le x < 2. \end{cases}$$

We know, a priori, that all of the terms a_n in the Fourier series of $f(x)$ will be 0 since $f(x)$ is odd. Similarly, we know that

$$b_n = \frac{1}{2} \int_{-2}^{2} f(x) \sin\left(\frac{\pi}{2} n\, x\right) dx = \frac{2}{2} \int_{0}^{2} f(x) \sin\left(\frac{\pi}{2} n\, x\right) dx.$$

So it remains to compute this integral. Since $f(x)$ is defined piece-wise, we break up the integral as

$$\frac{2}{2} \int_{0}^{2} f(x) \sin\left(\frac{\pi}{2} n\, x\right) dx = \int_{0}^{1} x \sin\left(\frac{\pi}{2} n\, x\right) dx + \int_{1}^{2} \sin\left(\frac{\pi}{2} n\, x\right) dx.$$

Use integration by parts to compute that

$$\begin{aligned} \int_{0}^{1} x \sin\left(\frac{\pi}{2} n\, x\right) dx &= \frac{4}{\pi^2 n^2} \sin\left(\frac{\pi}{2} n\right) - \frac{2}{\pi n} \cos\left(\frac{\pi}{2} n\right) \\ \int_{1}^{2} \sin\left(\frac{\pi}{2} n\, x\right) dx &= \frac{2}{\pi n} \left(\cos\left(\frac{\pi}{2} n\right) - \cos(\pi n)\right). \end{aligned}$$

Adding these together we get $b_n = \frac{4}{\pi^2 n^2}\sin\left(\frac{\pi}{2}n\right) - \frac{2}{\pi n}(-1)^n$. This is already an acceptable answer. The first few terms of the Fourier sine series are

$$f(x) = \frac{2}{\pi^2}(2+\pi)\sin\left(\frac{\pi}{2}\,x\right) - \frac{1}{\pi}\sin(\pi\,x) + \frac{2}{9\pi^2}(\pi-2)\sin\left(\frac{3\pi}{2}\,x\right) - \frac{1}{2\pi}\sin(2\pi\,x) + \cdots$$

but it is a little bit awkward to write down a closed-form summation that does not involve $\sin\left(\frac{\pi}{2}n\right)$. However, I will show you how to do this in case you want to know.

By division with remainder, we can uniquely write $n = 4k + r$, where $1 \le r \le 4$ (*NOTE: this is not the usual way of using division with remainder, since we usually require* $0 \le r \le 3$). Then

$$\sin\left(\frac{\pi}{2}n\right) = \begin{cases} 0 & \text{if } r = 2, 4 \\ 1 & \text{if } r = 1 \\ -1 & \text{if } r = 3. \end{cases}$$

It follows that

$$b_n = \begin{cases} -\frac{2}{\pi n} & n = 4k+2, 4k+4 \\ \frac{4}{\pi^2 n^2} + \frac{2}{\pi n} & n = 4k+1 \\ \frac{-4}{\pi^2 n^2} + \frac{2}{\pi n} & n = 4k+3. \end{cases}$$

Since the Fourier sine series begins with $n = 1$, we can begin with $k = 0$, and sum four terms at-a-time:

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} b_{4k+1}\sin\left(\frac{\pi}{2}(4k+1)\,x\right) + b_{4k+2}\sin\left(\frac{\pi}{2}(4k+2)\,x\right) \\ &\quad + b_{4k+3}\sin\left(\frac{\pi}{2}(4k+3)\,x\right) + b_{4k+4}\sin\left(\frac{\pi}{2}(4k+4)\,x\right) \\ &= \sum_{k=0}^{\infty}\left(\frac{4}{\pi^2(4k+1)^2} + \frac{2}{\pi(4k+1)}\right)\sin\left(\frac{\pi}{2}(4k+1)\,x\right) - \left(\frac{2}{\pi(4k+2)}\right)\sin\left(\frac{\pi}{2}(4k+2)\,x\right) \\ &\quad + \left(\frac{-4}{\pi^2(4k+3)^2} + \frac{2}{\pi(4k+3)}\right)\sin\left(\frac{\pi}{2}(4k+3)\,t\right) - \frac{2}{\pi(4k+4)}\sin\left(\frac{\pi}{2}(4k+4)\,t\right). \end{aligned}$$

This is hardly an improvement; it is better to just leave the solution in terms of $\sin\left(\frac{\pi}{2}n\right)$. □

Specialized Extensions

We present here two special extensions of a function from the interval $[0, L]$ to the interval $[-2L, 2L]$. First we present a "special odd" version, and then a "special even" version. Suppose throughout that $f(t)$ is a piece-wise continuous function, with piece-wise continuous derivative, originally defined on $(0, L]$.

For the special odd extension, we first extend $f(t)$ to $(0, 2L]$ by *reflecting the graph of $f(t)$ across the vertical line $t = L$*. In otherwords, we extend $f(t)$ for $L \le t \le 2L$ by the formula

$$f(t) = f(L - (t - L)) = f(2L - t).$$

To define $f(t)$ on the symmetric interval $[-2L, 2L]$, we must extend again; we choose the odd extension from above:

$$\text{odd extension: } f(-t) = -f(t), \qquad t \text{ in } (0, 2L].$$

Our goal is to compute the Fourier coefficients in this case. Since $f(t)$ is odd on $[-2L, 2L]$, it is true that $a_n = 0$ for all $n \ge 0$. We compute b_n:

$$\begin{aligned} b_n &= \frac{2}{2L}\int_0^{2L} f(t)\sin\left(\frac{\pi}{2L}n\,t\right)dt = \frac{1}{L}\int_0^{L} f(t)\sin\left(\frac{\pi}{2L}n\,t\right)dt + \frac{1}{L}\int_L^{2L} f(t)\sin\left(\frac{\pi}{2L}n\,t\right)dt \\ &= \frac{1}{L}\int_0^{L} f(t)\sin\left(\frac{\pi}{2L}n\,t\right)dt + \frac{1}{L}\int_{t=L}^{t=2L} f(2L-t)\sin\left(\frac{\pi}{2L}n\,t\right)dt. \end{aligned}$$

The second integral suggests we try the substitution $u = 2L - t$ (so $du = -dt$); we obtain then

$$\begin{aligned} b_n &= \frac{1}{L}\int_0^{L} f(t)\sin\left(\frac{\pi}{2L}n\,t\right)dt - \frac{1}{L}\int_{u=L}^{u=0} f(u)\sin\left(\frac{\pi}{2L}n\,(2L-u)\right)du \\ &= \frac{1}{L}\int_0^{L} f(t)\sin\left(\frac{\pi}{2L}n\,t\right)dt + \frac{1}{L}\int_{u=0}^{u=L} f(u)\sin\left(\pi n - \frac{\pi}{2L}n\,u\right)du \\ &= \frac{1}{L}\int_0^{L} f(t)\left(\sin\left(\frac{\pi}{2L}n\,t\right) - \sin\left(\frac{\pi}{2L}n\,t - \pi n\right)\right)dt \\ &= \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{2}{L}\int_0^L f(t)\sin\left(\frac{\pi}{2L}n\,t\right)dt & \text{if } n \text{ is odd.} \end{cases} \end{aligned}$$

In particular, the Fourier series for this specialized extension has the form

$$f(t) = \sum_{k=0}^{\infty} b_{2k+1}\sin\left(\frac{\pi}{L}\frac{2k+1}{2}\,t\right). \tag{2.1}$$

Note this form, as it will come up later in the chapter when we study PDEs.

The analogous specialized extension, the "even version" of the previous one, is next. First, extend $f(t)$ from $[0, L]$ to $[0, 2L]$ by anti-reflecting it over the line $x = L$, i.e. for $L < t \le 2L$, we have

$$f(t) = -f(2L - t).$$

Now extend $f(t)$ to all of $[-2L, 2L]$ by using the even extension: for $-2L < t \le 0$:

$$f(t) = f(-t).$$

Our goal is to compute the Fourier coefficients. Since $f(t)$ is even, we only need to compute the a_n terms. By symmetry, we know that $a_0 = 0$. Compute that

$$
\begin{aligned}
a_n &= \frac{2}{2L}\int_0^{2L} f(t)\cos\left(\frac{\pi}{2L}n\,t\right)dt = \frac{1}{L}\int_0^{L} f(t)\cos\left(\frac{\pi}{2L}n\,t\right)dt - \frac{1}{L}\int_L^{2L} f(2L-t)\cos\left(\frac{\pi}{2L}n\,t\right)dt \\
&= \frac{1}{L}\int_0^{L} f(t)\cos\left(\frac{\pi}{2L}n\,t\right)dt - \frac{1}{L}\int_{u=0}^{u=L} f(u)\cos\left(\pi n - \frac{\pi}{2L}n\,u\right)du \\
&= \frac{1}{L}\int_0^{L} f(t)\left[\cos\left(\frac{\pi}{2L}n\,t\right) - \cos\left(\frac{\pi}{2L}n\,t - \pi n\right)\right]dt \\
&= \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{2}{L}\int_0^L f(t)\cos\left(\frac{\pi}{2L}n\,t\right)dt & \text{if } n \text{ is odd.} \end{cases}
\end{aligned}
$$

In particular, the Fourier series for this specialized extension has the form

$$
f(t) = \sum_{k=0}^{\infty} a_{2k+1}\cos\left(\frac{\pi}{L}\frac{2k+1}{2}\,t\right) \tag{2.2}
$$

4 PDE's - Separation of Variables

Consider a function $u(x,t)$ of two independent variables x and t. Remember that the partial derivative of u with respect to (e.g.) x is

$$
\frac{\partial u}{\partial x}(x_0,t_0) := \lim_{h\longrightarrow 0}\frac{u(x_0+h,t_0)-u(x_0,t_0)}{h}.
$$

By fixing $t = t_0$ and considering $u(x,t_0)$ as a function of x alone, this is just the usual derivative at $x = x_0$. The partial derivative is again a function of the various variables, so one can take the partial derivative of a partial derivative, e.g.

$$
\frac{\partial}{\partial t}\frac{\partial}{\partial x}\frac{\partial u}{\partial t}.
$$

A partial differential equation is one involving functions of x and t, and various partial derivatives of u (of different orders) with respect to x and t. For example:

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

is a partial differential equation. We will often use the notation u_x for $\frac{\partial}{\partial x}u$; note that

$$
\frac{\partial}{\partial x}\frac{\partial}{\partial t}u = \frac{\partial}{\partial x}(u_t) = (u_t)_x = u_{tx};
$$

However, if both u_{tx} and u_{xt} are continuous functions, then they are in fact *equal*. This is "equality of mixed partials" that you first met in multi-variable calculus. So in that case, you don't have to worry about the order u_{tx} vs u_{xt}.

In this class we will almost always restrict our attention to PDEs of one variable, u, that is a function of two independent variables. They might be x and t, or x and y - the names are irrelevant.

Our main technique for solving PDEs is the method of "separation of variables." The point of the method is to "convert" a partial differential equation into a system of ordinary differential equations. We present this technique in a few examples.

Example. Convert the differential equation

$$u_{xx} + tu_{tt} = 0$$

into a system of ordinary differential equations using separation of variables.

Solution: First, observe that $u(x,t) = 0$ is a trivial solution. This is true of every linear PDE. Suppose that $u(x,t) = X(x)T(t) \neq 0$ can be written as the product of a function *only* of x with a function *only* of t. Then $u_x = X'(x)T(t)$, and similarly $u_t = X(x) \cdot T'(t)$. Similarly, $u_{xx} = X''(x)T(t)$, and $u_{tt} = X(x)T''(t)$. Thus,

$$u_{xx} + tu_{tt} = 0 \Rightarrow X''T + t \cdot XT'' = 0,$$

which can be rearranged as $X''T = -tXT''$. Finally, we *separate the variables*:

$$\frac{X''}{X} = -t\frac{T''}{T}.$$

We know that both $X(x)$ and $T(t)$ are not the zero-function, because $u(x,t) = X(x)T(t) \neq 0$, so the quotients make sense. Now stare at the previous equation for a second. The left-hand side is a function only of x; while the right-hand side is a function only of t. Hold $t = t_0$ fixed and let x vary. We see that no matter what x is,

$$\frac{X''(x)}{X(x)} = -t_0\frac{T''(t_0)}{T(t_0)} = \text{ a constant.}$$

Similarly, by holding $x = x_0$ fixed and letting t vary, we get that

$$\frac{X''(x_0)}{X(x_0)} = -t\frac{T''(t)}{T(t)} = \text{ a constant.}$$

Moreover, these two constants have to be the same! If we call that constant $-\lambda$, we obtain the system of ordinary differential equations

$$\begin{cases} X''(x) + \lambda X(x) = 0 \\ T''(t) - \lambda t \cdot T(t) = 0. \end{cases}$$ □

It is important to realize that λ in $\mathbb{C}$ is arbitrary. So the previous solution is really a parametrized family of systems of equations.

Example. Convert the following PDE

$$[p(x)u_x]_x + r(x)u_t = 0$$

into a system of ordinary differential equations.

Solution: Let $u(x,t) = X(x)T(t)$, as before. We use the product rule to re-write the PDE as

$$p'(x)u_x + p(x)u_{xx} + r(x)u_t = 0 \Rightarrow p'(x)X'(x)T(t) + p(x)X''(x)T(t) + r(x)X(x)T'(t) = 0.$$

We re-arrange terms to help separate variables:

$$[p'(x)X'(x) + p(x)X''(x)] \cdot T(t) = -r(x)X(x)T'(t),$$

and then we separate:

$$\frac{p'(x)X'(x) + p(x)X''(x)}{r(x)X(x)} = -\frac{T'(t)}{T(t)}.$$

Since the left-hand side is only a function of x, and the right-hand side is only a function of t, we know that both sides are equal to the same constant, which we call for convenience $-\lambda$. Then

$$\frac{p'(x)X'(x) + p(x)X''(x)}{r(x)X(x)} = -\lambda \qquad \Rightarrow \qquad p(x)X'' + p'(x)X'(x) + \lambda r(x)X(x) = 0$$

$$-\frac{T'(t)}{T(t)} = -\lambda \qquad \Rightarrow \qquad T'(t) - \lambda T(t) = 0.$$

So again, we find a (parametrized) system of ordinary differential equations. □

5 The Heat Equation

Consider a very thin rod of length L, placed with its left end at $x = 0$ and right end at $x = L$, and let $u(x,t)$ be the temperature of the rod at position x, at time t ($0 \le x \le L$, $t \ge 0$). We begin by deriving the heat equation - a PDE that our temperature function $u(x,t)$ must satisfy.

As always in modeling, we must start with something empirical. In this case, suppose that two bodies of the same cross-sectional area A are separated by a distance d. If u_1 and u_2 are their temperatures, then the rate of heat transfer from one to the other is proportional to

$$\frac{A}{d}|u_1 - u_2|,$$

and, of course, heat flows from the object of higher temperature to the object of lower temperature. We call the proportionality constant κ, the *thermal conductivity* of the material. If Q denotes heat energy, we have just said that

$$\frac{\Delta Q}{\Delta t} = \kappa \frac{A}{d}|u_1 - u_2|.$$

We begin analyzing the one-dimensional rod by fixing a point x_0, and considering a small slice of width Δx. We consider Δx small enough so that the "slice" of rod from $x = x_0$ to $x = x_0 + \Delta x$ is approximately the same temperature everywhere. Now, we compute how much heat energy enters this slice in a small time Δt. The heat transferred from the part of the rod to the left of the slice is given by

$$(\Delta Q)_{in} = \lim_{d \longrightarrow 0} \kappa \frac{A}{d} |u(x_0 - d, t) - u(x_0, t)| \Delta t = -\kappa A u_x(x_0, t) \Delta t.$$

Similarly, the heat that leaves the slice and goes to the part of the rod to the right is given by

$$(\Delta Q)_{out} = \lim_{d \longrightarrow 0} \kappa \frac{A}{d} |u(x_0 + \Delta x + d, t) - u(x_0 + \Delta x, t)| = -\kappa A u_x(x_0 + \Delta x, t) \Delta t.$$

Thus, the total increase in heat energy into the slice in time Δt is

$$\Delta Q = (\Delta Q)_{in} - (\Delta Q)_{out} = \kappa A (u_x(x_0 + \Delta x, t) - u_x(x_0, t)) \Delta t.$$

Now, we need to know how this increase in heat energy affects the temperature of the slice. We use the following relation:

$$\Delta Q = s \cdot m \cdot \Delta u,$$

where m is the mass of the object, s is its *specific heat*, and Δu is the change in temperature, i.e. $\Delta u = u(x_0, t + \Delta t) - u(x_0, t)$. This makes sense intuitively - more mass should be able to absorb more heat without rising as much in temperature. The constant s is just a proportionality constant that depends on the type of material (and also possibly the temperature, but we ignore this for simplicity). If the rod has constant density ρ, then the mass is ρ times the volume; the volume is just $\Delta x \cdot A$. So we get:

$$\Delta Q = s \rho A \Delta x (u(x_0, t + \Delta t) - u(x_0, t)) = \kappa A (u_x(x_0 + \Delta x, t) - u_x(x_0, t)) \Delta t.$$

Canceling, and re-arranging, we get

$$\frac{u(x_0, t + \Delta t) - u(x_0, t)}{\Delta t} = \frac{\kappa}{s\rho} \cdot \frac{u_x(x_0 + \Delta x, t) - u_x(x_0, t)}{\Delta x}.$$

Now, we take the limit as Δt and Δx go to 0:

$$u_t(x_0, t) = \frac{\kappa}{s\rho} u_{xx}(x_0, t).$$

It is easy to check that all of the constants κ, s, ρ are positive, so we can write $\frac{\kappa}{s\rho} =: \alpha^2$. This is the *Heat equation*:

$$\boxed{u_t = \alpha^2 u_{xx}.}$$

The rod with homogeneous boundary conditions

We model the temperature of the rod under the assumption that the endpoints of the rod are fixed at 0 temperature, i.e.

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

The initial temperature of the rod is some (infinitely differentiable) function, $f(x) = u(x,0)$. In otherwords, we want to find *all* solutions to the following system of equations:

$$\begin{cases} u_t = \alpha^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = f(x). \end{cases}$$

We use separation of variables to convert the PDE into a system of ordinary differential equations. So guess that $u(x,t) = X(x) \cdot T(t) \neq 0$, for some functions $X(x)$ and $T(t)$; then the differential equation $u_t = \alpha^2 u_{xx}$ implies that

$$X(x)T'(t) = \alpha^2 X''(x)T(t).$$

Now, *separate the variables*

$$\frac{1}{\alpha^2}\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} = const.$$

If we call that constant $-\lambda$, we obtain the system of equations

$$\begin{aligned} X''(x) + \lambda X(x) &= 0 \\ T'(t) + \lambda\alpha^2 T(t) &= 0. \end{aligned}$$

We must interpret the two boundary conditions on $u(x,t)$. Since $u(0,t) = X(0)T(t) = 0$, and $T(t) \neq 0$, we have that $X(0) = 0$. Similary, $X(L) = 0$. Thus, $X(x)$ must satisfy the boundary value problem (BVP)

$$\begin{cases} X''(x) + \lambda X(x) = 0 \\ X(0) = 0 = X(L). \end{cases}$$

We have solved this problem already; we know the only constants λ that admit non-zero solutions to this equation are $\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2$, for $n \geq 1$. The corresponding eigenfunctions are $X_n(x) = \sin\left(\frac{\pi}{L} n\, x\right)$. Notice then that we only have to solve the ordinary differential equation $T' + \lambda\alpha^2 T = 0$ for $\lambda = \lambda_n$. We get that $T(t)$ is a constant multiple of

$$T_n(t) = e^{-\lambda_n \alpha^2 t} = e^{-(\pi/L)^2 n^2 \alpha^2 t}.$$

Thus,

$$u_n(x,t) := X_n(t)T_n(t) = e^{-(\pi/L)^2 n^2 \alpha^2 t} \sin\left(\frac{\pi}{L} n\, x\right)$$

is a solution to the PDE $u_t = \alpha^2 u_{xx}$, *and* to the boundary conditions $u(0,t) = u(L,t) = 0$ for all t. Since these are both linear conditions, an arbitrary linear combination of these solutions is again a solution. Thus, we get a "general solution"

$$u(x,t) = \sum_{n=1}^{\infty} b_n u_n = \sum_{n=1}^{\infty} b_n e^{-\lambda_n \alpha^2 t} \sin\left(\frac{\pi}{L} n\, x\right).$$

It is not clear yet that this is the general solution to the problem, only that it is a solution with many arbitrary parameters b_n that we are free to choose. (It will turn out later that this really is the general solution.)

We would like to find constants b_n so that our solution satisfies the initial condition, i.e. that the initial temperature should be

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L} n\, x\right).$$

But we recognize the expression on the right as the Fourier sine series of $f(x)$. In otherwords, since $f(x)$ is only defined from $[0,L]$, we are free to do an odd extension of $f(x)$ to $[-L,L]$ and then compute its Fourier series. Since the extension is odd, the Fourier series will be exactly

$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L} n\, x\right), \qquad b_n = \frac{2}{L}\int_0^L f(x) \sin\left(\frac{\pi}{L} n\, x\right) dx.$$

With this choice of b_n, we have found the temperature $u(x,t)$.

The rod with insulated ends

We model the temperature $u(x,t)$ of a one-dimensional rod of length L where the ends of the rod are perfectly *insulated*. This is an idealized situation; insulated means that there is no heat transfer at either end of the rod. If the ends are located at $x = 0$ and $x = L$, this means that

$$u_x(0,t) = 0 = u_x(L,t).$$

These are the new boundary conditions. The rest of the model is the same as the previous case:

$$\begin{cases} u_t = \alpha^2 u_{xx} \\ u(x,0) = f(x). \end{cases}$$

The first equation is the PDE modeling temperature in one-dimension, and the second equation is the initial temperature.

We solve this using the method of separation of variables, just as before. Let $u(x,t) = X(x)T(t) \neq 0$; then $u_t = \alpha^2 u_{xx} \Rightarrow$

$$XT' = \alpha^2 X''T \qquad \Leftrightarrow \qquad \frac{1}{\alpha^2}\frac{T'}{T} = \frac{X''}{X}.$$

Since the left-hand side depends only on t, while the right-hand side depends only on x, both sides are equal to the same constant:

$$\begin{cases} \dfrac{1}{\alpha^2}\dfrac{T'}{T} = -\lambda \\ \dfrac{X''}{X} = -\lambda \end{cases} \quad \Leftrightarrow \quad \begin{matrix} T' + \lambda\alpha^2 T = 0 \\ X'' + \lambda X = 0. \end{matrix}$$

The two boundary conditions $u_x(0,t) = 0 = u_x(L,t)$ imply that $X'(0)\cdot T(t) = 0 = X'(L)\cdot T(t)$; since $T(t) \neq 0$, we conclude that $X'(0) = 0 = X'(L)$.

Thus $X(x)$ is a solution to the boundary value problem (BVP)

$$\begin{cases} X'' + \lambda X = 0 \\ X'(0) = 0 = X'(L). \end{cases}$$

But we know that the only eigenvalues λ of this problem are $\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2$, $n \geq 1$, and $\lambda_0 = 0$. The corresponding eigenfunctions are $X_n(x) = \cos\left(\frac{\pi}{L}nx\right)$. If $\lambda = \lambda_n$, solving

$$T' + \lambda_n\alpha^2 T = 0$$

for T, we get

$$T = ce^{-\lambda_n\alpha^2 t} = ce^{-(\pi/L)^2 n^2\alpha^2 t},$$

so that for the eigenvalue $\lambda = \lambda_n$ we get the solution

$$u_n(x,t) = X_n(x)T_n(t) = \cos\left(\frac{\pi}{L}n\,x\right)e^{-(\pi/L)^2n^2\alpha^2 t}.$$

It follows that an arbitrary linear combination of these these solutions is again a solution:

$$u(x,t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n e^{-(\pi/L)^2n^2\alpha^2 t}\cos\left(\frac{\pi}{L}n\,x\right);$$

notice that while $u_0(x,t) = 1$, we have written $a_0/2$ instead of a_0. Both are just arbitrary constants, but this choice conforms to our convention for Fourier series.

It turns out that this is the general solution to the PDE with those boundary conditions. For us to obtain a solution with our initial temperature $f(x)$, we need to find the constants a_n, $n \geq 0$ so that

$$u(x,0) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L}nx\right) = f(x).$$

But the expression on the left is just a Fourier cosine series; since $f(x)$ is only defined for $[0,L]$, we may extend $f(x)$ to be even on $[-L,L]$ and then compute its Fourier (cosine) series. In particular, we know that

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L}nx\right)$$

when $a_n = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{\pi}{L}nx\right)dx$.

Example. Consider the insulated rod of length L, with initial temperature given by $f(x) = 1$. What is the temperature $u(x,t)$ for any time t?

Solution: The general solution is

$$u(x,t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n e^{-(\pi/L)^2 n^2 \alpha^2 t} \cos\left(\frac{\pi}{L} n x\right).$$

We need to compute the coefficients a_n, for $n \geq 0$. First, $a_0 = \frac{2}{L}\int_0^L f(x)dx = 2$. If $n \geq 1$,

$$\begin{aligned} a_n &= \frac{2}{L}\int_0^L f(x)\cos\left(\frac{\pi}{L}nx\right)dx \\ &= \frac{2}{L}\frac{L}{\pi n}\sin\left(\frac{\pi}{L}nx\right)\Big|_{x=0}^{x=L} \\ &= 0. \end{aligned}$$

So $u(x,t) = \frac{2}{2} = 1$. □

This should confirm your physical intuition: an insulated rod of constant temperature should remain at constant temperature for all time.

Example. Suppose that an insulated rod of length L has initial temperature $f(x) = \cos\left(\frac{\pi}{L} m\, x\right)$ for $m \geq 1$. What is the temperature $u(x,t)$ for any time t?

Just like before, we compute the Fourier cosine series of $f(x)$: but this is trivial because $f(x)$ is itself one of the basis elements in $\mathcal{B}$:

$$\cos\left(\frac{\pi}{L} m\, x\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\, x\right) \qquad \Rightarrow \qquad \begin{cases} a_n = 0 & \text{if } n \neq m \\ a_n = 1 & \text{if } n = m. \end{cases}$$

Thus, the solution is

$$u(x,t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n e^{-(\pi/L)^2 n^2 \alpha^2 t} \cos\left(\frac{\pi}{L} nx\right) = e^{-(\pi/L)^2 m^2 \alpha^2 t} \cos\left(\frac{\pi}{L} mx\right).$$ □

Notice that the previous solution exhibits exponential decay as $t \longrightarrow \infty$, so that

$$\lim_{t\longrightarrow\infty} u(x,t) = 0.$$

5.1 Proposition. *Consider an insulated rod of length L, with temperature $u(x,t)$. Then the average temperature*

$$Avg(t) := \frac{1}{L}\int_0^L u(x,t)dx$$

is constant (i.e. independent of t), and is equal to $\frac{a_0}{2}$ in the Fourier cosine series of the initial temperature $u(x,0)$.

Proof. To show that $Avg(t)$ is constant, we can prove that its derivative is 0. Indeed,

$$\begin{aligned}\frac{d}{dt}Avg(t) &= \frac{1}{L}\int_0^L \left[\frac{d}{dt}u(x,t)\right]dx = \frac{1}{L}\int_0^L u_t dx \\ &= \frac{1}{L}\int_0^L \alpha^2 u_{xx} dx = \frac{\alpha^2}{L}\left[u_x(L,t) - u_x(0,t)\right] \\ &= 0.\end{aligned}$$

We used the fact that $u_t = \alpha^2 u_{xx}$ (since $u(x,t)$ satisfies the Heat equation) and the boundary conditions for the insulated rod ($u_x(0,t) = 0 = u_x(L,t)$ for all t).

By computing the Fourier cosine series for $u(x,0)$, the 0th term

$$a_0 = \frac{2}{L}\int_0^L u(x,0)dx = 2 \cdot Avg(0).$$ □

5.2 Theorem. *Consider the following PDE, which models the insulated rod of length L with initial temperature $f(x)$:*

$$\begin{cases} u_t = \alpha^2 u_{xx} \\ u_x(0,t) = 0 = u_x(L,t) \\ u(x,0) = f(x). \end{cases}$$

There exists a unique *solution.*

Proof. The existence of the solution was given above. Since $f(x)$ is continuous, its even extension is equal to its Fourier cosine series (except possibly at $x = \pm L$)

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\, x\right),$$

so that

$$u(x,t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n e^{-(\pi/L)^2\alpha^2 t} \cos\left(\frac{\pi}{L} n\, x\right)$$

is a solution. We want to show that this solution is *unique.*

Let $w(x,t)$ be another solution. Observe that the difference, $d(x,t) := u(x,t) - w(x,t)$ satisfies the PDE and the homogeneous boundary conditions, and that

$$d(x,0) = u(x,0) - w(x,0) = f(x) - f(x) = 0.$$

In otherwords, $d(x,t)$ is a solution to the problem when the initial temperature is 0 at every point on the rod.

We want to prove that $u(x,t)$ is the unique solution, i.e. $w(x,t)$ is necessarily equal to $u(x,t)$. Thus, we must show that $d(x,t) = 0$. (Check your physical intuition - this is very believable.) To this end, we define the function

$$E(t) := \frac{1}{2}\int_0^L d(x,t)^2 dx.$$

Physically, $E(t)$ models the total heat energy in the rod. Mathematically, $E(t)$ is just a function that will help us prove uniqueness. Observe that $E(t) \geq 0$ for all t, since the integrand is a non-negative function. Also, observe that

$$E(0) = \frac{1}{2}\int_0^L d(x,0)^2 dx = 0,$$

since $d(x,0) = 0$. Now we compute:

$$\begin{aligned}
\frac{d}{dt}E(t) &= \frac{1}{2}\int_0^L \left[\frac{d}{dt}d(x,t)^2\right] dx \\
&= \int_0^L d(x,t)\cdot d_t(x,t)dx = \int_0^L d(x,t)\cdot \alpha^2 d_{xx}(x,t)dx \\
&= \alpha^2 d(x,t)\cdot d_x(x,t)\big|_{x=0}^{x=L} - \alpha^2\int_0^L d_x(x,t)\cdot d_x(x,t)dx \\
&= -\alpha^2\int_0^L d_x(x,t)^2 dx \leq 0.
\end{aligned}$$

Since $E(0) = 0$, and $E'(t) \leq 0$ for every t, we conclude that $E(t) \leq 0$ for every t. But we already knew that $E(t) \geq 0$, so $E(t) = 0$ for all t. Thus,

$$0 = 2E(t) = \langle d(x,t), d(x,t)\rangle \qquad \Rightarrow \qquad d(x,t) = 0. \qquad \square$$

The method of the previous proof is called the *Energy method* - it is ubiquitous in PDE theory. Observe that the same proof works if we consider the previous boundary conditions

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

instead of those for the insulated rod.

Non-homogeneous boundary conditions

Consider the heat equation $u_t = \alpha^2 u_{xx}$. Observe that any linear polynomial in x,

$$v(x,t) = p_0 + p_1 x$$

is a solution to this equation. If we demand the usual homogeneous boundary conditions,

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

then both p_0 and p_1 must be zero, which is not interesting. But if we instead consider the non-homogeneous boundary conditions $u(0,t) = T_0$ and $u(L,t) = T_L$, then

$$v(x,t) = T_0 + \frac{T_L - T_0}{L} x$$

is a solution to the PDE *and* the two boundary conditions.

5.3 Definition. If $u(x,t)$ is a solution to the heat equation, the limit

$$v(x) := \lim_{t \longrightarrow \infty} u(x,t)$$

is called its *steady-state* solution. The difference

$$w(x,t) := u(x,t) - v(x)$$

is called its *transient* solution.

We want to solve the following PDE:

$$\begin{cases} u_t = \alpha^2 u_{xx} \\ u(0,t) = T_0 \\ u(L,t) = T_L \\ u(x,0) = f(x). \end{cases}$$

If $v(x,t) = T_0 + \frac{T_L - T_0}{L} x$, then observe that the difference

$$d(x,t) := u(x,t) - v(x)$$

between any solution $u(x,t)$ and $v(x,t)$ still satisfies the first equation, since $v(x,t)$ does. Also notice that

$$d(0,t) = u(0,t) - v(0,t) = T_0 - T_0 = 0,$$

and likewise $d(L,t) = 0$. Finally, $d(x,0) = f(x) - v(x,0)$. This is usually not zero, but rather just determines some initial temperature for $d(x,t)$. In summary, the difference $d(x,t)$ satisfies the system

$$\begin{cases} u_t = \alpha^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = f(x) - v(x) \end{cases}$$

We know that this system of equations admits a *unique* solution, say $w(x,t)$. To compute that solution, compute the Fourier sine series

$$f(x) - v(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L} n\, x\right);$$

then we know that

$$w(x,t) = \sum_{n=1}^{\infty} b_n e^{-(\pi/L)^2 n^2 \alpha^2 t} \sin\left(\frac{\pi}{L} n\, x\right).$$

Finally, $u(x,t) = v(x) + w(x,t)$.

Looking at the form of $w(x,t)$ (and those exponentially-decaying terms), we see that $\lim_{t\longrightarrow\infty} w(x,t) = 0$. Thus $v(x,t)$ really is the steady-state solution $v(x)$ and $w(x,t)$ is the transient solution. We also can see that solutions are unique, since if $u_2(x,t)$ is some other solution, we can write $u_2(x,t) = v(x,t) + w_2(x,t)$, where $w_2(x,t)$ satisfies the same PDE as $w(x,t)$. But the solutions to that system are unique! Thus, $w(x,t) = w_2(x,t)$, and so $u(x,t) = u_2(x,t)$.

The half-insulated rod

Suppose now that the left end of the rod, at $x = 0$, is kept at temperature 0, while the right end at $x = L$ is insulated. Then the temperature $u(x,t)$ must satisfy

$$\begin{cases} u_t = \alpha^2 u_{xx} \\ u(0,t) = 0 \\ u_x(L,t) = 0 \\ u(x,0) = f(x). \end{cases}$$

Using separation of variables just as before, we obtain

$$\begin{cases} X'' + \lambda X = 0 \\ X(0) = 0 = X'(L), \end{cases} \qquad T' + \lambda\alpha^2 T = 0.$$

We must find the eigenvalues λ of the BVP for $X(x)$. It is easy to determine that $\lambda = 0$ is *not* an eigenvalue. Then, a general solution for $\lambda \neq 0$ is given by

$$X(x) = c_1 e^{i\sqrt{\lambda}x} + c_2 e^{-i\sqrt{\lambda}x}.$$

The condition $X(0) = 0$ implies that $c_2 = -c_1$; $X'(L) = 0$ implies that

$$e^{i\sqrt{\lambda}L} = -e^{-i\sqrt{\lambda}L} \qquad \Rightarrow \qquad e^{2i\sqrt{\lambda}L} = -1.$$

Check, using Euler's formula, that $e^{i\pi} = -1$. It follows that $-1 = e^{\pi i + 2\pi i n}$ for every integer n. So comparing exponents, we get that

$$2i\sqrt{\lambda}L = \pi i + 2\pi i n \qquad \Rightarrow \qquad \lambda_n = \left(\frac{\pi}{2L}\right)^2 (2n+1)^2.$$

Check that the corresponding eigenfunction is $X_n(x) = \sin\left(\frac{\pi}{2L}(2n+1)\, x\right)$. The corresponding solution for $T(t)$ is

$$T_n(t) = e^{-\lambda_n \alpha^2 t} = e^{-(\pi/2L)^2 (2n+1)^2 \alpha^2 t}.$$

By taking a linear combination of these solutions, we get the general solution

$$u(x,t) = \sum_{n=1}^{\infty} b_n e^{-(\pi/2L)^2(2n+1)^2\alpha^2 t} \sin\left(\frac{\pi}{2L}(2n+1)\,x\right).$$

We still have to find the constants b_n so that $u(x,t)$ satisfies the initial condition:

$$u(x,0) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{2L}(2n+1)\,x\right) = f(x).$$

But this is exactly the odd specialized extension of $f(x)$. We have shown that if

$$b_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{\pi}{2L}(2n+1)x\right)dx,$$

then $f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{2L}(2n+1)\,x\right)$ (except at finitely many points).

6 The Wave Equation

Consider a very thin string of length L, strung tight with its left end at $x = 0$ and its right end is at $x = L$. Let $u(x,t)$ be the height of the string at position x and time t, ($0 \leq x \leq L$, and $t \geq 0$). We begin by deriving the wave equation in one dimension - a PDE that the height function $u(x,t)$ must satisfy.

Similar to the previous derivation, we consider a "slice" of string at position x_0 of length Δx at some time t_0. We consider Δx small enough so that our slice of string is a straight line. Let $T(x,t)$ be the force of tension on the string at position x and time t. We ignore gravity; the only forces acting on our slice of string is the tension at the right, $T(x_0+\Delta x, t_0)$ and the tension at the left, $T(x_0, t_0)$.

We resolve the tension into vertical and horizontal components

$$V(x,t) = T(x,t)\sin(\theta) \qquad H(x,t) = T(x,t)\cos(\theta),$$

where θ is the angle that our string element makes with the equilibrium position. Observe then that

$$V(x,t) = (T(x,t)\cos(\theta))\tan(\theta) = H(x,t)\tan(\theta) = H(x,t)u_x(x,t).$$

This last equality is obvious: just think of "rise over run." Now we consider the horizontal and vertical accelerations:

$$\begin{aligned} V(x_0+\Delta x, t_0) - V(x_0,t_0) &= ma_{vert} = \rho\Delta x u_{tt}(x_0,t_0) \\ H(x_0+\Delta x, t_0) - H(x_0,t_0) &= ma_{horiz} = 0. \end{aligned}$$

The second equation is zero because we are assuming that the string only moves *in one dimension*, i.e. vertically. It follows that $H(x,t)$ is a constant. The first equation can be re-written as

$$\frac{V(x_0+\Delta x, t_0) - V(x_0,t_0)}{\Delta x} = \rho u_{tt}(x_0,t_0),$$

and taking the limit as $\Delta x \longrightarrow 0$, we get

$$V_x(x_0,t_0) = \rho u_{tt}(x_0,t_0).$$

Since $V = Hu_x$, and H is a constant, we conclude that

$$V_x = (Hu_x)_x = Hu_{xx} = \rho u_{tt}.$$

Defining $a^2 = \frac{H}{\rho}$, we get the *wave equation*:

$$\boxed{u_{tt} = a^2 u_{xx}.}$$

Note that a has the units of velocity, and is called the *wave velocity*. It is the speed of a wave traveling through the string.

Endpoints fixed, no initial velocity

To solve the wave equation, we must impose some boundary conditions. Suppose that the endpoints are fixed at the equilibrium height, i.e.

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

Also, suppose that the string is displaced from equilibrium, and let go with no initial velocity. Then $u(x,0) = f(x)$, and $u_t(x,0) = 0$. Altogether, we have the system

$$\begin{cases} u_{tt} = a^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = f(x), \quad u_t(x,0) = 0 \end{cases}$$

We solve this system using separation of variables. Guessing, $u(x,t) = X(x)T(t) \neq 0$, we obtain the (parametrized family of) systems of ordinary differential equations

$$\begin{cases} X''(x) + \lambda X(x) = 0 & T''(t) + a^2\lambda T(t) = 0 \\ X(0) = 0 = X(L) & T'(0) = 0. \end{cases}$$

The BVP for $X(x)$ has eigenvalues $\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2$, $n \geq 1$, and eigenfunctions $X_n(x) = \sin\left(\frac{\pi}{L} n\, x\right)$. The corresponding general solution for $T(t)$ is

$$T(t) = c_1 \cos\left(a\frac{\pi}{L} n\, t\right) + c_2 \sin\left(a\frac{\pi}{L} n\, t\right).$$

Since $T_n'(0) = 0$, we conclude that $c_2 = 0$. Thus, any solution $T(t)$ for the eigenvalue λ_n is a multiple of

$$T_n(t) = \cos\left(a\frac{\pi}{L}n\,t\right).$$

So $u_n(x,t) = X_n(x)T_n(t) = \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(a\frac{\pi}{L}n\,t\right)$ is a solution to the PDE, boundary values, and has zero initial velocity. Then the same is true of the formal linear combination

$$u(x,t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(a\frac{\pi}{L}n\,t\right).$$

We will check later that this is the general solution; at this point, we just want to find constants b_n so that the last condition $u(x,0) = f(x)$ is satisfied. Then we want

$$u(x,0) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right) = f(x).$$

Recognize this as the Fourier sine series for $f(x)$ of period $2L$; in other words, we can do the odd extension of $f(x)$ to $[-L, L]$ and compute its Fourier series. If $b_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{\pi}{L}n\,x\right)dx$, then

$$u(x,t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(a\frac{\pi}{L}n\,t\right)$$

is a solution to our system of equations.

How do we know that this solution converges when $t \neq 0$? In the case of the heat equation, convergence was obvious because of the exponentially decaying functions. It is not as clear in this case. We will show a nice way to re-write the solution above so that convergence is clear.

6.1 Proposition. *Consider the formal sum*

$$u(x,t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(a\frac{\pi}{L}n\,t\right),$$

and let $u(x,0) = f(x)$. Let $\hat{f}(x)$ be the odd extension of $f_{av}(x)$ to $[-L, L]$, further extended periodically to $\mathbb{R}$ of period $2L$. If $\hat{f}(x)$ is in $\mathcal{F}$, then

$$u(x,t) = \frac{\hat{f}(x-at) + \hat{f}(x+at)}{2}.$$

In particular, $u(x,t)$ is convergent for all $0 \leq x \leq L$, and t in $\mathbb{R}$.

Proof. Since $\hat{f}(x)$ is in $\mathcal{F}$, the Fourier (sine) series converges, and we may identify

$$\hat{f}(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right).$$

Using the trig identity $\sin(A \pm B) = \sin(A)\cos(B) \pm \cos(A)\sin(B)$:

$$\hat{f}(x-at) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,(x-at)\right) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(\frac{\pi}{L}na\,t\right) - \cos\left(\frac{\pi}{L}n\,x\right)\sin\left(\frac{\pi}{L}n\,at\right)$$

$$\hat{f}(x+at) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,(x+at)\right) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(\frac{\pi}{L}na\,t\right) + \cos\left(\frac{\pi}{L}n\,x\right)\sin\left(\frac{\pi}{L}n\,at\right).$$

By summing, we get

$$\hat{f}(x-at) + \hat{f}(x+at) = 2\sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(\frac{\pi}{L}na\,t\right) = 2u(x,t). \qquad \square$$

Returning to the string problem, we see that if the initial position function $f(x)$ is piecewise continuous, then $\hat{f}(x)$ is in $\mathcal{F}$, and we have a solution

$$u(x,t) = \frac{\hat{f}(x-at) + \hat{f}(x+at)}{2} = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi}{L}n\,x\right)\cos\left(a\frac{\pi}{L}n\,t\right).$$

It is possible to prove that $u(x,t) = \frac{1}{2}\left[\hat{f}(x-at) + \hat{f}(x+at)\right]$ is a solution directly, without appealing to separation of variables and Fourier series. Why did we bother solving the equation with separation of variables?

First, it is hard to guess a closed-form solution like $\frac{\hat{f}(x-at)+\hat{f}(x+at)}{2}$, and it is unusual that a PDE should have such a solution. So it is instructive to understand more general methods like separation of variables since they apply to more problems. Second, the eigenvalues and eigenfunctions have physical significance. Fix an eigenvalue $\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2$. The corresponding temporal function $T_n(t) = \cos\left(a\frac{\pi}{L}n\,t\right)$ has frequency

$$\frac{a\pi}{L}n.$$

This is a *natural frequency* of the string – a frequency at which the string will freely vibrate. The displacement pattern that you observe when vibrating a string at its natural frequency is (a multiple of) $X_n(x) = \sin\left(\frac{\pi}{L}n\,x\right)$. This is called a *natural mode* of vibration. In particular, there are $n+1$ zeroes in $[0,L]$ for $n \geq 1$ that are easily observed. We understand then that the total motion of the string is a superposition of its natural modes of vibration.

Endpoints fixed, zero initial position

Now suppose begins from the equilibrium without any displacement, and with some initial velocity $g(x)$. We still fix the endpoints at $u = 0$, so altogether we have the system

$$\begin{cases} u_{tt} = a^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = 0, \quad u_t(x,0) = g(x) \end{cases}$$

We solve this system using separation of variables. Guessing, $u(x,t) = X(x)T(t) \neq 0$, we obtain the (parametrized family of) systems of ordinary differential equations

$$\begin{cases} X''(x) + \lambda X(x) = 0 & T''(t) + a^2\lambda T(t) = 0 \\ X(0) = 0 = X(L) & T(0) = 0. \end{cases}$$

The BVP for $X(x)$ has eigenvalues $\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2$, $n \geq 1$, and eigenfunctions $X_n(x) = \sin\left(\frac{\pi}{L} n\, x\right)$. The corresponding general solution for $T(t)$ is

$$T(t) = c_1 \cos\left(a\frac{\pi}{L} n\, t\right) + c_2 \sin\left(a\frac{\pi}{L} n\, t\right).$$

Since $T_n(0) = 0$, we conclude that $c_1 = 0$. Thus, any solution $T(t)$ for the eigenvalue λ_n is a multiple of

$$T_n(t) = \sin\left(a\frac{\pi}{L} n\, t\right).$$

So $u_n(x,t) = X_n(x)T_n(t) = \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(a\frac{\pi}{L} n\, t\right)$ is a solution to the PDE, boundary values, and has zero initial position. Then the same is true of the formal linear combination

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(a\frac{\pi}{L} n\, t\right).$$

We will check later that this is the general solution; at this point, we just want to find constants c_n so that the last condition $u_t(x,0) = g(x)$ is satisfied. Then we want

$$u_t(x,0) = \sum_{n=1}^{\infty} c_n \left(a\frac{\pi}{L} n\right) \cdot \sin\left(\frac{\pi}{L} n\, x\right) = g(x).$$

This looks vaguely like the Fourier sine series for $g(x)$ of period $2L$, but with the coefficients scaled; in other words, we can do the odd extension of $g(x)$ to $[-L, L]$ and compute its Fourier series. Then we want

$$c_n \left(a\frac{\pi}{L} n\right) = b_n = \frac{2}{L} \int_0^L g(x) \sin\left(\frac{\pi}{L} n\, x\right) dx;$$

in otherwords, choosing

$$c_n = \frac{2}{a\pi n} \int_0^L g(x) \sin\left(\frac{\pi}{L} n\, x\right),$$

we have that

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(a\frac{\pi}{L} n\, t\right)$$

is a solution to our system of equations.

We run into the same issue of convergence.

6.2 Proposition. *Consider the formal sum*

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(a \frac{\pi}{L} n\, t\right),$$

and assume that $u_t(x,0) = g(x)$ and its derivative are both piece-wise continuous. Let $\hat{g}(x)$ be the odd extension of $g_{av}(x)$ to $[-L, L]$, further extended periodically to $\mathbb{R}$ of period $2L$. Since $\hat{g}(x)$ is odd and periodic, its average value on any period is 0. Thus, an anti-derivative $G(x)$ of $\hat{g}(x)$ is periodic, and

$$u(x,t) = \frac{G(x+at) - G(x-at)}{2a}.$$

In particular, $u(x,t)$ is convergent for all $0 \le x \le L$, and t in $\mathbb{R}$.

Proof. Since $G(x)$ is even and is in $\mathcal{F}$, it has a Fourier cosine series

$$G(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\, x\right).$$

Compute that

$$\begin{aligned} G'(x) &= \sum_{n=1}^{\infty} -a_n \frac{\pi}{L} n \sin\left(\frac{\pi}{L} n\, x\right) \\ u_t(x,t) &= \sum_{n=1}^{\infty} c_n \frac{\pi}{L} na \cdot \sin\left(\frac{\pi}{L} n\, x\right) \cos\left(a \frac{\pi}{L} n\, t\right), \end{aligned}$$

Since $G'(x) = u_t(x,0)$, we see that $-a_n = a\, c_n$. Then compute

$$\begin{aligned} G(x+at) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\,(x+at)\right) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\, x\right) \cos\left(\frac{\pi}{L} na\, t\right) - a_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(\frac{\pi}{L} na\, t\right) \\ G(x-at) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\,(x-at)\right) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi}{L} n\, x\right) \cos\left(\frac{\pi}{L} na\, t\right) + a_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(\frac{\pi}{L} na\, t\right). \end{aligned}$$

Subtracting, we get

$$\begin{aligned} G(x+at) - G(x-at) &= 2 \sum_{n=1}^{\infty} -a_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(\frac{\pi}{L} na\, t\right) \\ &= 2 \sum_{n=1}^{\infty} a\, c_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(\frac{\pi}{L} na\, t\right) = 2a\, u(x,t). \end{aligned}$$

□

We may assume that the initial velocity of the string is piece-wise continuous, so that the previous proposition implies that

$$u(x,t) = \frac{G(x+at) - G(x-at)}{2a} = \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi}{L} n\, x\right) \sin\left(a \frac{\pi}{L} n\, t\right)$$

is a solution to our system of equations, where

$$G'(x) = g(x) \qquad \text{and} \qquad c_n = \frac{2}{a\pi n} \int_0^L g(x) \sin\left(\frac{\pi}{L} n\, x\right) dx.$$

Fixed ends, general case

It is easy to combine the two previous cases. If the endpoints remain fixed at $u = 0$, but the string is initially displaced at $f(x)$ with some initial velocity $g(x)$, we can solve the system

$$\begin{cases} u_{tt} = a^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = f(x), \quad u_t(x,0) = g(x) \end{cases}$$

by summing solutions to the corresponding systems

$$\begin{cases} u_{tt} = a^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = f(x), \quad u_t(x,0) = 0 \end{cases} \qquad \begin{matrix} u_{tt} = a^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = 0, \quad u_t(x,0) = g(x) \end{matrix}$$

In particular, we have the closed form solution

$$u(x,t) = \frac{f(x+at) + f(x-at)}{2} + \frac{G(x+at) - G(x-at)}{2},$$

where we do odd extensions of $f(x)$ and $g(x)$, extend periodically, and let $G(x)$ be an anti-derivative of $g(x)$.

Uniqueness

As we did for the heat equation, we would like to check that our model admits *unique* solutions. We use the energy method. The total energy in the string is given by the sum of its potential and kinetic energy at any time. The kinetic energy should be "$\frac{1}{2}\, m\, v^2$" - for our slice of string this is $\frac{1}{2}\rho \Delta x u_t^2$. Adding this up over the string we get

$$KE(t) = \frac{1}{2} \int_0^L \rho u_t^2 dx.$$

The potential energy is a bit more complicated. Consider the original picture we had when deriving the wave equation. We had a small right triangle with base Δx, height Δu, and angle $\theta = \tan^{-1}(\Delta u/\Delta x)$ (i.e. $\tan\theta = u_x$). If you consider the work required to move the slice of string from the base, where its length is the natural length Δx, to the hypotenuse, where its length is $\sqrt{\Delta x^2 + \Delta u^2}$, the only work is done by stretching the string "against" the force of tension. The tension force is always tangent to the string. So if the tension force is the constant T, then the work ("force×distance") is

$$\begin{aligned} W &= T \cdot \left[\sqrt{\Delta x^2 + \Delta u^2} - \Delta x\right] \\ &= T \cdot \left[\sqrt{1 + \left(\frac{\Delta u}{\Delta x}\right)^2} - 1\right] \Delta x \\ &= T \cdot \left[\left(1 + \frac{1}{2}\left(\frac{\Delta u}{\Delta x}\right)^2 - \frac{1}{8}\left(\frac{\Delta u}{\Delta x}\right)^4 + \cdots\right) - 1\right] \Delta x \\ &\approx \frac{T}{2} u_x^2 \Delta x, \end{aligned}$$

as $\Delta x \longrightarrow 0$. Here, we have used the Taylor series

$$\sqrt{1+t^2} = \sum_{n=0}^{\infty} \binom{1/2}{n} t^{2n}, \qquad |t| < 1$$

with $t = \Delta u/\Delta x$. Summing over all the slices in the string, we get that

$$PE(t) = \frac{T}{2} \int_0^L u_x^2 dx.$$

Then the energy function is $E(t) = \frac{1}{2}\int_0^L \left[\rho u_t^2 + T u_x^2\right] dx$.

6.3 Theorem. *Consider the following PDE, which models the vibrating string of length L with initial position $f(x)$ and initial velocity $g(x)$:*

$$\begin{cases} u_{tt} = a^2 u_{xx} \\ u(0,t) = 0 = u(L,t) \\ u(x,0) = f(x) \quad u_t(x,0) = g(x). \end{cases}$$

There exists a unique *solution.*

Proof. By subtracting solutions (like in the case with the heat equation), it is enough to prove uniqueness when $f(x) = g(x) = 0$. In that case, we must show that 0 is the only solution. To that end, consider any solution $u(x,t)$, and define the energy

$$E(t) = \frac{1}{2} \int_0^L \left[\rho u_t^2 + T u_x^2\right] dx$$

in the string at any time t. Clearly $E(t) \geq 0$. Using the assumed initial conditions, we compute

$$E(0) = \frac{1}{2}\int_0^L \left[\rho u_t(x,0)^2 + Tu_x(x,0)^2\right] dx = 0.$$

We will prove that $E'(t) = 0$. Indeed, by differentiating under the integral we have

$$\begin{aligned}
E'(t) &= \int_0^L [\rho u_t u_{tt} + Tu_x u_{xt}]\, dx \\
&= \rho \int_0^L u_t u_{tt}\, dx + T\int_0^L u_x u_{tx}\, dx \\
&= \rho \int_0^L u_t u_{tt}\, dx + Tu_x \cdot u_t|_{x=0}^{x=L} - T\int_0^L u_t u_{xx}\, dx \\
&= \int_0^L u_t(\rho u_{tt} - Tu_{xx})\, dx + Tu_x \cdot u_t|_{x=0}^{x=L} \\
&= Tu_x \cdot u_t|_{x=0}^{x=L} = 0.
\end{aligned}$$

Indeed, the boundary conditions $u(0,t) = 0 = u(L,t)$ imply that $u_t(0,t) = u_t(L,t) = 0$, which gives the very last equality. Since $E'(t) = 0$, $E(t)$ is a constant, and so must be 0 since $E(0) = 0$. It follows that both $u_t(x,t))$ and $u_x(x,t)$ are zero. Since $u_x(x,t) = u_t(x,t) = 0$, so $u(x,t)$ is constant. Since $u(x,0) = 0$, we have $u(x,t) = 0$. $\square$

It is worth noticing that this theorem also applies to the case where we let one (or both) of the sides move vertically; for instance, if the right side of the string at $x = L$ is free, then the right boundary condition is $u_x(L,t) = 0$ for all $t \geq 0$. You will consider this kind of problem in the homework.

7 Laplace's Equation

It is true that the heat equation in two dimensions is

$$u_t = \alpha^2(u_{xx} + u_{yy}).$$

We could use this to model heat transfer in a very thin sheet of metal, for example. However, this is a PDE of 3 variables, x, y and t, and in this class we will restrict to just studying PDE's of 2 variables.

But we could at least look at the *steady-state* solution to the above equation. This is exactly the part of any solution which is independent of time, implying that left-hand side is 0. In otherwords, a steady-state solution to the heat equation in two variables is a solution to

$$\boxed{u_{xx} + u_{yy} = 0.}$$

This is known as *Laplace's equation*.

Before we go through the usual yoga of separating variables and obtaining series solutions, it is worthwhile to compare this with some of the concepts that you covered in multivariable calculus (Math 126). Much of this is just terminology, but it occurs often in "engineering math" and physics.

7.1 Definition. Consider a differentiable real-valued function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. The *gradient* of f, written $\text{grad}(f)$ or ∇f is the function from $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by

$$\text{grad}(f) = \left[\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right].$$

The symbol ∇ is pronounced "del," though it is called the nabla symbol. The gradient gives the best linear approximation to the function f at any point; you can think of it as a more general derivative: it gives the slope of the tangent line approximation to f, for all of the various lines determined by fixing all but one of the independent variables.

If we instead start with a function $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, we can think of g as a collection of m functions $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, for $1 \le i \le m$, by projecting onto the ith coordinate of the range of g. If we consider the gradient of each g_i, we get a collection of m vector-valued functions (vectors of length n). It is natural to organize all of this data into an $m \times n$ matrix:

$$\text{jac}(g) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

This matrix is called the *Jacobian* of g. It gives the best linear approximation to g at any point.

7.2 Definition. Consider a function $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. The *divergence* of g is the real-valued function

$$\text{div}(g) = \frac{\partial g_1}{\partial x_1} + \frac{\partial g_2}{\partial x_2} + \cdots + \frac{\partial g_n}{\partial x_n}.$$

This is often written as $\nabla \cdot g$, as a mnemonic device, where $\nabla = \left[\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right]$.

Thinking of $\nabla = \left[\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right]$ as an operator can be intuitively useful. We can try to make sense of $\nabla \cdot \nabla$: if $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a real-valued function,

$$\nabla \cdot \nabla f = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_i} f = f_{x_1 x_1} + \cdots + f_{x_n x_n}.$$

This operator $\Delta = \nabla \cdot \nabla$, sometimes written ∇^2, is called the *Laplacian*. The previous two formulas are just shorthand for the composition $\Delta = \text{div} \circ \text{grad}$. The symbol Δ is a capital "delta." If $u : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a twice-differentiable function of two variables, then Laplace's equation is exactly

$$\Delta u = 0.$$

7.3 Definition. A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is called *harmonic* if its Laplacian

$$\Delta f = 0.$$

Harmonic functions turn out to be very important. In physics, they are steady-state solutions to the heat equation in n-variables. In mathematics, there are a great deal of very interesting theorems concerning harmonic functions. This subject is called *harmonic analysis*. For example, a function f that is twice-differentiable and harmonic must actually be infinitely differentiable. Another example: on any closed and bounded set in the domain, a harmonic function must attain both its maximum and minimum values on the boundary of that set.

For any region Ω in $\mathbb{R}^n$, we can hope to find a harmonic function on that region with some prescribed boundary values. This problem is called the *Dirichlet problem* for Ω.

Dirichlet problem for a rectangle

We now want to solve Laplace's equation $u_{xx} + u_{yy} = 0$ on the closed rectangle $0 \leq x \leq V$, and $0 \leq y \leq L$. We impose some boundary conditions: suppose that the restriction of $u(x, y)$ to the top, bottom and left edge is 0, while the restriction of $u(x, y)$ to the vertical line $x = V$ is some function $f_R(y)$. (The R stands for "right," i.e. right edge).

Employ the usual separation of variables technique: guess $u(x, y) = X(x)Y(y) \neq 0$, and get $-X''Y = XY''$, or

$$-\frac{X''}{X} = \frac{Y''}{Y} = -\lambda.$$

The 3 homogeneous boundary conditions give

Top : $0 = X(x)Y(L)$ implies $Y(L) = 0$

Bottom : $0 = X(x)Y(0)$ implies $Y(0) = 0$

Left : $0 = X(0)Y(y)$ implies $X(0) = 0$.

Thus we get the two systems

$$\begin{cases} X'' - \lambda X = 0 \\ X(0) = 0 \end{cases} \qquad \begin{matrix} Y'' + \lambda Y = 0 \\ Y(0) = 0 = Y(L). \end{matrix}$$

The BVP for Y is the familiar one. We know that $\lambda_n = \left(\frac{\pi}{L}\right)^2 n^2$ for $n \geq 1$, and the corresponding eigenfunction is $Y_n(y) = \sin\left(\frac{\pi}{L} n\, y\right)$. The general solution for $X(x)$ is

$$X(x) = c_1 \cosh\left(\frac{\pi}{L} n\, x\right) + c_2 \sinh\left(\frac{\pi}{L} n\, x\right).$$

However, the condition $X(0) = 0$ implies that $c_1 = 0$. Thus we may choose $X_n(x) = \sinh\left(\frac{\pi}{L} n\, x\right)$; then $u_n(X) = X_n(x)Y_n(x) = \sinh\left(\frac{\pi}{L} n\, x\right) \sin\left(\frac{\pi}{L} n\, y\right)$. A more general solution is given by a convergent linear combination

$$u(x, y) = \sum_{n=1}^{\infty} c_n u_n = \sum_{n=1}^{\infty} c_n \sinh\left(\frac{\pi}{L} n\, x\right) \sin\left(\frac{\pi}{L} n\, y\right).$$

Finally, we must choose the constants c_n in order to satisfy the initial condition $u(V, y) = f_R(y)$. So we want

$$u(V, y) = \sum_{n=1}^{\infty} c_n \sinh\left(\frac{\pi}{L} n\, V\right) \sin\left(\frac{\pi}{L} n\, y\right) = f_R(y).$$

Comparing this to the Fourier sine series of $f_R(y)$, we want

$$c_n \sinh\left(\frac{\pi}{L} n\, V\right) = b_n = \frac{2}{L} \int_0^L f_R(y) \sin\left(\frac{\pi}{L} n\, y\right) dy,$$

i.e. $c_n = \frac{1}{\sinh\left(\frac{\pi}{L} n\, V\right)} \cdot \frac{2}{L} \int_0^L f_R(y) \sin\left(\frac{\pi}{L} n\, y\right) dy$. With this choice of c_n, the resulting series is convergent, so is a solution to the problem.

Of course, more generally we could have non-zero boundary conditions at more than one edge. The idea is to solve the problem one edge at a time, as we have just done, and then add the resulting solutions together. This sum will still be a harmonic function, and will satisfy all four boundary conditions.

Dirichlet problem for the disk

Consider the unit disk $D \subset \mathbb{R}^2$, i.e. $D = \{(x, y) \mid x^2 + y^2 \leq 1\}$. We want to find a harmonic function u such that the restriction of u to the boundary circle is a prescribed function f. This is easier to describe if we switch to polar coordinates. In polar coordinates, we want a harmonic function $u(r, \theta)$ so that $u(1, \theta) = f(\theta)$ for some (2π-periodic) function f.

However, we must interpret Laplace's equation in polar coordinates! There is a function $g : \mathbb{R}^2_p \longrightarrow \mathbb{R}^2_r$ which takes a vector in polar coordinates and gives the vector in rectangular coordinates. Indeed, since $x = r\cos(\theta)$, and $y = r\sin(\theta)$, we have that

$$g(r, \theta) = [r\cos(\theta), r\sin(\theta)].$$

You can verify using the chain rule that

$$\begin{bmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{bmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -r\sin(\theta) & r\cos(\theta) \end{pmatrix} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix}$$

(This matrix is the transpose of the Jacobian of g.) It follows that

$$\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -r\sin(\theta) & r\cos(\theta) \end{pmatrix}^{-1} \begin{bmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{bmatrix} = \begin{pmatrix} \cos(\theta) & -\frac{1}{r}\sin(\theta) \\ \sin(\theta) & \frac{1}{r}\cos(\theta) \end{pmatrix} \begin{bmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{bmatrix}$$

It follows that

$$\begin{aligned} \frac{\partial}{\partial x} \circ \frac{\partial}{\partial x} &= \left[\cos(\theta)\frac{\partial}{\partial r} - \frac{1}{r}\sin(\theta)\frac{\partial}{\partial \theta}\right] \circ \left[\cos(\theta)\frac{\partial}{\partial r} - \frac{1}{r}\sin(\theta)\frac{\partial}{\partial \theta}\right] \\ &= \cos^2(\theta)\partial_{rr} - \frac{1}{r}\cos(\theta)\partial_{r\theta} + \frac{1}{r^2}\sin(\theta)\partial_\theta + \frac{1}{r}\sin^2(\theta)\partial_r \\ &\quad - \frac{1}{r}\sin(\theta)\cos(\theta)\partial_{\theta r} + \frac{1}{r^2}\sin(\theta)\cos(\theta)\partial_\theta + \frac{1}{r^2}\sin^2(\theta)\partial_{\theta\theta}; \\ \frac{\partial}{\partial y} \circ \frac{\partial}{\partial y} &= \left[\sin(\theta)\frac{\partial}{\partial r} + \frac{1}{r}\cos(\theta)\frac{\partial}{\partial \theta}\right] \circ \left[\sin(\theta)\frac{\partial}{\partial r} + \frac{1}{r}\cos(\theta)\frac{\partial}{\partial \theta}\right] \\ &= \sin^2(\theta)\partial_{rr} - \frac{1}{r^2}\cos(\theta)\sin(\theta)\partial_\theta + \frac{1}{r}\sin(\theta)\cos(\theta)\partial_{r\theta} + \frac{1}{r}\cos^2(\theta)\partial_r \\ &\quad + \frac{1}{r}\cos(\theta)\sin(\theta)\partial_{\theta r} - \frac{1}{r^2}\cos(\theta)\sin(\theta)\partial_\theta + \frac{1}{r^2}\cos^2(\theta)\partial_{\theta\theta}. \end{aligned}$$

It is a small miracle that summing the two simplifies as

$$\Delta = \partial_{xx} + \partial_{yy} = \partial_{rr} + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_{\theta\theta},$$

So that

$$\Delta u = u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta} = 0$$

is the heat equation in polar coordinates. We proceed to solve this using separation of variables.

Guess $u(r,\theta) = R(r)\Theta(\theta) \neq 0$, so that $R''\Theta + \frac{1}{r}R'\Theta + \frac{1}{r^2}R\Theta'' = 0$, or

$$r^2\frac{R''}{R} + r\frac{R'}{R} = -\frac{\Theta''}{\Theta} = \lambda.$$

This yields the systems

$$\begin{cases} \Theta'' + \lambda\Theta = 0 \qquad r^2R'' + rR' - \lambda R = 0 \\ \end{cases}$$

On the face of it, it seems that we have no boundary conditions! But we do have the condition that Θ must be periodic of period 2π; since only periodic solutions occur when $\lambda \geq 0$, we immediately restrict to that situation. Then the typical solutions $\sin(\sqrt{\lambda}\,\theta)$ and $\cos(\sqrt{\lambda}\,\theta)$ are periodic of period $2\pi/\sqrt{\lambda}$. If 2π is to be a period, it must be an integer multiple of this fundamental period

$$2\pi = \left(\frac{2\pi}{\sqrt{\lambda}}\right)n,$$

i.e. $\lambda_n = n^2$ for $n \geq 0$. The eigenspace for each (positive) eigenvalue is dimension 2: any eigenfunction is a linear combination

$$\Theta_n(\theta) = a_n \cos(n\theta) + b_n \sin(n\theta).$$

Finally, we must solve the ordinary differential equations

$$r^2 R'' + rR' - n^2 R = 0.$$

These are *Euler equations*, which one can solve by guessing $R(r) = r^\alpha$. By plugging in, you will find that

$$\alpha(\alpha - 1) + \alpha - n^2 = 0 \qquad \Leftrightarrow \qquad \alpha = \pm n.$$

So $R(r) = c_1 r^n + c_2 r^{-n}$. However, we demand that $R(r)$ be continuous at the origin, enforcing that $c_2 = 0$. Thus, we can fix $R_n(r) = r^n$.

Then $u_n(r, \theta) = r^n (a_n \cos(n\theta) + b_n \sin(n\theta))$. Forming a more general solution, we get the linear combination

$$u(r, \theta) = a_0 + \sum_{n=1}^{\infty} r^n \left[a_n \cos(n\theta) + b_n \sin(n\theta)\right].$$

We might as well replace a_0 with $a_0/2$, to make it look like the usual Fourier series. Then if we evaluate $u(r, \theta)$ at the boundary $r = 1$, we want

$$u(1, \theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta) = f(\theta).$$

This is exactly the usual fourier series for the boundary function $f(\theta)$, so we must choose the constants as usual:

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) \cos(n\theta) d\theta \qquad b_n = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) \sin(n\theta) d\theta.$$

Uniqueness

We want to show that the solutions to a given Dirichlet problem are unique. To do this, we will employ the following theorem

7.4 Theorem (Divergence Theorem)**.** *Let D be a bounded domain in $\mathbb{R}^2$, and let $f : D \cup \partial D \longrightarrow \mathbb{R}^n$ be a continuously differentiable function. Then*

$$\int_D \operatorname{div}(F) dV = \int_{\partial D} (F \cdot n) dS.$$

Here, the integral on the left is a volume integral ("dV"), while the integral on the right is a surface integral ("dS"); in practice these integrals would be computed by parametrizing the region D and its boundary. Also, n is the outword-pointing normal vector to the boundary ∂D. Intuitively, $F \cdot n$ measures the flow of F across the boundary of ∂D.

7.5 Theorem. *Let $u(x, y)$ be a harmonic function in a region D such that $u|_{\partial D} = 0$. Then $u = 0$.*

Proof. We employ the "energy method" again; this time with

$$E = \int_D |\operatorname{grad}(u)|^2 dV.$$

First, observe that for any function $u : \mathbb{R}^n \longrightarrow \mathbb{R}$,

$$\operatorname{div}(u \operatorname{grad}(u)) = |\operatorname{grad}(u)|^2 + u \operatorname{div}(\operatorname{grad}(u)) = |\operatorname{grad}(u)|^2 + u\Delta u.$$

This is an easy computation, and only uses the product rule for partial derivatives.

Now since $\Delta u = 0$, we have

$$E = \int_D |\operatorname{grad}(u)|^2 dV + \int_D u\Delta u dV = \int_D \operatorname{div}(u \operatorname{grad}(u)) du.$$

Using the divergence theorem, we see then that

$$E = \int_D \operatorname{div}(u \operatorname{grad}(u)) du = \int_{\partial D} u \operatorname{grad}(u) \cdot n dS = 0,$$

because $u|_{\partial D} = 0$.

Since $E = \int_D |\operatorname{grad}(u)|^2 dV = 0$, we conclude that $\operatorname{grad}(u) = 0$, i.e. u is constant. Since $u|_{\partial D} = 0$, $u = 0$. □

Uniqueness for any Dirichlet problem follows by considering the difference of any two solutions. The previous theorem applies to the difference, and shows that the difference is zero. Thus any two solutions must be equal.

We now prove one of the special properties of Harmonic functions. This gives an alternate proof of the uniqueness of solutions to the Dirichlet problem, which does not appeal to any involved theorems like the divergence theorem.

7.6 Theorem (Weak Maximum Principle)**.** *Consider a bounded domain $D \subset \mathbb{R}^2$, and Let $f : D \cup \partial D \longrightarrow \mathbb{R}$ be a harmonic function. Then f attains its maximum value on ∂D.*

Proof. We know from calculus (Math 126) that if f attains a local maximum at (x_0, y_0), then both of the second partials $f_{xx}(x_0, y_0)$ and $f_{yy}(x_0, y_0)$ are less than 0 (think concave down). In particular, this implies that

$$\Delta f = f_{xx} + f_{yy} \leq 0 \qquad \text{at } (x_0, y_0).$$

It follows that any function $g(x, y)$ which satisfies $\Delta g > 0$ for all points in the domain D must not have a local maximum in D.

Let $M = \max_{(x,y)\in\partial D} f(x, y)$. Define $g(x, y) = x^2 + y^2$. Then $\Delta g = 4 > 0$ for every point in D, so we know that g attains its maximum value on the boundary ∂D. Call this maximum value M'. If $\varepsilon > 0$ is any positive number, define $f_\varepsilon(x, y) := f(x, y) + \varepsilon g(x, y)$, and note that

$$\Delta f_\varepsilon = 0 + 4\varepsilon > 0.$$

Again, we conclude that f_ε does not have a local maximum in D; it therefore attains its maximum value, on the boundary ∂D. In otherwords, there is some point (x_0, y_0) in ∂D such that

$$f_\varepsilon(x, y) \leq f_\varepsilon(x_0, y_0)$$

for every point $(x, y) \in D \cup \partial D$. But we know that

$$f_\varepsilon(x_0, y_0) = f(x_0, y_0) + \varepsilon g(x_0, y_0) \leq M + \varepsilon M',$$

since (x_0, y_0) is on the boundary; thus we have

$$f_\varepsilon(x, y) \leq M + \varepsilon M'.$$

Taking the limit as $\varepsilon \longrightarrow 0$, we get

$$f(x, y) \leq M = \max_{(x,y)\in\partial D} f(x, y). \qquad \square$$

Extra Credit Prove the uniqueness theorem for the solution to a Dirichlet problem in a bounded domain D without using the divergence theorem. Hint: Use the weak maximum principle.

8 A classification

So far we have studied 3 types of 2nd-order linear homogeneous PDEs (in 2 variables):

Heat Equation $u_t = \alpha^2 u_{xx}$

Wave Equation $u_{tt} = a^2 u_{xx}$

Laplace Equation $u_{tt} + u_{xx} = 0$

For any real constants $\{A, B, C, D, E, F\} \subset \mathbb{R}$, it turns out that the equation

$$Au_{xx} + Bu_{xt} + Cu_{tt} + Du_x + Eu_t + Fu = 0$$

can be transformed into an equation very similar to one of the equations above. In this section, we see how this is done. This is clearly useful to us, since we already know how to solve all the equations above. Also, directly using separation of variables will often fail on an equation in this general form.

Step 1: Eliminate the cross-term

The first trick is the following. For constants a, b, c, we have

$$au_{xx} + 2bu_{xt} + cu_{tt} = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial t} \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial t} \end{bmatrix} u(x,t).$$

Just "multiply" out the right-hand side and see what you get! Let $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be the coefficient matrix above.

We consider the effect of a linear change of coordinates on the form of the differential equation. Let $M = (m_{ij})$ be a 2×2 invertible matrix, and define

$$\begin{pmatrix} y \\ s \end{pmatrix} := M \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} m_{11}x + m_{12}t \\ m_{21}x + m_{22}t \end{pmatrix}.$$

Compute (using the chain rule) that

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial y}{\partial x}\frac{\partial}{\partial y} + \frac{\partial s}{\partial x}\frac{\partial}{\partial s} = m_{11}\frac{\partial}{\partial y} + m_{21}\frac{\partial}{\partial s} \\ \frac{\partial}{\partial t} &= \frac{\partial y}{\partial t}\frac{\partial}{\partial y} + \frac{\partial s}{\partial t}\frac{\partial}{\partial s} = m_{12}\frac{\partial}{\partial y} + m_{22}\frac{\partial}{\partial s}, \end{aligned}$$

or in matrix notation,

$$\begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial t} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial s} \end{bmatrix} M.$$

Using the trick above, the change of coordinates transforms $au_{xx} + 2bu_{xt} + cu_{tt}$ as

$$\begin{aligned} \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial t} \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial t} \end{bmatrix} u(x,t) &= \begin{bmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial s} \end{bmatrix} M \begin{bmatrix} a & b \\ b & c \end{bmatrix} M^T \begin{bmatrix} \frac{\partial}{\partial y} \\ \frac{\partial}{\partial s} \end{bmatrix} u(y,s) \\ &= \begin{bmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial s} \end{bmatrix} MAM^T \begin{bmatrix} \frac{\partial}{\partial y} \\ \frac{\partial}{\partial s} \end{bmatrix} u(y,s). \end{aligned}$$

It remains to pick M in a clever way. Eliminating the cross-term entails making the off-diagonal entries in the coefficient matrix 0, i.e. we want to make MAM^T diagonal. But how? Since A is a real-symmetric matrix, we can find an orthogonal matrix U (i.e. a unitary matrix with real-valued entries) so that

$$U^*AU = U^TAU = D \qquad \text{is diagonal.}$$

Fix such a matrix U, and let $M = U^T$. Then $MAM^T = U^TAU^{TT} = U^TAU = D$ is diagonal, as desired. Using this new coordinate system, the cross-term is eliminated. Indeed, if λ_1 and λ_2 are the two (real) eigenvalues of A, then $MAM^T = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, so

$$au_{xx} + 2bu_{xt} + cu_{tt} = \begin{bmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial s} \end{bmatrix} MAM^T \begin{bmatrix} \frac{\partial}{\partial y} \\ \frac{\partial}{\partial s} \end{bmatrix} u(y,s) = \lambda_1 u_{yy} + \lambda_2 u_{ss}.$$

The signs of the eigenvalues will end up determining which of the three equations we can transform the general equation into. There are three cases:

Parabolic One eigenvalue is 0 and the other is non-zero. (Heat equation)

Hyperbolic The eigenvalues have opposite signs. (Wave equation)

Elliptic The eigenvalues have the same sign. (Laplace equation)

Remember that the eigenvalues of A are just solutions to the characteristic equation

$$\lambda^2 - (a+c)\lambda + (ac - b^2) = 0.$$

In particular, $\lambda = \frac{a+c}{2} \pm \frac{\sqrt{(a-c)^2+b^2)}}{2}$; observe that we have one eigenvalue $\lambda = 0$ when $(a+c) = \sqrt{(a-c)^2 + b^2}$, i.e. when

$$4ac = b^2.$$

The quantity $b^2 - 4ac$ is called the *discriminant*. When it is zero, we have a parabolic equation; when it is positive, we have a hyperbolic equation, and when it is negative, we have an elliptic equation. These names come from plane geometry, because the discriminant of the conic section

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

is $B^2 - 4AC$, and the discriminant controls whether the shape of the curve determined by that equation is a parabola, hyperbola, or ellipse in the same way.

Step 2: Eliminate the first-order terms

Consider the equation

$$Au_{xx} + Cu_{tt} + Du_x + Eu_t + Fu = 0.$$

We may assume that one of the constants, A or C, is not zero - as otherwise we would have a first-order equation. By interchanging x and t, if necessary, we may assume that $A \neq 0$. Then dividing the equation by A, we may assume that $A = 1$.

Consider the substitution $u(x,t) = e^{\alpha x + \beta t} w(x,t)$. Using the product rule, compute:

$$u_x = e^{\alpha x+\beta t}[\alpha w(x,t) + w_x(x,t)], \qquad u_t = e^{\alpha x+\beta t}[\beta w(x,t) + w_t(x,t)]$$
$$u_{xx} = e^{\alpha x+\beta t}[\alpha^2 w(x,t) + 2\alpha w_x(x,t) + w_{xx}(x,t)], \quad u_{tt} = e^{\alpha x+\beta t}[\beta^2 w(x,t) + 2\beta w_t(x,t) + w_{tt}(x,t)].$$

Then

$$\begin{aligned}0 &= u_{xx} + Cu_{tt} + Du_x + Eu_t + Fu \\ &= e^{\alpha x+\beta t}\left[(\alpha^2 w + 2\alpha w_x + w_{xx}) + C(\beta^2 w + 2\beta w_t + w_{tt}) + D(\alpha w + w_x) + E(\beta w + w_t) + Fw\right] \\ &= e^{\alpha x+\beta t}\left[w_{xx} + Cw_{tt} + (2\alpha + D)w_x + (2\beta C + E)w_t + (\alpha^2 + C\beta^2 + D\alpha + E\beta + F)w\right]\end{aligned}$$

Of course we can cancel $e^{\alpha x+\beta t}$ since it is nowhere zero. Set $\alpha = -D/2$, and $\beta = -E/2C$ to kill the cross-terms. Of course, this only works when $C \neq 0$. Then the cases are:

i) $(C \neq 0)$ $w_{xx} + Cw_{tt} + F'w = 0$

ii) $(C = 0)$ $w_{xx} + Ew_t + F'w = 0$,

where $F' = \alpha^2 + C\beta^2 + D\alpha + E\beta + F$. We may assume that $E \neq 0$, because otherwise that equation is just an ordinary differential equation.

To summarize our progress so far, using both steps we can reduce our equation to one of the two types above (or to an ordinary differential equation).

Step 3: Normalization

We can do one last change of variables to normalize the various constants. Observe that if $s = \alpha t$, then $\frac{\partial}{\partial t} = \frac{\partial s}{\partial t}\frac{\partial}{\partial s} = \alpha\frac{\partial}{\partial s}$. It follows that $w_t = \frac{\partial}{\partial t}w = \alpha\frac{\partial}{\partial s}w = \alpha w_s$. (This is just a simple case of the linear change of coordinates that we computed in step 1.) Similarly, $w_{tt} = \alpha^2 w_{ss}$. We now apply this kind of transformation to each of the two reduced cases above.

First, for the case $C \neq 0$: by letting $s = \frac{1}{\sqrt{|C|}}t$, we have

$$w_{xx} + Cw_{tt} + F'w = 0 \qquad \mapsto \qquad w_{xx} \pm w_{ss} + F'w = 0.$$

Here, the plus or minus sign depends on whether C is positive or negative, which ultimately depends on the discriminant being negative or positive.

For the case $C = 0$, choosing $s = \frac{1}{-E}w_t$ we have

$$w_{xx} + Ew_t + F'w = 0 \qquad \mapsto \qquad w_{xx} - w_s + F'w = 0.$$

These final equations can finally be put in a *canonical form*:

i) (Heat) $w_s = w_{xx} + c\,w$

ii) (Wave) $w_{ss} = w_{xx} + c\,w$

iii) (Laplace) $w_{xx} + w_{ss} = c\,w$

We have solved each of these equations when $c = 0$. In fact, this form of the heat equation can be further simplified by a dependent variable substitution to get rid of the cw term at the end. However, the other two equations cannot be so reduced. As all of these equations are separable, it is easy to solve them with the methods we have studied thus far. Once we have solved these equations for w, it is possible (though in practice tedious) to invert all of the transformations and recover a solution $u(x, t)$ to the original equation.

Index