

Math 344 - Linear Analysis II

Notes

Math 344 - Linear Analysis II

Notes

Ryan Tully-Doyle
Cal Poly, SLO

January 6, 2021

Contents

1	What you should know	1
1.1	Introduction.	1
1.2	Calculus	2
1.3	Linear algebra	2
1.4	Differential Equations	6
2	Series solutions	8
2.1	Power series	8
2.2	Series solutions at ordinary points	13
2.3	Legendre series	17
2.4	Series solutions at regular singularities	21
2.5	More on Frobenius methods (partial)	26
2.6	Bessel equations	27
2.7	Series as vectors (an introduction to L^2)	32
3	All things Fourier	34
3.1	Periodicity	34
3.2	Complex exponentials and trig functions	35
3.3	Sums of periodic functions	40
3.4	Questions of convergence	40
3.5	First applications	40
3.6	Fourier series and linear algebra	40
3.7	The Fourier transform	40
3.8	Convolution	40
3.9	Applications of convolution	40
4	The Laplace Transform	41
4.1	The Laplace Transform.	41
4.2	Periodic functions.	44
4.3	The Laplace transform of derivatives	44
4.4	The shifting theorems	44
4.5	Heaviside functions	44
4.6	Impulses - the Dirac delta “function”	44
4.7	Convolution	44

Chapter 1

What you should know

1.1 Introduction

This course is the second course in a sequence that uses the language and techniques of linear algebra to study problems in differential equations. The following sections give a broad overview of ideas from calculus, differential equations, and linear algebra that will be used in the course. An overview of ideas you should have seen before might look like

1. Calculus
 - Taylor series
 - Interval of convergence and the ratio test
 - Power series of basic functions
 - Integrating and differentiating power series
 - Improper integrals
2. Differential equations
 - Second order constant coefficient linear differential equations
 - Characteristic equations
 - Reduction of order
 - Harmonic motion (homogeneous and forced motion)
3. Linear Algebra
 - Vector spaces
 - Linear independence
 - Basis of a vector space
 - Inner products (beyond the dot product)
 - Orthogonality
 - Vector projections

You don't need to be an expert, but having seen and used these ideas in the past is going to make this course quite a bit more interesting and productive.

1.2 Calculus

The first of the basic ideas we're going to need is **improper integration**, which is a fancy way of saying an integral with infinity as a limit. Do not be fooled by the notation

$$\int_a^\infty f(x) dx.$$

Infinity is not a number, and the notation for improper integrals is merely shorthand for a limit (this is a recurring theme with infinity). That is, we say that an improper integral in infinity **converges** if

$$\int_a^\infty f(x) dx = \lim_{N \rightarrow \infty} \int_a^N f(x) dx \text{ exists and is finite.}$$

DO NOT treat infinity like a number.

An infinite series is also a limit. (The clue is that infinity is present in the symbol). That is,

$$\sum_{n=0}^\infty a_n = \lim_{N \rightarrow \infty} \sum_{n=0}^N a_n.$$

This is sometimes called the “limit of partial sums”. An infinite series exists if the limit of partial sums exists and is finite. An infinite series $\sum a_n$ is said to **converge absolutely** if $\sum |a_n|$ is a convergent series.

You learned many ways of determining when an infinite series converges in calculus. The most important test for us in this course is the **ratio test**, which states that a series converges (absolutely) if

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < \infty.$$

1.3 Linear algebra

1.3.1 Motivation

One of the most profound ideas of linear algebra is that *any finite dimensional vector space over \mathbb{R} or \mathbb{C} is secretly \mathbb{R}^n or \mathbb{C}^n* . This insight allows us to reduce the study of vector spaces and the maps between them to the study of matrices.

The key idea is that every finite dimensional vector space can be represented in coordinates once we choose a basis. We denote the representation of a vector $v \in V$ with respect to a basis \mathcal{V} by $v_{\mathcal{V}}$. Better yet, that basis can be chosen to be orthonormal by way of the Gram-Schmidt process and the dot product structure of Euclidean space. The coordinatization of V also gives unique representations of linear maps between those spaces.

Theorem 1.3.1 *Let V, W be finite dimensional vector spaces with bases \mathcal{V}, \mathcal{W} . Then any linear map $T : V \rightarrow W$ has a unique matrix representation with respect to \mathcal{V}, \mathcal{W} by*

$$T(x) = Ax$$

with

$$A = [T(v_1)_{\mathcal{W}} \quad \dots \quad T(v_n)_{\mathcal{W}}]$$

Typical examples introduced in a linear algebra course include the space of polynomials of degree less than or equal to n . At the same time, we usually also

get to see a very suggestive example of a useful linear map and the representation of that map in matrix form.

Let P_n denote the space of polynomials of degree $\leq n$. Consider the map $D : P_3 \rightarrow P_2$ defined by

$$D(a_0 + a_1t + a_2t^2 + a_3t^3) = a_1 + 2a_2t + 3a_3t^2.$$

That is, D is the map that takes the derivative of a polynomial. It isn't hard to use the standard basis for P_n to get the matrix representation

$$D(p) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

for the action of D on P_3 .

This example is a good place to begin asking questions about how far we can push finite dimensional linear algebra. The fact that differentiation of polynomials is wonderful - but what else can we apply it to? Nice functions have power series that converge absolutely, and we like to think of an absolutely convergent power series as sort of an “infinite polynomial”. Our intuition might lead us to make a connection with calculus at this point. When we learn to work with power series, we learn that for a convergent power series,

$$\frac{d}{dx} \sum a_n(x-a)^n = \sum na_n(x-a)^{n-1}.$$

In analogy with our example about polynomials above, we're tempted to write, for a function f defined by a convergent power series, that

$$D(f) = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ \vdots & & & & \ddots \end{bmatrix}}_A \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \vdots \end{bmatrix} = a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + \dots$$

This idea is shot through with issues that need to be addressed.

- The object A is some kind of $\infty \times \infty$ matrix. How does that make sense?
- What are the vector spaces that D is mapping between?
- Does the idea of coordinatization still work?
- If it does, what exactly is “ \mathbb{R}^∞ ” supposed to be?
- Do infinite dimensional vector spaces and bases make sense at all?

1.3.2 Inner products

The **dot product** of two vectors in \mathbb{C}^n is

$$x \cdot y = \sum_{i=1}^n \bar{y}_i x_i. \quad (1.3.1)$$

Standard notation for the dot product is $\langle x, y \rangle$ and in \mathbb{C}^n is equivalent to y^*x , where $*$ designates the conjugate transpose of a matrix. The dot product has the following properties:

1. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ conjugate symmetry
2. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ linearity in the first term
3. $\langle x, x \rangle \geq 0$ non-negativity

Once we have the dot product, we can start building the geometry of \mathbb{C}^n . First, note that

$$\|x\|^2 = \langle x, x \rangle = \sum_{i=1}^n |x_i|^2. \quad (1.3.2)$$

Motivated by the real case, we say that two vectors x, y are **orthogonal** and write $x \perp y$ if $\langle x, y \rangle = 0$.

Another important inequality is indicated by the relationship between angles and the dot product in \mathbb{R}^n , where we have

$$\langle x, y \rangle = \|x\| \|y\| \cos \theta,$$

where θ is the angle between the vectors. While the idea of “angle” doesn’t make sense in \mathbb{C}^n (at least in the same way), we still have the **Cauchy-Schwarz** inequality

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \quad (1.3.3)$$

Orthogonality also underlies the vector version of the **Pythagorean theorem**,

$$\|x\|^2 + \|y\|^2 = \|x + y\|^2 \iff x \perp y. \quad (1.3.4)$$

Finally, it would be remiss to leave out the single most important inequality in mathematics, our old friend the **triangle inequality**, which in vector terms can be expressed

$$\|x + y\| \leq \|x\| + \|y\| \quad (1.3.5)$$

Because finite dimensional vector spaces have representations in coordinates as \mathbb{R}^n or \mathbb{C}^n , all finite dimensional vector spaces carry the geometric structure delineated above.

1.3.3 Basis and coordinates

Let V be a vector space over a field \mathbb{F} . Recall that a (finite) set of vectors $S \subset V$ is **linearly independent** if only the trivial solution exists for the equation

$$0 = \sum_{\mathcal{I}} c_i v_i. \quad (1.3.6)$$

A set S of vectors in V is said to **span** V if every vector in V can be realized as a linear combination of vectors in S . That is, given $v \in V$, there exist coefficients c_i so that

$$v = \sum_{\mathcal{I}} c_i v_i.$$

A basis \mathcal{V} for V is a subset of V so that \mathcal{V} is linearly independent and \mathcal{V} spans V . It is a major result that every vector space has a basis. The full result requires the invocation of [Zorn’s Lemma](#) or other equivalents of the [axiom of choice](#) and will not be proven here. (A nice argument can be found [here](#).) Our interest is in modeling vector spaces that carry the logic and structure of Euclidean space. The **dimension** of V is the order of a basis \mathcal{V} . If the basis has a finite number of elements, say n , then V is called finite dimensional. In particular, (and clearly providing motivation for the definition), $\dim \mathbb{R}^n = n$.

Suppose that V is a finite dimensional vector space with a basis \mathcal{V} . Let v be a vector in V . Then the **coordinates of v with respect to \mathcal{V}** are the constants c_i so that $v = \sum_{\mathcal{I}} c_i v_i$. These coordinates are *unique* once we have fixed a basis \mathcal{V} . That is, we have a bijective relationship between the vectors

$v \in V$ and the coordinate representations $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^n$. In \mathbb{F}^n , the coordinate representation of a vector is straightforward to compute using the dot product.

Theorem 1.3.2 *Let e_1, \dots, e_m be an orthonormal basis for \mathbb{F}^m and $v \in \mathbb{F}^n$. Then the n th coordinate of v with respect to the basis is $\langle v, e_n \rangle$, and the expansion of v with respect to the basis is*

$$v = \sum_1^m \langle v, e_i \rangle e_i.$$

Furthermore, we can use the coordinate representation to write representing matrices for linear functions $T : V \rightarrow W$. Suppose that V, W are vector spaces of dimension m, n respectively over \mathbb{F} . Then

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow i_V & & i_W^{-1} \uparrow \\ \mathbb{F}^m & \xrightarrow{A} & \mathbb{F}^n \end{array}$$

where A is the matrix that represents T and i is the natural bijection- the coordinatization - between V, W and $\mathbb{F}^m, \mathbb{F}^n$ respectively. We should note that matrix multiplication is defined so that

$$\begin{array}{ccccc} U & \xrightarrow{S} & V & \xrightarrow{T} & W \\ \downarrow i_U & & & & i_W^{-1} \uparrow \\ \mathbb{F}^r & \xrightarrow{A} & \mathbb{F}^m & \xrightarrow{B} & \mathbb{F}^n \end{array}$$

reduces to the diagram

$$\begin{array}{ccc} U & \xrightarrow{S \circ T} & W \\ \downarrow i_U & & i_W^{-1} \uparrow \\ \mathbb{F}^r & \xrightarrow{BA} & \mathbb{F}^n \end{array}$$

That is, the representing matrix of a composition is the product of the representing matrices of the functions.

Any basis of a vector space can be replaced with an equivalent basis of orthonormal vectors - the algorithm for creating an orthonormal basis from a basis is called the **Gram-Schmidt process**.

1.3.4 Operators

When a linear function maps V into itself, special things happen. First, the matrix that represents $T : \mathbb{F}^n \rightarrow \mathbb{F}^n$ is square. There are a large number of equivalences between the structure of square matrices, linear maps, and sets of vectors. Many of these are captured in the **invertible matrix theorem**, one

of the central objects of study in elementary linear algebra.

Theorem 1.3.3 Invertible matrix theorem. *Let A be an $n \times n$ matrix. If any of the following conditions hold, all of them do. If any of them are false, they are all.*

1. A is invertible.
2. A row reduces to the identity matrix I .
3. A has n pivot positions.
4. $\text{rank} A = n$.
5. The equation $A\vec{x} = \vec{0}$ has only the trivial solution.
6. The columns of A are linearly independent.
7. The function $T(\vec{x}) = A\vec{x}$ is one-to-one.
8. The equation $A\vec{x} = \vec{v}$ is consistent for all $\vec{v} \in \mathbb{F}^n$.
9. The columns of A span \mathbb{F}^n .
10. The function $T(\vec{x}) = A\vec{x}$ is onto.
11. There is a matrix C so that $CA = I$.
12. There is a matrix D so that $AD = I$.
13. A^T is invertible.
14. $\det A \neq 0$.

Operators contain more information than the invertibility of the functions that they represent. For the following discussion, let us fix a basis of a vector space V and let A be the matrix that represents a function $T : V \rightarrow V$. A scalar λ and a vector v are said to be an **eigenpair** for A if

$$Av = \lambda v.$$

It is straightforward to see that the set of all vectors v for which the eigenvector equation holds is a subspace of V , called the **eigenspace** associated with λ . The eigenspaces of the matrix A are its **invariant subspaces**, which is to say that a vector in an eigenspace is mapped by A to the same eigenspace. It turns out that knowing the invariant subspaces of A are often enough to completely characterize A . If A is $n \times n$ and A has n linearly independent eigenvectors (that is, one can find a basis of \mathbb{F}^n consisting of eigenvectors of A), then

$$A = SDS^{-1},$$

where S is a matrix of eigenvectors and D is a diagonal matrix of the associated eigenvalues (including repetition of course). (One should think of S as a change of basis matrix under which the operator A becomes diagonal.)

Many operators are not diagonalizable, even very simple ones. For example, $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ only has a one-dimensional eigenspace. Diagonalizability is so useful that we give characterizations of those operators a special name, the **Spectral Theorem**. An operator on a real vector space is called **symmetric** if $A^T = A$. An operator on a complex vector space is called **Hermitian** (or conjugate symmetric) if $A^* = \overline{A^T} = A$. One of the major theorems of

elementary linear algebra is that such operators are diagonalizable and that there exists an orthonormal basis of eigenvectors for V .

Theorem 1.3.4 *Let A be an $n \times n$ real (complex) matrix. Then A is diagonalizable with respect to an orthonormal basis of eigenvectors if and only if A is symmetric (hermitian).*

For complex operators, one can say more. A is called **normal** if $AA^* = A^*A$. One reason that complex vector spaces are so much nicer than real vector spaces is that normal operators turn out to have orthonormal diagonalizations.

Theorem 1.3.5 Complex (finite) spectral theorem. *Let A be an operator on a finite dimensional Hilbert space V . Then A is normal if and only if A can be diagonalized with respect to an orthonormal basis of eigenvectors for V .*

1.4 Differential Equations

A homogeneous second order linear differential equation is of the form

$$y'' + p(x)y' + q(x)y = 0.$$

The most important differential equations of this type are **linear constant coefficient differential equations**, usually introduced in the second order. For example,

$$y'' + 3y' + 2y = \cos x$$

is such an equation. More generally, these equations have the form

$$a_n y^{(n)}(x) + \dots + a_1 y'(x) + a_0 y(x) = f(x).$$

(This can be written

$$\mathcal{F}(y) = f(x)$$

where \mathcal{F} is the operator that takes y to the differential expression on the left hand side.) These equations are important because they are the general equations that describe harmonic motions - vibration, oscillation, and rotations.

The method of solution we develop in earlier courses is to find the solution by working in steps. First, we solve the associated **homogeneous equation** in order to identify any functions that are sent to the zero function by the differential operator. This is called the **homogeneous solution** y_h . To do so, we use the method of the characteristic equation or the method of annihilators. In either case, we have to factor potentially high degree polynomials.

Example 1.4.1 To solve the homogeneous equation

$$y'' + 3y' + 2y = 0,$$

we first consider the characteristic equation

$$m^2 + 3m + 2 = 0$$

which factors as

$$(m + 2)(m + 1) = 0$$

with solutions $m = -2, m = -1$. Then the principle of superposition tells us that the homogeneous solution is

$$y_h = c_1 e^{-2x} + c_2 e^{-x}.$$

□

Now, for the second step, we need to find a **particular solution** that we can feed into the differential equation so that $\mathcal{F}(y_p) = f(x)$. We learn several methods for this in first courses that consider differential equations, usually including the method of undetermined coefficients and the method of variation of parameters. In both cases, we're limited to either forcing functions nice enough that we can guess the form of the solution (in the case of undetermined coefficients) or functions that we can integrate (in the case of variation of parameters). However, in physical systems we often want to consider driving or forcing functions $f(x)$ that represent physical conditions that don't have nice integrals: for example, switches that turn on and off, or impulses that are applied to the system instantaneously (such as a strike). While the methods above can be adapted to deal with these, it is easier to **transform** the problem into a different mathematical setup that will simplify the process of dealing with these common physical situations.

The model problem to keep in mind is the following mass-spring type setup:

$$y'' + ay' + by = F$$

where F might switch on and off or act on the system instantaneously (and so cannot be a continuous function). By transforming the problem into an equivalent domain, we can deal with these fundamental situations very easily.

Chapter 2

Series solutions

2.1 Power series

2.1.1 Motivation

One of the limitations of the elementary techniques of differential equations is the extremely limited number of types of equations that can be solved. Once we are dealing with equations that are second order or higher, we are pretty much restricted to equations with constant coefficients ($ay'' + by' + cy = 0$), or equations in special forms (like the Cauchy-Euler equations $ax^2y'' + bxy' + cy = 0$).

Example 2.1.1 Airy's equation. A seemingly simple equation that cannot be solved with elementary techniques is **Airy's equation**, which is of the form

$$y'' - xy = 0.$$

□

One problem that we run into is that many functions cannot be expressed as combinations of the elementary functions that we learn in calculus (for example, $f(x) = \int e^{x^2} dx$), and these functions are frequently the solutions to differential equations like the one in the example above. We turn, as we did in calculus, to **power series** representations of functions. Power series are one of the most important ideas developed in introductory calculus (though it is often the case that WHY you are learning about them isn't immediately obvious). We'll begin by recalling the definition of a power series, the question of when and where such a series converges, and some basic operations that we can perform

2.1.2 Definition of power series and convergence

Definition 2.1.2 An infinite series of the form

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n$$

is called a power series centered at $x = x_0$.

Because we can always make the substitution $u = x - x_0$, it is typically sufficient to consider series centered at 0,

$$\sum_{n=0}^{\infty} a_n x^n. \tag{2.1.1}$$

◇

A power series **converges** at a point x if the sequence of partial sums converges - that is,

$$\lim_{N \rightarrow \infty} \sum_{n=0}^N a_n x^n$$

exists and is equal to a finite value. The **interval of convergence** of a power series (centered at 0) is the largest interval of the form $I = (-r, r)$ for which the series converges for each $x \in (-r, r)$. The quantity r is called the **radius of convergence**. There are three possibilities for the value of r .

Theorem 2.1.3 For a series of the form given in (2.1.1), exactly one of the following holds.

$$1. r = 0 : \sum_{n=0}^{\infty} a_n x^n \text{ converges only for } x = 0.$$

$$2. r = \infty : \sum_{n=0}^{\infty} a_n x^n \text{ converges for every } x \in \mathbb{R}.$$

$$3. r \text{ is finite: There is a constant } r > 0 \text{ so that } \sum_{n=0}^{\infty} a_n x^n \text{ converges for } x < |r| \text{ and diverges for } x > |r|.$$

The tool from calculus that is most useful in computing a radius of convergence for a power series is the **ratio test** (the reason is that only a single x will remain after computing the ratio).

Definition 2.1.4 For the power series $\sum_{n=0}^{\infty} a_n x^n$, the radius of convergence r is given by $r = \frac{1}{L}$ where L is the limit

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L.$$

If $L = 0$, then $r = \infty$ (that is, the series converges for all values of x). If $L = \infty$, then $r = 0$ (that is, the series only converges trivially at $x = 0$). ◇

Example 2.1.5 Finding the radius of convergence. Compute the radius of convergence of

$$\sum_{n=0}^{\infty} \frac{2^{n+1}}{5^n} x^n.$$

We compute the limit

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \frac{2^{n+2}}{5^{n+1}} \cdot \frac{5^n}{2^{n+1}} = \frac{2}{5},$$

and so $r = \frac{5}{2}$. □

2.1.3 Algebra and calculus of power series

The beauty and utility in power series is that when they converge, they can be treated almost like giant polynomials. That is, we can perform algebra and calculus operations on convergent power series. To do so, first we can make a function out of a power series by restricting the domain to the interval of

convergence (so that the function is well-defined). That is, if r is the radius of convergence for $\sum_{n=0}^{\infty} a_n x^n$, then we can define

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad x \in (-r, r).$$

Before we discuss operations involving power series, we note a fundamental fact (related to the identical fact about polynomials) - two power series are equal if and only if their corresponding coefficients are equal. That is,

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n \text{ if and only if } a_n = b_n \text{ for all } n.$$

When we combine power series, we have to make sure that we work on a domain where both functions are defined - that is, if f and g have radii of convergence r_1, r_2 respectively, then the radius of convergence of an algebraic combination of f, g will be $r = \min\{r_1, r_2\}$.

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $g(x) = \sum_{n=0}^{\infty} b_n x^n$ with common radius of convergence r .

1. Addition: $f(x) + g(x) = \sum_{n=0}^{\infty} (a_n + b_n) x^n$.

2. Scalar multiplication: $cf(x) = \sum_{n=0}^{\infty} (ca_n) x^n$

3. Multiplication of series: $f(x)g(x) = \sum_{n=0}^{\infty} c_n x^n$ where the coefficients

c_n are computed by $c_n = \sum_{k=0}^n a_{n-k} b_k$ (this is essentially the result of a giant distribution and collection of like terms)

Division can be defined as well (as long as the function in the denominator is not zero), but is used less often in what we will be studying

One of our first applications of power series will be to plug them into equations, in order to figure out the coefficients. If we know the coefficients of a function given by a power series, then we know the function. Typically, finding the coefficients will involve solving a **recurrence relation**, which is a recursive process for determining coefficients from some number of known coefficients.

Example 2.1.6 Solving a recurrence relation. Assume that the coefficients in the power series

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

satisfy the equation

$$\sum_{n=1}^{\infty} n a_n x^n - \sum_{n=0}^{\infty} a_n x^{n+1} = 0.$$

Find a formula that gives any a_n in terms of a_0 .

First, to combine power series, we'll need both the powers of the like terms to match, and the series to have the same number of terms. Notice that if we replace $n \rightarrow n + 1$ in the second series, and we change the range of the index accordingly, we get

$$\sum_{n=1}^{\infty} na_n x^n - \sum_{n=1}^{\infty} a_{n-1} x^n = 0.$$

As the powers of the corresponding terms match, and the range of indices match, we can combine the series using addition to get

$$\sum_{n=1}^{\infty} (na_n - a_{n-1})x^n = 0.$$

Because a convergent power series can only be equal to zero if the coefficients are all zero, we get the recurrence relation

$$na_n - a_{n-1} = 0, \quad n = 1, 2, 3, \dots,$$

which can be written

$$a_n = \frac{1}{n}a_{n-1}, \quad n = 1, 2, 3, \dots$$

This gives a family of equations that we can use to find each coefficient in terms of a_0 and eventually guess a formula that won't require us to work recursively.

$$n = 1 : a_1 = a_0.$$

$$n = 2 : a_2 = \frac{1}{2}a_1 = \frac{1}{2}a_0.$$

$$n = 3 : a_3 = \frac{1}{3}a_2 = \frac{1}{3 \cdot 2}a_0.$$

$$n = 4 : a_4 = \frac{1}{4}a_3 = \frac{1}{4 \cdot 3 \cdot 2}a_0.$$

We can guess that the general formula for a_n is

$$a_n = \frac{1}{n!}a_0.$$

(Notice that I haven't claimed to have proved this, because that would require mathematical induction. For our purposes in this course, if a pattern looks to hold, we will assume that it continues indefinitely.)

Under the assumption that our formula for a_n holds for all n , we can say

$$f(x) = a_0 \sum_{n=0}^{\infty} \frac{1}{n!} x^n,$$

which you might notice is the Taylor series at 0 for the function $f(x) = a_0 e^x$. \square

Just like algebra, we can perform calculus on power series in the obvious way (treating them like giant polynomials). Suppose that f is defined by a power series with domain equal to the interval of convergence, so that

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad |x| < r.$$

Then we can differentiate f term by term to any order of derivative (after all, we're never going to run out of x es):

$$f'(x) = \sum_{n=1}^{\infty} na_n x^{n-1}$$

$$f''(x) = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}$$

and so on. Power series can also be integrated, which is straightforward to write a formula for. (You might consider doing it as an exercise.)

2.1.4 Analytic functions (a connection with complex analysis)

A function with a power series defined on some interval (a, b) is particularly nice because the power series implies the existence of derivatives of all orders. Such functions are called **smooth**. (Note that smooth functions need not have a convergent power series, which is a weird but deep fact of real analysis!). Power series representations are so special, we name functions that have them as a special class.

Definition 2.1.7 A function f with a convergent power series representation on some non-trivial interval of convergence (a, b) is called an **analytic** function on (a, b) . \diamond

Analytic functions are the nicest behaved functions in calculus and its applications. One of the major theorems of calculus is that if a function is analytic, then it has a unique power series representation, called the **Taylor series** for f . Even better, the series can be computed!

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n.$$

The functions of introductory calculus are nearly all analytic, many on the entire real line. You should be familiar with the power series expansions of those functions derived from exponential functions.

1. $e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots = \sum_{n=0}^{\infty} \frac{1}{n!}x^n$
2. $\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \dots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}x^{2n}$
3. $\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \dots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1}$

Because we can combine power series using algebra to get new power series, we can do the same with the analytic functions that they represent. That is, if f, g are analytic, so are algebraic combinations of them.

Theorem 2.1.8 If f, g are analytic functions on a common interval (a, b) , then so are $f + g, f - g, fg$, and $\frac{f}{g}$ (as long as $g(x) \neq 0$).

The most useful examples of analytic functions are polynomials! (From this perspective, polynomials are just finite power series.) Polynomials are analytic at every point, and so therefore are $f + g$, $f - g$ and fg . Things are a bit more complicated in the case of rational functions $\frac{f}{g}$, because the zeroes of g determine the radius of convergence of the quotient.

Here we come across one of the first places that the analysis of analytic functions is actually secretly taking place in the complex plane. Consider the function

$$f(x) = \frac{1}{1+x^2}.$$

$1+x^2$ doesn't possess any real zeroes. Does this mean that the power series for f converges everywhere? Let's see. Using the fact that $\frac{1}{1-x} = 1 + x + x^2 + \dots$ for $|x| < 1$, we can show that

$$\frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = 1 + (-x^2) + (-x^2)^2 + (-x^2)^3 + \dots = \sum_{n=0}^{\infty} (-1)^n x^{2n},$$

as long as $|-x^2| < 1$ which is just $|x| < 1$. So even though $1+x^2$ has no real zeros, the power series for f only had radius of convergence 1! What's going on?

It turns out that when we're looking at quotients of analytic functions, we need to consider all of the zeros of the denominator, real and complex. For our example, $1+x^2=0$ has solutions $\pm i$, which are a distance of 1 away from the origin in the complex plane (0 on the real line). Since the closest zero to the center of our power series is 1 unit away, the radius of convergence is 1.

Theorem 2.1.9 *If $p(x), q(x)$ are polynomials and $q(x_0) \neq 0$, then the radius of convergence of the power series representation is the distance from x_0 to the closest root (real or complex) of q .*

Example 2.1.10 Radius of convergence for a rational function. Compute the radius of convergence of the power series centered at $x = 0$ of

$$f(x) = \frac{x^2}{(x+3)(x^2+3)}.$$

The zeroes of q are $x_1 = -3, x_2 = i\sqrt{3}, x_3 = -i\sqrt{3}$. Using the distance formula, we can compute that the distances to the origin are $d_1 = 3, d_2 = \sqrt{3}, d_3 = \sqrt{3}$, and so the radius of convergence is $r = \sqrt{3}$. \square

2.2 Series solutions at ordinary points

2.2.1 Ordinary points

Consider the second order linear differential equation

$$y'' + p(x)y' + q(x)y = 0.$$

The fact that we have functions as “coefficients” for the terms means that our previous tools do not apply. Instead, we'll try to discover solutions using power series and recurrence relations. It will be much easier to do this if p and q are well-behaved where we center our power series.

Definition 2.2.1 A point $x = x_0$ is an **ordinary point** for the differential equation

$$y'' + p(x)y' + q(x)y = 0$$

if p and q are analytic at $x = x_0$. If a point is not ordinary, then it is called a **singular point** of the equation. \diamond

Example 2.2.2 Finding ordinary points. Find the ordinary points of the differential equation

$$y'' + \frac{1}{x^2 - 1}y' + \frac{1}{x - 3}y = 0.$$

Because p, q fail to be analytic at $x = \pm 1, 3$ the equation is singular at those three points. Every other point is an ordinary point. \square

2.2.2 Power series solutions at ordinary points

We now give an example of solving a differential equation with series solutions. We'll start with a problem that we already know the solution to, before moving on to solving Airy's equation. It will be useful to recall the expressions for the derivatives of convergent power series. If $y = \sum_{n=0}^{\infty} a_n x^n$, then

$$y' = \sum_{n=1}^{\infty} n a_n x^{n-1},$$

$$y'' = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2}.$$

Example 2.2.3 Finding series solutions. Find the general solution to the differential equation $y'' + y = 0$ using series methods.

We expect to find two linearly independent solutions, and we expect the general solution to be of the form $y = c_1 y_1 + c_2 y_2$ for arbitrary constants c_1, c_2 because we have a second order equation. Notice that $p = 0$ and $q = 1$ which are analytic at every point, so we will work at the center $x = 0$ and assume that our solution is of the form $y = \sum_{n=0}^{\infty} a_n x^n$.

On taking derivatives and plugging into the differential equation, we get

$$\sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} + \sum_{n=0}^{\infty} a_n x^n = 0.$$

We want to combine these into a single sum and use a recurrence relation to find the solutions, but we need the terms and the index ranges to match. If we shift the index on the first sum by letting $n - 2 \rightarrow n$, we get

$$\sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} x^n + \sum_{n=0}^{\infty} a_n x^n = 0.$$

Now we can combine the series to get

$$\sum_{n=0}^{\infty} [(n+2)(n+1) a_{n+2} + a_n] x^n = 0$$

which gives the family of coefficient equations

$$(n+2)(n+1) a_{n+2} + a_n = 0$$

which is convenient to rearrange as

$$a_{n+2} = -\frac{1}{(n+2)(n+1)} a_n, \quad n = 0, 1, 2, \dots$$

Notice that we need two pieces of information to determine all of the information contained in this recursion: $n = 0$ starts a chain that gives the coefficients for the even values of n , and $n = 1$ starts a chain that gives that information for the odd values (corresponding to the two separate independent solutions). We consider each family separately.

$$\begin{aligned} n = 0 : a_2 &= -\frac{1}{2 \cdot 1} a_0 = -\frac{1}{2!} a_0 \\ n = 2 : a_4 &= -\frac{1}{4 \cdot 3} a_2 = \frac{1}{4 \cdot 3 \cdot 2 \cdot 1} a_0 = \frac{1}{4!} a_0 \\ n = 4 : a_6 &= -\frac{1}{6 \cdot 5} a_4 = -\frac{1}{6!} a_0. \end{aligned}$$

That is, the even terms seem to be of the form $a_{2k} = \frac{(-1)^k}{(2k)!} a_0$.

For the odd values of n , we get

$$\begin{aligned} n = 1 : a_3 &= -\frac{1}{3 \cdot 2} a_1 = -\frac{1}{3!} a_1 \\ n = 3 : a_5 &= -\frac{1}{5 \cdot 4} a_3 = \frac{1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} a_1 = \frac{1}{5!} a_1 \\ n = 5 : a_7 &= -\frac{1}{7 \cdot 6} a_5 = -\frac{1}{7!} a_1. \end{aligned}$$

which looks like $a_{2k+1} = \frac{(-1)^k}{(2k+1)!} a_1$.

Now that we know the coefficients of the power series, we can write

$$\begin{aligned} y &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots \\ &= a_0 + a_1 x - \frac{1}{2!} a_0 x^2 - \frac{1}{3!} a_1 x^3 + \dots \\ &= a_0 \left(1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \dots\right) + a_1 \left(x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \dots\right) \end{aligned}$$

which allows us to write $y = a_0 y_1 + a_1 y_2$ where

$$y_1(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \quad y_2 = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}.$$

An easy application of the ratio test shows that each of these series converges for all values of x (that is, the radius of converge is ∞).

Finally, we need to verify that these series are linearly independent. With two objects, this will be true as long as the functions aren't scalar multiples of each other, which is indeed the case.

(We might note that none of this is surprising in this case because the series in question have very nice closed forms you should recognize on sight!) \square

At an ordinary point, we will always follow the steps in the example above.

1. Assume that a solution of the form $y = \sum_{n=0}^{\infty} a_n x^n$ exists.
2. Plug the series into the equation to get a recurrence relation and find the coefficients in terms of a_0 and a_1 .
3. Use the ratio test to find the radius of convergence (which is where our solution is valid).
4. Verify that the solutions are linearly independent.

Theorem 2.2.4 *Let p, q be analytic functions on some common interval $|x - x_0| < r$ centered at x_0 . Then the general solution to the second order differential equation*

$$y'' + py' + qy = 0$$

can be represented as power series centered at x_0 on the same interval. The coefficients in the series can be described in terms of the initial data a_0, a_1 , and the general solution can be rearranged into the form

$$y = a_0 y_1 + a_1 y_2$$

where y_1, y_2 are linearly independent.

We conclude this chapter with the solution of Airy's equation.

Example 2.2.5 Airy's equation. Solve the equation

$$y'' - xy = 0.$$

Since $p = 0$ and $q = x$, every point is an ordinary point. So we assume that a series solution exists of the form $y = \sum_{n=0}^{\infty} a_n x^n$. Plugging into the equation, we get

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - x \sum_{n=0}^{\infty} a_n x^n = 0.$$

Multiplying through, we get

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - \sum_{n=0}^{\infty} a_n x^{n+1} = 0.$$

First, we'll make the powers match, then we'll shift indicies as necessary.

$$\sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2} x^n - \sum_{n=1}^{\infty} a_{n-1} x^n = 0.$$

Now that the powers of x match, the easiest way to make the index ranges match is to pop the first term off of the first sum.

$$2a_2 + \sum_{n=1}^{\infty} (n+2)(n+1)a_{n+2} x^n - \sum_{n=1}^{\infty} a_{n-1} x^n = 0.$$

Before we combine, notice that right away we can see that $a_2 = 0$. This is going to have an interesting effect on the resulting series.

$$2a_2 + \sum_{n=1}^{\infty} [(n+2)(n+1)a_{n+2} - a_{n-1}] x^n = 0.$$

Setting the coefficients to be 0, we get the family of relations

$$a_{n+2} = \frac{1}{(n+2)(n+1)} a_{n-1}.$$

Because there is a separation of three index elements between the terms in the recursion, there will be three families of equations. Immediately, we can see that because $a_2 = 0$, the recursion implies that $a_5, a_8, a_{11}, \dots = 0$ as well. (From the theorem, we know that we should be able to get all the information we need in terms of a_0 and a_1 .)

For the set of coefficients beginning with a_0 , we have

$$\begin{aligned} n = 1 : a_3 &= \frac{1}{3 \cdot 2} a_0 \\ n = 4 : a_6 &= \frac{1}{6 \cdot 5} a_3 = \frac{1}{6 \cdot 5 \cdot 3 \cdot 2} a_0 \\ n = 7 : a_9 &= \frac{1}{9 \cdot 8} a_6 = \frac{1}{9 \cdot 8 \cdot 6 \cdot 5 \cdot 3 \cdot 2} a_0. \end{aligned}$$

There is a clear pattern, though it is annoying to write in closed form.

For the set of coefficients beginning with a_1 , we have

$$\begin{aligned} n = 2 : a_4 &= \frac{1}{4 \cdot 3} a_1 \\ n = 5 : a_7 &= \frac{1}{7 \cdot 6} a_4 = \frac{1}{7 \cdot 6 \cdot 4 \cdot 3} a_1 \\ n = 8 : a_{10} &= \frac{1}{10 \cdot 9} a_7 = \frac{1}{10 \cdot 9 \cdot 7 \cdot 6 \cdot 4 \cdot 3} a_1. \end{aligned}$$

Then the general solution to the equation is

$$y = a_0 y_1(x) + a_1 y_2(x)$$

where

$$\begin{aligned} y_1(x) &= 1 + \frac{1}{3 \cdot 2} x^3 + \frac{1}{6 \cdot 5 \cdot 3 \cdot 2} x^6 + \dots \\ y_2(x) &= x + \frac{1}{4 \cdot 3} x^4 + \frac{1}{7 \cdot 6 \cdot 4 \cdot 3} x^7 + \dots \end{aligned}$$

These power series are convergent everywhere, obviously linearly independent, and unfortunately have no closed form; that is, series representations is the best that we can do.

```
var( 'z_y_x'); y = 1 + 1/6 *x^3 + 1/180 *x^6
z = x + 1/12 *x^4 + 1/504 *x^7
a = plot([],title="(Approx)_Solutions_to_Airy's_equation")
a += plot(y, x, (-2,2), color='red')
a += plot(z, x, (-2,2))
show(a)
```

□

2.3 Legendre series

2.3.1 Legendre's equation

Physics and applied mathematics provide a wealth of differential equations, some of which are sufficiently common to carry special names, notably

$$(1 - x^2)y'' - 2xy' + c(c + 1)y = 0, \quad (\text{Legendre's equation})$$

$$y'' - 2xy' + 2cy = 0, \quad (\text{Hermite's equation})$$

$$(1 - x^2)y'' - xy' + c^2y = 0 \quad (\text{Chebyshev's equation})$$

where c is some parameter that can take on any real value. All of these equations have an ordinary point at $x = 0$, and so have series solutions of the form $y = \sum_{n=0}^{\infty} a_n x^n$. The solutions to these equations for different values of c

have special properties, particularly when c is an integer.

In this section, we'll investigate the solutions and properties of the Legendre equation. First, note that the equation as written is not in our standard form. We'll need to rewrite it to get an estimate for the interval of convergence of our series solution (though it will be convenient to leave it in its original form when we attempt to solve it.) In the form

$$y'' - \frac{2x}{1-x^2}y' + \frac{c(c+1)}{1-x^2}y = 0,$$

we can see that p, q are analytic at the very least on the interval $(-1, 1)$, as the power series for $\frac{1}{1-x^2}$ has radius of convergence 1.

Armed with the knowledge that our solutions exist, we can follow the usual technique to derive the solutions.

Checkpoint 2.3.1 Derive the general solutions to the Legendre equation using recurrence relations.

In general, the solutions to Legendre's equation for a fixed constant c are of the form

$$y = y_1(x) + y_2(x)$$

with

$$y_1(x) = a_0 \left[1 - \frac{c(c+1)}{2}x^2 + \frac{(c-2)c(c+1)(c+3)}{4!}x^4 - \frac{(c-4)(c-2)c(c+1)(c+3)(c+5)}{6!}x^6 + \dots \right]$$

$$y_2(x) = a_1 \left[x - \frac{(c-1)(c+2)}{3!}x^3 + \frac{(c-3)(c-1)(c+2)(c+4)}{5!}x^5 + \dots \right]$$

2.3.2 Legendre polynomial

It turns out that the most important case of the Legendre equation occurs when the constant c is an integer, which we emphasize by calling it $c = N$. In that case, the recursion that determines the solution to the equation is

$$a_{n+2} = -\frac{n(n+1) - N(N+1)}{(n+1)(n+2)}a_n.$$

Since N is fixed, this means that $a_{N+2} = 0$, and moreover that the subsequent terms in the recursion are zero as well:

$$a_{N+2} = a_{N+4} = a_{N+6} = \dots = 0.$$

That is, either y_1 or y_2 is a finite power series - a polynomial - depending on if N is even or odd. These polynomials turn out to be quite useful.

Definition 2.3.2 Let N be a nonnegative integer. The **Legendre polynomial of degree N** , denoted $P_N(x)$, is the polynomial solution to

$$(1-x^2)y'' - 2xy' + N(N+1)y = 0,$$

normalized so that $P_N(1) = 1$.

◇

Example 2.3.3 The first few Legendre polynomials are easy to write down. Remember that each polynomial is normalized to have $P_N(1) = 1$.

$$N = 0 : y_1(x) = a_0 \Rightarrow P_1(x) = 1.$$

$$N = 1 : y_2(x) = a_1x \Rightarrow P_2(x) = x.$$

$$N = 2 : y_1(x) = a_0(1 - 3x^2) \Rightarrow P_2(x) = \frac{1}{2}(3x^2 - 1).$$

□

For larger values of N , we need a better method to produce these polynomials. Fortunately for us, computer algebra systems have functions that produce these polynomials very quickly. (The following is in the mathematical language Sage. Similar functions exist in MATLAB, Mathematica, Octave, python, etc.)

```
var('x')
legendre_P(2, x)
```

If we wish to proceed by hand, there is an explicit formula for generating $P_N(x)$ known as Rodrigues's formula:

$$P_N(x) = \frac{1}{2^N N!} \frac{d^N}{dx^N} (x^2 - 1)^N, \quad N = 0, 1, 2, \dots$$

2.3.3 Legendre polynomials and vector spaces

Here we reach an instance of one of the most important larger themes of mathematical analysis - we can view functions with nice properties as vectors sitting inside an infinite dimensional vector space. With certain modifications, these spaces act very similarly to the familiar Euclidean spaces \mathbb{R}^n . One of the most important features of vector spaces is that vectors can be decomposed into orthogonal pieces with respect to some inner product in terms of a basis. For example, we can think of the Taylor series for a function as expressing that function in terms of the vectors $1, x, x^2, \dots$, where the coefficients $\frac{f^{(n)}(a)}{n!}$ are really the coordinates of the vector.

The continuous functions on the interval $[-1, 1]$ are denoted $C^0[-1, 1]$. (We're using symmetric intervals in this case because the Legendre polynomials are even or odd.) This space is a (infinite dimensional) inner product space with the dot product defined as

$$f \cdot g = \langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx.$$

Following the analogy with \mathbb{R}^n , we define two vectors in an inner product space to be **orthogonal** if $\langle f, g \rangle = 0$. Beyond a notion of "angle", the inner product also gives a way to measure length with the **norm** defined in this case by

$$\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_{-1}^1 f(x)^2 dx}.$$

Theorem 2.3.4 *The Legendre polynomials are mutually orthogonal; that is,*

$$\langle P_j, P_k \rangle = \int_{-1}^1 P_j(x)P_k(x) dx = 0 \text{ if } j \neq k.$$

Thus, the polynomials P_0, P_1, \dots, P_n are an orthogonal basis for the space of polynomials of degree n or less.

Proof. See Annin, Goode, p. 744. ■

Furthermore, one can show that for a fixed integer j that

$$\|P_j\| = \sqrt{\frac{2}{2j+1}}.$$

One of the most important properties of an orthogonal basis for an inner product space is that generic vectors can be decomposed into orthogonal pieces. Suppose that we have the set of Legendre polynomials P_0, \dots, P_N as an orthogonal basis for the polynomials of degree N or less. Then we can compute the expansion of a polynomial in terms of Legendre polynomials by the formula

$$p(x) = \sum_{i=0}^N a_i P_i = \sum_{i=0}^n \frac{\langle p, P_i \rangle}{\|P_i\|^2} P_i.$$

This should look familiar, as we are finding the coefficients a_i by using the vector projection formula $\text{proj}_v u = \frac{u \cdot v}{\|v\|^2} v$. Since we know that $\|P_i\|^2 = \frac{2}{2i+1}$, we get the specific expansion

$$p(x) = \sum_{i=0}^N a_i P_i(x), \text{ where } a_i = \frac{2i+1}{2} \int_{-1}^1 p(x) P_i(x) dx.$$

The real power of the expansion idea is that it holds for general functions, not just polynomials, under the additional assumption that $f \in C^1(-1, 1)$; that is, both f, f' are continuous.

Theorem 2.3.5 *Suppose that f, f' are continuous on $(-1, 1)$. Then f has a Legendre series expansion on $(-1, 1)$ given by*

$$f(x) = \sum_{i=0}^{\infty} a_i P_i(x)$$

where each a_i is given by the vector projection

$$a_i = \frac{\langle f, P_i \rangle}{\|P_i\|^2} = \frac{2i+1}{2} \int_{-1}^1 f(x) P_i(x) dx.$$

It is not at all obvious that an infinite series of functions should converge, nor what it means for a series of functions to converge to another function. Mathematicians wrestled with these ideas for more than a century establishing the basics of modern function theory. Proof of the theorem above will require you to consult an advanced textbook on spaces of functions. For now, accept the assertion that the theorem above does indeed have a proof and that mathematicians, physicists, and engineers have been using these sorts of ideas for far longer than they have been rigorously demonstrable.

You might wonder why we would prefer to work with Legendre series rather than the Taylor series that are so easy to compute - first, notice that Taylor series only exist for *analytic* functions, where here we have Legendre series for functions that are merely continuously differentiable. Secondly, it turns out that working numerically with Taylor approximations is actually pretty terrible. The term used in numerical analysis is **ill-conditioned**. Legendre expansions are significantly more stable in computations. In many cases, Legendre series also converge to the function being approximated much more quickly than Taylor series, which we can illustrate with the following example.

Example 2.3.6 Taylor and Legendre approximation of $\sin \pi x$. The function $\sin \pi x$ completes one period on the interval $(-1, 1)$ and obviously has a continuous derivative. The Taylor expansion of $\sin \pi x$ to degree 5 is given by the formula

$$T_5(x) = \pi x - \frac{\pi^3 x^3}{6} + \frac{\pi^5 x^5}{120}$$

It is straightforward to compute the first few terms in the Legendre expansion as follows: First, note that since sine is odd, there will be no terms of even degree. That is, $\sin x = a_1 P_1(x) + a_3 P_3(x) + a_5 P_5(x)$.

$$\begin{aligned} a_1 &= \frac{2(1)+1}{2} \int_{-1}^1 x \sin \pi x \, dx = \frac{3}{\pi} \\ a_3 &= \frac{2(3)+1}{2} \int_{-1}^1 \frac{1}{2}(3x^2 - 1) \sin \pi x \, dx = \frac{7}{\pi^3}(\pi^2 - 15) \\ a_5 &= \frac{2(5)+1}{2} \int_{-1}^1 P_5(x) \, dx = \frac{11}{\pi^5}(945 - 105\pi^2 + \pi^4) \end{aligned}$$

Let us compare these approximations.

```
var ('x')
f(x) = sin(x)
g = plot(sin(pi*x), (-1,1))
P1 = legendre_P(1, x)
P3 = legendre_P(3, x)
P5 = legendre_P(5, x)
P(x) = 3/pi * P1 + 7/pi**3*(pi**2 - 15)*P3 + 11/pi**5*(945 -
105*pi**2 + pi**4)*P5
g+= plot(P(x), (-1,1), color='red')
T(x) = pi*x - pi**3*x**3/6 + pi**5*x**5/120
g+= plot(T(x), (-1,1), color = 'green')
show(g)
```

Notice how much deviation there is in the Taylor approximation towards the end of the interval, as compared to the Legendre approximation which even at 5th degree fits right into the sine function. \square

2.4 Series solutions at regular singularities

2.4.1 Types of singular points

So far, we've looked at differential equations at ordinary points, specifically

$$y'' + p(x)y' + q(x)y = 0.$$

That is, we've assumed that analytic solutions exist and proceeded with power series methods. However, a major area of interest in many applied settings is what happens at points where p and q fail to be analytic. These **singular points** are often very important to consider when analyzing a system. (You may need to know what happens to the behavior of the system as the input approaches a singular point.)

Definition 2.4.1 A point $x = x_0$ is a **regular singular point** of the equation

$$y'' + p(x)y' + q(x)y = 0$$

if

1. x_0 is a singular point of the equation,

2. and both

$$\hat{p}(x) = (x - x_0)p(x), \quad q(\hat{x}) = (x - x_0)^2 q(x)$$

are analytic at $x = x_0$

A singular point that does not satisfy these conditions is called an **irregular singular point**. \diamond

To give you some terminology (and a connection with complex analysis), a function $f(x)$ for which $f(x)(x - x_0)^n$ is analytic at x_0 for $n = N$ but singular at x_0 for $n < N$ is said to have a **pole of order N at x_0** . The point of the definition above is that for a second order equation, essentially the worst singular behavior we can have and still work with series directly is second order poles. Similar statements apply to higher order equations.

Example 2.4.2 Classifying points. Classify the points of the equation

$$y'' + \frac{1}{x(x-1)^2}y' + \frac{x}{x(x-1)^3}y = 0$$

The equation clearly has singularities at $x = 0, 1$. For $x = 0$,

$$\hat{p}(x) = xp(x) = \frac{1}{x-1}, \quad \hat{q}(x) = xq(x) = \frac{1}{(x-1)^3},$$

both of which are analytic at $x = 0$. Thus, $x = 0$ is a regular singular point.

On the other hand, at $x = 1$, we have

$$\hat{p}(x) = (x-1)p(x) = \frac{1}{x(x-1)},$$

which is singular at $x = 1$. Thus, $x = 1$ is an irregular singular point. \square

2.4.2 Series solutions at regular singular points

Suppose that $x = x_0$ is a regular singular point for the equation

$$y'' + p(x)y' + q(x)y = 0.$$

Essentially we can remove the singular behavior by multiplying through by $(x - x_0)^2$ to get an equation of the form

$$(x - x_0)^2 y'' + (x - x_0)\hat{p}(x)y' + \hat{q}(x)y = 0$$

where $\hat{p} = (x - x_0)p(x)$, $\hat{q} = (x - x_0)^2 q(x)$ are analytic at x_0 . Going one step further, we can make the substitution $u = x - x_0$, which moves the regular singular point to $x = 0$, and so we can restrict our attention to equations of the form

$$x^2 y'' + xp(x)y' + q(x)y = 0$$

where p, q are analytic at 0. (We've essentially just pulled the singular behavior out of the coefficient functions to make it easier to analyze.)

The most elementary equation in this family is the Cauchy-Euler equation, where p, q are just constants p_0, q_0 . In that case, we have a second order linear constant coefficient homogeneous equation, which we can solve using the method of the indicial equation.

Theorem 2.4.3 Suppose we have a Cauchy-Euler equation

$$x^2 y'' + xp_0 y' + q_0 y = 0.$$

The equation has solutions on $(0, \infty)$ of the form $y = x^r$, where r is a solution to the indicial equation

$$r(r-1) + p_0r + q_0 = 0 \quad (2.4.1)$$

We can use the constant case to guide our method for more general coefficients. Suppose we have an equation with coefficient functions

$$x^2y'' + xp(x)y' + q(x)y = 0$$

with p, q analytic at 0. Then we can assume that there is some positive radius of convergence R for which p, q both have convergent series expansions

$$\begin{aligned} p(x) &= p_0 + p_1x + p_2x^2 + \dots \\ q(x) &= q_0 + q_1x + q_2x^2 + \dots \end{aligned}$$

We can plug these in to get the equation

$$x^2y'' + x(p_0 + p_1x + p_2x^2 + \dots)y' + (q_0 + q_1x + q_2x^2 + \dots)y = 0.$$

Now we have to make a guess. Notice that if x is very small, then $p(x) \approx p_0$ and $q(x) \approx q_0$, which means that near 0, the equation is approximately the Cauchy-Euler equation:

$$x^2y'' + x(p_0 + p_1x + p_2x^2 + \dots)y' + (q_0 + q_1x + q_2x^2 + \dots)y \approx x^2y'' + p_0xy' + q_0y = 0.$$

So we guess that the form of the series solution includes a term x^r that describes the behavior of the function near $x = 0$ in $(0, R)$ (where we're looking at a Cauchy-Euler equation) and a piece in terms of a power series that picks up the behavior away from 0.

That is, using this intuition, we can guess that for the equation

$$x^2y'' + xp(x)y' + q(x)y = 0,$$

solutions will be of the form

$$y(x) = x^r \sum_{n=0}^{\infty} a_n x^n, \quad a_0 \neq 0 \quad (2.4.2)$$

where as before, r is a solution to the indicial equation

$$r(r-1) + p_0r + q_0 = 0.$$

This turns out to be the right idea, and series of the form (2.4.2) are called **Frobenius series**.

Theorem 2.4.4 For $x > 0$, suppose that

$$x^2y'' + xp(x)y' + q(x)y = 0 \quad (2.4.3)$$

and that p, q are analytic at 0 with a mutual radius of convergence R . Write $p(x) = \sum p_n x^n$, $q(x) = \sum q_n x^n$. Let r_1, r_2 be roots of the indicial equation (2.4.1), and assume that $r_1 \geq r_2$ if the roots are real.

Then (2.4.3) has a Frobenius series solution of the form

$$y_1(x) = x^{r_1} \sum_{n=0}^{\infty} a_n x^n, \quad a_0 \neq 0$$

which converges at least on $(0, R)$.

If r_1, r_2 are distinct and do not differ by an integer, then there exists a second linearly independent Frobenius series solution of the form

$$y_2(x) = x^{r_2} \sum_{n=0}^{\infty} b_n x^n, \quad b_0 \neq 0.$$

The reason that we require the roots not to differ by an integer is that if they do, the resulting solutions will not be linearly independent. We'll address that case in the next section. First, let's look at an example of finding Frobenius series solutions. It is useful to note the following derivatives of Frobenius series:

$$y(x) = x^r \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+r}$$

$$y'(x) = \sum_{n=0}^{\infty} (n+r) a_n x^{n+r-1} \quad (2.4.4)$$

$$y''(x) = \sum_{n=0}^{\infty} (n+r)(n+r-1) a_n x^{n+r-2}. \quad (2.4.5)$$

Example 2.4.5 Using recursion to find a Frobenius series. Find two linearly independent solutions to the equation

$$4x^2 y'' + 3xy' + xy = 0. \quad (2.4.6)$$

To check the form of the solutions, we'll put the problem into standard form

$$x^2 y'' + x\left(\frac{3}{4}\right)y' + xy = 0,$$

so that $p(x) = \frac{3}{4}, q(x) = x$, both of which are analytic on $(0, \infty)$. We need $p_0 = p(0) = \frac{3}{4}, q_0 = q(0) = 0$ to find r_1, r_2 . The indicial equation is therefore

$$r(r-1) + \frac{3}{4}r + 0 = 0$$

which has roots $r = 0, r = \frac{1}{4}$. These are distinct real roots that do not differ by an integer, so we know that we have two linearly independent Frobenius solutions of the form

$$y_1(x) = \sum_{n=0}^{\infty} a_n x^n, \quad y_2(x) = x^{1/4} \sum_{n=0}^{\infty} b_n x^n.$$

Plugging in the expressions for the series derivatives, we get

$$\begin{aligned} 0 &= 4x^2 \sum_{n=0}^{\infty} (n+r)(n+r-1) a_n x^{n+r-2} + 3x \sum_{n=0}^{\infty} (n+r) a_n x^{n+r-1} + x \sum_{n=0}^{\infty} a_n x^{n+r} \\ &= \sum_{n=0}^{\infty} 4(n+r)(n+r-1) a_n x^{n+r} + \sum_{n=0}^{\infty} 3(n+r) a_n x^{n+r} + \sum_{n=0}^{\infty} a_n x^{n+r+1} \\ &= \sum_{n=0}^{\infty} 4(n+r)(n+r-1) a_{n-1} x^{n+r} + \sum_{n=0}^{\infty} 3(n+r) a_n x^{n+r} + \sum_{n=1}^{\infty} a_{n-1} x^{n+r} \end{aligned}$$

To combine the series, we can pull off the term for $n = 0$ on the left sums and get the expression

$$x^r(4r(r-1) + 3r)a_0 = 0$$

which we should note leads directly to the indicial equation. For the remaining terms, we combine the series and get the recurrence relation

$$[4(n+r)(n+r-1) + 3(n+r)]a_n + a_{n-1} = 0$$

which simplifies to

$$a_n = \frac{1}{4(n+r)(n+r-1) + 3(n+r)} a_{n-1}$$

Now we can use the values of r that we determined before to write the Frobenius series for each value.

$r = 0$: Here, we have a standard power series recursion

$$a_n = \frac{1}{4n(n-1) + 3n} a_{n-1} = \frac{1}{n(4n-1)} a_{n-1}$$

which gives

$$\begin{aligned} n = 1 : a_1 &= \frac{1}{1 \cdot 3} a_0 \\ n = 2 : a_2 &= \frac{1}{2 \cdot 7} a_1 = \frac{1}{1 \cdot 3 \cdot 2 \cdot 7} a_0 = \frac{1}{2!(3 \cdot 7)} a_0 \\ n = 3 : a_3 &= \frac{1}{3 \cdot 11} a_2 = \frac{1}{3!(3 \cdot 7 \cdot 11)} a_0 \end{aligned}$$

In general, this pattern looks something like

$$a_n = \frac{1}{n!(3 \cdot 7 \cdot \dots \cdot (4n-1))} a_0$$

Thus, the Frobenius series is

$$y_1(x) = a_0 x^0 \sum_{n=0}^{\infty} \frac{1}{n!(3 \cdot 7 \cdot \dots \cdot (4n-1))} x^n, \quad x \in (0, \infty).$$

$r = \frac{1}{4}$: We'll use b_i to denote the coefficients of the second solution. Plugging in r , we get the recursion

$$\begin{aligned} a_n &= \frac{1}{4(n + \frac{1}{4})(n + \frac{1}{4} - 1) + 3(n + \frac{1}{4})} a_{n-1} \\ &= \frac{1}{n(4n+1)} a_{n-1} \end{aligned}$$

which will give

$$a_n = \frac{1}{n!(5 \cdot 9 \cdot \dots \cdot (4n+1))} a_0.$$

Thus, a second independent solution is

$$y_2(x) = b_0 x^{1/4} \sum_{n=1}^{\infty} \frac{1}{n!(5 \cdot 9 \cdot \dots \cdot (4n+1))} x^n, \quad x \in (0, \infty).$$

We could choose any value for a_0, b_0 and have a solution to (2.4.6), so it is convenient to let both be 1, and then write

$$\begin{aligned} y_1(x) &= \sum_{n=0}^{\infty} \frac{1}{n!(3 \cdot 7 \cdot \dots \cdot (4n-1))} x^n, \quad x \in (0, \infty) \\ y_2(x) &= x^{1/4} \sum_{n=1}^{\infty} \frac{1}{n!(5 \cdot 9 \cdot \dots \cdot (4n+1))} x^n, \quad x \in (0, \infty) \end{aligned}$$

□

2.5 More on Frobenius methods (partial)

2.5.1 Frobenius theory at regular singular points

More may end up here eventually. For now, it is sufficient to note that the method of reduction of order can be used to produce a second linearly independent solution from the first Frobenius series obtained. The form of the resulting answer will vary depending on the nature of the roots.

Theorem 2.5.1 *For $x > 0$, suppose that*

$$x^2 y'' + xp(x)y' + q(x)y = 0 \quad (2.5.1)$$

and that p, q are analytic at 0 with a mutual radius of convergence R . Write $p(x) = \sum p_n x^n, q(x) = \sum q_n x^n$. Let r_1, r_2 be roots of the indicial equation (2.4.1), and assume that $r_1 \geq r_2$ if the roots are real.

Then (2.5.1) has two linearly independent solutions on at least the interval $(0, R)$. Exactly one of the following cases holds.

1. $r_1 - r_2$ is not an integer.

$$y_1(x) = x^{r_1} \sum_{n=0}^{\infty} a_n x^n, \quad a_0 \neq 0$$

$$y_2(x) = x^{r_2} \sum_{n=0}^{\infty} b_n x^n, \quad a_0 \neq 0.$$

2. $r_1 = r_2 = r$.

$$y_1(x) = x^r \sum_{n=0}^{\infty} a_n x^n, \quad a_0 \neq 0$$

$$y_2(x) = y_1(x) \ln x + x^r \sum_{n=0}^{\infty} b_n x^n, \quad a_0 \neq 0.$$

3. $r_1 - r_2$ a positive integer.

$$y_1(x) = x^{r_1} \sum_{n=0}^{\infty} a_n x^n, \quad a_0 \neq 0$$

$$y_2(x) = A y_1(x) \ln x + x^{r_2} \sum_{n=0}^{\infty} b_n x^n, \quad a_0 \neq 0.$$

2.5.2 Irregular points

Irregular points are important but outside the scope of this class. See [this lecture](#) for example as a starting point. The main idea is that solutions still involve Frobenius series, but now multiplied by exponential factors. That is, irregular points give rise to solutions that blow up or collapse!

2.6 Bessel equations

2.6.1 Bessel's differential equation

One of the most important differential equations in applied mathematics and physical science is the [Bessel differential equation](#), which has the form

$$x^2 y'' + xy' + (x^2 - \alpha^2)y = 0 \quad (2.6.1)$$

for some complex constant α , which is called the **order** of the solutions arising from the associated equation. Bessel's equation and the associated solutions are a core technique in the analysis of the propagation of waves, for example.

In general, we have to write the solutions to this equation in terms of integrals; that is, there are not closed form solutions. But since (2.6.1) has a regular singular point at 0, in the case where α is a nonnegative real constant, we can use Frobenius techniques to write series solutions.

Since $p(0) = 1$ and $q(0) = -\alpha^2$, we get the indicial equation

$$r(r-1) + r - \alpha^2,$$

which has solutions $\pm\alpha$. As long as 2α is not an integer, we can find two Frobenius series solutions of the form

$$y(x) = x^r \sum_{n=0}^{\infty} a_n x^n,$$

and again it is useful to recall the derivative formulas (2.4.4) and (2.4.5).

We proceed with our standard technique and recover a series solution from a recurrence relation. Substituting into the (2.6.1), we get

$$\sum_{n=0}^{\infty} [(r+n)^2 - \alpha^2] a_n x^{n+r} + \sum_{n=2}^{\infty} a_{n-2} x^{n+r} = 0.$$

We can peel off the first two terms of the left sum. For $n = 0$, we get the indicial equation. For $n = 1$, we get

$$[(r+1)^2 - \alpha^2] a_1 = 0.$$

For $n \geq 2$, we get the recurrence relation

$$[(n+r)^2 - \alpha^2] a_n = a_{n-2}.$$

To construct the first solution, consider the root $r = \alpha$. The recurrence relation becomes

$$a_n = -\frac{1}{n(2\alpha+n)} a_{n-2}.$$

Since the gap between the terms in the relation is two, there will be two families of coefficients. Since $a_1 = 0$, it must be that $a_3, a_5, \dots = 0$. On the other hand, with a little work we get that the even terms are characterized by

$$a_{2k} = \frac{(-1)^k}{2 \cdot 4 \cdots (2k) \cdot (2\alpha+2) \cdots (2\alpha+2k)} a_0,$$

which is a bit of a bear. Next, we'll try to reduce this into a more palatable form. First, we can pull out a whole pile of 2s to get something a bit more condensed, and a form that might help guide a useful choice for a_0 .

$$a_{2k} = \frac{(-1)^k}{2^{2k} k! (\alpha+1) \cdots (\alpha+k)} a_0, k = 1, 2, \dots$$

Then we have a Frobenius series solution to (2.6.1) of the form

$$y_1(x) = a_0 x^\alpha \left[1 + \frac{(-1)^k}{2^{2k} k! (\alpha + 1) \cdots (\alpha + k)} x^{2k} \right]. \quad (2.6.2)$$

2.6.2 Bessel functions of the first kind

When $\alpha = N$ is an integer, the formula collapses even further if we make a clever choice for the arbitrary constant a_0 . Let

$$a_0 = \frac{1}{N! 2^N}.$$

Then the solution is denoted $J_N(x)$ and is called the **Bessel function of the first kind of order N** and has the form

$$J_N(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! (N+k)!} \left(\frac{x}{2}\right)^{2k+N}.$$

The following code uses prebuilt libraries to construct and plot the first few examples of J_N

```
var('x')
g = plot(bessel_J(0, x), (0, 50), color = 'red')
g += plot(bessel_J(1, x), (0, 50), color = 'blue')
g += plot(bessel_J(2, x), (0, 50), color = 'green')
show(g)
```

The graphs above should suggest *why* these functions are useful at modeling wave propagation - each function oscillates with decreasing amplitude and has an infinite number of zeros on $(0, \infty)$.

To deal with the case where $\alpha \in (0, \infty)$ is not an integer, we need to introduce a generalization of the factorial function called the **gamma function**, one of the most famous special functions in mathematics. The basic idea of the gamma function is to “fill in the spaces” between the factorials with a nice, smooth function that retains the special properties of the factorial function (that is, a nice function y for which $y(x) = xy(x-1)$). It turns out that the best way to do that is with a function defined in terms of an integral.

Definition 2.6.1 The gamma function $\Gamma : (0, \infty) \rightarrow \mathbb{R}$ is defined by the integral formula

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt. \quad (2.6.3)$$

◇

The gamma function is actually defined on all complex numbers outside of the negative integers, but we do not need that fact here.

Theorem 2.6.2 For $x > 0$,

$$\Gamma(x+1) = (x+1)\Gamma(x).$$

Proof. This is an immediate consequence of integration by parts.

$$\begin{aligned} \Gamma(x+1) &= \int_0^{\infty} t^x e^{-t} dt \\ &= [-t^x e^{-t}]_0^{\infty} + x \int_0^{\infty} t^{x-1} e^{-t} dt \end{aligned}$$

$$= x\Gamma(x)$$

■

If we had a starting point, we could show that the gamma function generates the factorials. Fortunately, it is very easy to show that

$$\Gamma(1) = \int_0^\infty e^{-t} dt = 1.$$

Then induction on (2.6.3) gets us

$$\Gamma(2) = 2\Gamma(1) = 2!, \quad \Gamma(3) = 3\Gamma(2) = 3!,$$

and in general that for a positive integer N that

$$\Gamma(N) = N!.$$

At the same time, we can get the factorial-like formula

$$\Gamma(x+k+1) = [(x+1)(x+2)\cdots(x+k)]\Gamma(x). \quad (2.6.4)$$

Now we can extend the gamma function to negative numbers that aren't integers by using the relation

$$\Gamma(x) = \frac{\Gamma(x+1)}{x} \quad (2.6.5)$$

If $x \in (-1, 0)$, then $x+1 \in (0, 1)$, and so for $x \in (-1, 0)$ we use (2.6.5) as the definition of $\Gamma(x)$. We can continue in this way inductively: if $x \in (-2, -1)$, then $x+1 \in (-1, 0)$, which is now defined, and again we use formula (2.6.5). The resulting function is continuous off the negative integers and has the graph below.

```
var('x')
plot(gamma(x), (-5,5), detect_poles='true', ymin=-5, ymax =
    5)
```

The gamma function allows us to define the Bessel functions for non-integer orders. First, choose

$$a_0 = \frac{1}{2^\alpha \Gamma(\alpha+1)}$$

and plug into (2.6.2) to get

$$J_\alpha(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+1)\Gamma(\alpha+k+1)} \left(\frac{x}{2}\right)^{2k+\alpha}, \quad (2.6.6)$$

which by the definition of Γ reduces to $J_N(x)$ when N is an integer.

2.6.3 Bessel functions of the second kind

Because the differential equation (2.6.1) has lots of solutions, there are many ways to write down a pair of linearly independent functions that solve the equation. The function that we will consider here, we choose for convenience by convention of physicists and engineers, but there are several equivalent ways of formulating the general solution. First, note that as long as 2α isn't an integer, the second root to the indicial equation, $r = -\alpha$ will also have an associated

and linearly independent solution. It is straightforward that we can just replace α with $-\alpha$ in (2.6.6) to get

$$J_{-\alpha} = \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+1)\Gamma(k-\alpha+1)} \left(\frac{x}{2}\right)^{2k-\alpha}$$

and a general solution

$$y(x) = c_1 J_{\alpha}(x) + c_2 J_{-\alpha}(x).$$

In practice, we don't use $J_{-\alpha}$ as the second solution, but rather a linear combination of J_{α} and $J_{-\alpha}$. Thus we define the **Bessel function of the second kind of order α** by

$$Y_{\alpha}(x) = \frac{J_{\alpha}(x) \cos \alpha\pi - J_{-\alpha}(x)}{\sin \alpha\pi}.$$

The function is constructed this way because it extends naturally to the case where α is an integer, avoiding the complications of the reduction of order technique needed to derive the second solution from the first in the α is an integer case. Without doing the computations, it is enough to know that if $\alpha = n$, then we can define

$$Y_n(x) = \lim_{\alpha \rightarrow n} \left[\frac{J_{\alpha}(x) \cos \alpha\pi - J_{-\alpha}(x)}{\sin \alpha\pi} \right]$$

and that this sum can be computed.

Thus the general solution to Bessel's equation is

$$y(x) = c_1 J_{\alpha}(x) + c_2 Y_{\alpha}(x)$$

for $\alpha \in (0, \infty)$.

```
var('x')
g = plot(bessel_Y(0, x), (0,50), color = 'red')
g += plot(bessel_Y(1, x), (0,50), color = 'blue')
g += plot(bessel_Y(2,x), (0,50), color = 'green')
show(g)
```

Notice that the Y_{α} are unbounded at $x = 0$, which is typically not a desired characteristic in a physical problem.

2.6.4 Bessel-Fourier series

We've already seen that Legendre polynomials form an orthogonal basis for the C^1 functions on a finite interval with respect to the inner product

$$\langle f, g \rangle = \int f(x)g(x) dx.$$

It turns out that Bessel functions can be used to form a different orthogonal system of functions that serve as a basis for the C^1 functions (one of an infinite family), but with a slightly different inner product. In fact, just one function J_{α} can be used to generate the basis (where the choice of α lets us describe functions in terms of a solution to a particular physical scenario).

The graphs of J_{α} should be evidence that for a given α , J_{α} has infinite zeros (as it wiggles back and forth across the x -axis). For a fixed α , let $\lambda_{\alpha,n}$ denote the sequence of positive zeros of J_{α} .

To construct an orthonormal family, we'll need to use a modification of the standard inner product by introducing a **weight function**. One way of thinking about the standard inner product $x \cdot y = \sum x_i y_i$ or $\langle f, g \rangle = \int f g$ is that we give all terms equal importance. But there's no reason that we have to do that. A weight function $w(x) \geq 0$ essentially turns an inner product into a weighted average instead of a standard average. On \mathbb{R}^n , a weighted inner product has the form $x \cdot y = \sum x_i y_i w(i)$. On function spaces, we have the **weighted inner product with respect to w** which is given by

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

We do this because we change the geometry of the space when we use a new inner product. New vectors will be orthogonal. We will choose the weight function $w(x) = x$, and we get the following fact about the family $J_\alpha(\lambda_{\alpha,n}x)$.

Proposition 2.6.3 *Let α be a fixed parameter, and let $\lambda_{\alpha,n}$ denote the sequence of positive zeros of the Bessel function of the first kind J_α .*

Then with respect to the weighted inner product

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

the family of functions given by

$$\{J_\alpha(\lambda_{\alpha,n}x)\}_{n=1}^\infty$$

is an orthogonal set of vectors in $C^1[0, 1]$.

What we're really saying here is that for any $n \neq m$ that

$$\langle J_\alpha(\lambda_{\alpha,n}x), J_\alpha(\lambda_{\alpha,m}x) \rangle = \int_0^1 x J_\alpha(\lambda_{\alpha,n}x) J_\alpha(\lambda_{\alpha,m}x) dx = 0.$$

We can also show (with quite a bit of elbow grease) that

$$\|J_\alpha(\lambda_{\alpha,n}x)\| = \frac{1}{2} [J_{\alpha+1}(\lambda_{\alpha,n})]^2.$$

Even better, the family $J_\alpha(\lambda_{\alpha,n}x)$ turns out to be a basis for C^1 , and so we can express any C^1 function in terms of a series of Bessel functions.

Theorem 2.6.4 Fourier-Bessel expansion. *Suppose that $f \in C^1[0, 1]$. Then for $x \in [0, 1]$,*

$$f(x) = a_1 J_\alpha(\lambda_{\alpha,1}x) + a_2 J_\alpha(\lambda_{\alpha,2}x) + \dots = \sum_{n=1}^\infty a_n J_\alpha(\lambda_{\alpha,n}x),$$

where the coefficients a_i are given by the projection formula

$$a_i = \frac{\langle f, J_\alpha(\lambda_{\alpha,i}x) \rangle}{\|J_\alpha(\lambda_{\alpha,i}x)\|^2} = \frac{2}{[J_{\alpha+1}(\lambda_{\alpha,i})]^2} \int_0^1 x f(x) J_\alpha(\lambda_{\alpha,i}x) dx.$$

Again, the proof of convergence of these types of series is beyond the scope of the course and more suited to a course in real analysis or partial differential equations. We keep emphasizing the vector notation not just to impress with our ability to assemble huge blocks of mathematics, but to point out that all we're really doing here is linear algebra with fancier vectors.

2.7 Series as vectors (an introduction to L^2)

One of the major themes of this section is that functions are secretly vectors. Even better, functions with some level of regularity are vectors in an inner product space, which means that we can use basic ideas from linear algebra to understand them (once we know which space they are in). The key idea in all of this is that the integral is really just an infinite sum indexed by the input x , and so the dot product is really just a finite dimensional version of an integral (in fact, there are areas of math where this is made explicit, as in probability theory).

With Legendre polynomials and Bessel functions, we've demonstrated that C^1 functions live inside an infinite dimensional inner product space. Just as in finite dimensional linear algebra, we used the orthogonal expansion theorem to write f as

$$f(x) = \sum \frac{\langle f(x), b_n(x) \rangle}{\langle b_n(x), b_n(x) \rangle} b_n(x)$$

for some orthogonal basis b_n . We pick a set of basis function based on the particular problem we are trying to understand. For example, if we're doing numerical estimation of a function, we can choose b_n to be the Legendre polynomials and get better approximations than the basic Taylor series. If we're studying functions acting in a system associated with a Bessel equation, we can express the function in terms of the solutions to that system to understand how the wave behavior is driving the function.

The theoretical backdrop for all of this is the theory of **Hilbert space**, which is where questions of convergence are first addressed. One of the most important Hilbert spaces is called $L^2[a, b]$ - these are the functions f defined on an interval $[a, b]$ with finite energy. A function is in L^2 if

$$\sqrt{\int_a^b f(x)^2 dx} < \infty$$

(where dx is a more powerful type of differential called **Lebesgue measure**). But this is precisely the inner product we considered with the Legendre polynomials:

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

Taylor polynomials and Legendre polynomials are outstanding bases to use for L^2 if we're interested in looking at local behavior of a function (say zeros or extreme points). But lots of other functions live in L^2 that don't have nice behavior that we might be interested in looking at. For example, we might have an equation that models a circuit, and we want to feed in a square wave to see what the response might be.

```
var('x')
f1(x) = 1
f2(x) = 0
f(x) = piecewise([(0,1),f1],[(1,2),f2],[(2,3),f1])
plot(f, (0,3))
```

Now, we can come up with a Legendre polynomial expansion for this function (at least on the piecewise bits), but it isn't going to really capture what's interesting about the square wave - a square wave is **periodic**. If only there was a basis of functions for L^2 that was orthogonal and periodic and that

could be used to model functions like the square wave. Such a basis would be extremely useful for solving differential equations involving oscillations. If only....

Chapter 3

All things Fourier

3.1 Periodicity

As hinted at in the prior section, our concern for this chapter will be to study functions that are **periodic**, that is, functions that repeat their behavior after a fixed change in the input variable (called the **period** of the function). Periodic functions are some of the most important in the physical sciences, engineering, and mathematics. We're setting the stage for your eventual adventures in partial differential equations, and so we need to treat two kinds of input variables, space and time. A typical function we might consider could look something like $f(x, t) = y$, where the position x and time t both determine the behavior of the output y .

For now, we'll restrict our attention to functions of one variable. A function $f(t)$ is **periodic of period** T if there is some constant T so that

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

for all t . Of course, if one such T exists, many do. That is, any integer multiple of the period is also a period. (For example, $\cos t = \cos(t + 2n\pi)$ for every integer n .) The smallest such T for which $f(t) = f(t + T)$ is called the fundamental period.

One iteration of the function over a period is called a **cycle**. The number of cycles per unit input is referred to in earlier courses as the frequency of the functions, but this isn't really a correct idea unless we're talking about a simple sine or cosine. Consider the function

$$f(t) = \cos 2\pi t + \frac{2}{3} \cos 6\pi t.$$

The periods of the individual summands are 1 and $1/3$ but the period of the entire function is 1.

```
var('t')
f(t) = cos(2*pi*t) + 2/3*cos(6*pi*t)
plot(f, (-3,3))
```

While the period of f is 1, it doesn't really make sense to talk about the *frequency* of f as a single number, because f seems to carry two frequencies, one from each of the constituent cosines. Naturally, we might be led to the naive conclusion that any sum of period functions is periodic, but this can't be true. Functions with irrational periods make this impossible. Consider the function $g(t) = \cos t + \sin \sqrt{2}t$. No integer multiple of $\sqrt{2}$ will ever coincide

with an integer multiple of 1, and so these functions will never coincide at the end of a cycle. (In fact, the graph of this function is quite chaotic.)

```
var('t')
f(t) = cos(sqrt(2)*t) + cos(t)
plot(f, (-60,60))
```

On the other hand, it's pretty easy to show that any two functions with rational periods must eventually coincide.

Checkpoint 3.1.1 $fTgSf + g$

The classical example we see over and over again of time periodic systems is *harmonic oscillation*. We discover that the state of a harmonic oscillator is given by a function

$$x(t) = A \sin(2\pi\nu t + \phi)$$

for amplitude A , frequency ν , and phase ϕ . The period of the function is $1/\nu$ (which you should show by direct computation).

For spatial periodicity, the classical problem is the distribution of heat on a circular ring. A point on the ring is designated by an angle θ . Space periodicity comes from the fact that for any function that depends on the position on the ring $f(\theta)$, it will be the case that $f(\theta) = f(\theta + 2\pi)$. The problem is to give some initial distribution of heat on the ring. In the long run as $t \rightarrow \infty$, we expect the ring to be evenly distributed (constant for all θ). So how can we write down the short term distributions?

Fourier's idea was to try to model the answer with a sum of sines:

$$T(\theta, t) = \sum_{n=1}^{\infty} A_n(t) \sin(n\theta + \phi_n).$$

This idea basically launched the development of the language of modern physics and engineering.

A term $\sin(n\theta + \phi)$ is called a **harmonic** (from the mathematics of music). In the same way that a musical chord is made up of several harmonics combined into a whole, so too is the function $T(\theta, t)$ made up of harmonics that build the complete function. As weird as it seems, in some sense we're going to treat physical systems like music and try to pull the solutions apart in terms of their constituent notes.

3.2 Complex exponentials and trig functions

3.2.1 Complex numbers and arithmetic

I am a complex analyst by trade, so I will assert that the complex numbers are really the only way to go if you want to do mathematical analysis. The use of complex numbers and functions to represent real systems is common convention for several reasons. First, the math is WAY easier in the complex setting. Basically, without delving into complex analysis, we're going to be cheating by thinking of functions as complex-valued with real inputs.

First and foremost, I am a mathematician, not an engineer, so I will use the symbol i to represent the terribly named imaginary unit

$$i = \sqrt{-1}.$$

Feel free to use j if it makes your brain happy.

A complex number is a quantity

$$z = x + iy$$

where x, y are real numbers. x is called the **real part** of z and is written $x = \operatorname{Re} z$. Likewise, y is called the **imaginary part** of z (yes, the imaginary part is real) and is written $y = \operatorname{Im} z$. Notice that this means that we can think of the real numbers as living inside the complex numbers (it's the set with $y = 0$).

You should think of the complex numbers as essentially having the same structural features as the real numbers. The operations are defined on the real and imaginary parts, with the additional feature that $i^2 = -1$. Say that $z = x + iy$ and $w = a + ib$.

- Addition:

$$z + w = (x + a) + i(y + b)$$

- Multiplication

$$zw = (a + ib)(x + iy) = (ax - by) + i(bx + ay)$$

- Division:

$$\frac{z}{w} = \frac{x + iy}{a + ib} = \frac{(x + iy)(a - ib)}{(a + ib)(a - ib)} = \frac{(ax + by)}{a^2 + b^2} + i \frac{ay - bx}{a^2 + b^2}$$

If $z = x + iy$ then the **complex conjugate** of z is

$$\bar{z} = x - iy.$$

Complex conjugates follow the following “distribution” like rules:

$$\overline{z + w} = \bar{z} + \bar{w}, \quad \overline{zw} = \bar{z}\bar{w}, \quad \overline{\bar{z}} = z.$$

With complex conjugates, it is very easy to write formulas for the real and imaginary part of a complex number.

$$\operatorname{Re} z = \frac{z + \bar{z}}{2}, \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i}$$

Complex conjugates are very important in calculus and matrix settings, so we have an alternative notation that we apply to more complicated objects: $\bar{z} = z^*$. Because integrals are limits of finite sums, integration respects complex conjugation and we write

$$\left(\int f(t)g(t) dt \right)^* = \int f(t)^* g(t)^* dt.$$

A natural way to think of a complex number is as a vector, with the real part representing the “ x -coordinate” and the imaginary part the “ y ”. Because of the way that complex conjugates are defined, we can compute the norm or magnitude of a complex number by

$$|z| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}.$$

In practice, it is useful to write the above expression as $|z|^2 = z\bar{z}$.

Theorem 3.2.1 Parallelogram identity. *Let z, w be complex numbers. Then*

$$|z + w|^2 + |z - w|^2 = 2(|z|^2 + |w|^2).$$

Another useful related property is

$$|z + w|^2 = |z|^2 + 2 \operatorname{Re} z \bar{w} + |w|^2.$$

Finally, just as a vector in \mathbb{R}^2 can be thought of in both Cartesian and polar forms, so too can a complex number. Given $z = x + iy$, the length of the vector can be written $r = |z| = \sqrt{x^2 + y^2}$ and the angle from the positive real axis satisfies $\tan \theta = \frac{y}{x}$. The angle θ is called the **argument** of the complex number and is sometimes written $\arg z$. First note that this means that the complex numbers of radius 1 coincide precisely with the unit circle. Since $x = r \cos \theta$ and $y = r \sin \theta$, we can write

$$z = r \cos \theta + ir \sin \theta,$$

though we're quickly going to get a more useful form. The thing to note is that complex numbers naturally encode and support the trig functions that represent rotations. As complicated as complex numbers might seem, this is the major theme of our use of them: *complex numbers are rotations of real numbers*.

3.2.2 Complex exponentials and Euler's formula

The most important complex function is the complex exponential $f(z) = e^z$. Given a Taylor series with a positive interval of convergence, we can extend the Taylor series formula to complex numbers simply by plugging in complex numbers. The set of convergence will be a disk with diameter equivalent to the real interval of convergence. Hence, we make the definition

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

which converges for all $z \in \mathbb{C}$. It is an easy exercise (do it!) to show that

$$\overline{e^z} = e^{\bar{z}}.$$

As is the case in one variable, e^z is the unique function that is its own derivative. We will take on authority that the complex exponential satisfies the usual algebraic properties (as defining the complex logarithm is more appropriately done in a complex analysis course):

$$\begin{aligned} e^{z+w} &= e^z e^w \\ e^{zw} &= e^{zw} \end{aligned}$$

Using the properties of exponential functions, we can write

$$e^z = e^{x+iy} = e^x e^{iy}.$$

This doesn't seem like we've gained much, but something magical happens with e^{iy} . When the power of a complex exponential is purely imaginary, it corresponds to a point on the unit circle! That is,

$$e^{iy} = \cos y + i \sin y, \tag{3.2.1}$$

which is an almost unimaginably useful formula called **Euler's formula**. You might suspect such a relationship should hold given that the Taylor series

for sine and cosine *almost* look like they add to e^x in the real case, but the alternating signs are wrong. In the complex case, this relation holds.

What this means is that complex multiplication and division (and hence exponentiation) can be rewritten in form that is tractable. If $z_1 = r_1 e^{i\theta_1}$, $z_2 = r_2 e^{i\theta_2}$, then

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)},$$

and

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i\theta_1 - i\theta_2}.$$

It's easy to miss how weird and useful this idea is. Essentially, we've come up with a way to multiply vectors and produce a vector.

Since (3.2.1) expresses $e^{i\theta}$ in terms of sine and cosine, we can reverse this relation to define sine and cosine in terms of complex exponentials.

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

One of the themes of this course is that we can decompose functions into pieces that have useful properties. One way that you might not have seen before is to pull a function apart into symmetric pieces. A function is called **even** if $f(x) = f(-x)$ (corresponding to symmetry about the y -axis. A function is called **odd** if $f(x) = -f(-x)$ (corresponding to symmetry about the origin). We can build even and odd functions out of any function in the following way:

$$f_{\text{even}}(x) = \frac{f(x) + f(-x)}{2}, \quad f_{\text{odd}}(x) = \frac{f(x) - f(-x)}{2}.$$

Checkpoint 3.2.2 $f_{\text{even}} f_{\text{odd}}$

f_{even} and f_{odd} decompose *any* function into symmetric pieces, as

$$f_{\text{even}}(x) + f_{\text{odd}}(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2} = f(x).$$

It turns out that the even and odd pieces of the complex exponential are the trig functions.

$$\frac{e^{i\theta} + e^{-i\theta}}{2} = \cos \theta,$$

and

$$\frac{e^{i\theta} - e^{-i\theta}}{2} = i \sin \theta.$$

3.2.3 What complex exponentials mean and how to use them

Typically, we're going to be working with a real parameter t . So define the function

$$f(t) = e^{it} = \cos t + i \sin t.$$

As t moves along in time, what is f doing? If we think about the complex numbers as vectors (that is, $f(t) = (\text{Re } f(t), \text{Im } f(t))$), we can think of f as a parametric function where $x(t) = \cos t$ and $y(t) = \sin t$. Then as t increases, the graph of f traces out the unit circle, rotating in the counterclockwise direction. Likewise, negative complex exponentials trace out the unit circle but in the clockwise direction. One complete traversal of the circle takes place as t goes from 0 to 2π .

It is typical practice in applications to work with integer values for frequency, and so in engineering (and general Fourier analysis), the convention is to use as the basic exponential relation

$$e^{2\pi it} = \cos 2\pi t + i \sin 2\pi t,$$

as this gives a function with frequency 1 (that is, we complete one loop around the unit circle for $t \in [0, 1]$). In other words, $e^{2\pi it}$ has frequency 1 Hz. More generally,

$$e^{2\pi nit} = \cos 2\pi nt + i \sin 2\pi nt$$

has frequency n Hz.

Notice that $e^{-2\pi nit}$ also has frequency n Hz, but travels in the opposite direction. This will be an important theme in Fourier analysis - oscillations of a given frequency have *two* components, propagating in opposite directions.

Complex exponentials are vastly easier to work with than the sinusoids. Recall that the generic sine function is

$$f(t) = A \sin(2\pi \nu t + \phi)$$

with frequency ν , amplitude A , and phase ϕ . The complex exponential with the same information is

$$g(t) = Ae^{i(2\pi \nu t + \phi)}$$

which can be split into

$$g(t) = Ae^{i\phi} e^{2\pi \nu t}.$$

Notice that

$$|Ae^{i\phi} e^{2\pi \nu t}| = |A| = A$$

and that $e^{i\phi}$ describes a starting point on the unit circle.

Finally, we'll come to a fact that you may have seen in previous courses, though usually without explanation.

Fact 3.2.3

This is not at all obvious from a superimposed graph.

```
var('x')
f(x) = 2*sin(3*pi*x + pi/2)
g(x) = sin(3*pi*x)
pic = plot(f, color='red')
pic += plot(g, color='green')
show(pic)
```

But if we plot the sum, we get:

```
var('x')
f(x) = 2*sin(3*pi*x + pi/2)
g(x) = sin(3*pi*x)
pic = plot(f + g, color='blue')
show(pic)
```

Here the computation is massively simplified by the use of complex exponentials. Consider

$$A_1 e^{i(2\pi \nu + \phi_1)} + A_2 e^{i(2\pi \nu + \phi_2)}.$$

Using basic exponential algebra, we can compute

$$\begin{aligned} & A_1 e^{i(2\pi \nu + \phi_1)} + A_2 e^{i(2\pi \nu + \phi_2)} \\ &= e^{2\pi i \nu t} (A_1 e^{i\phi_1} + A_2 e^{i\phi_2}). \end{aligned}$$

That is, the sum of two exponentials with the same frequency is another exponential of the same frequency (but with a new phase and amplitude).

3.3 Sums of periodic functions

3.4 Questions of convergence

3.5 First applications

3.6 Fourier series and linear algebra

3.7 The Fourier transform

3.8 Convolution

3.9 Applications of convolution

Chapter 4

The Laplace Transform

4.1 The Laplace Transform

4.1.1 Definition of the Laplace Transform

Our first transform method is the **Laplace transform**, which takes a function $f : [0, \infty) \rightarrow \mathbb{R}$ and produces a function of a new variable $F(s)$.

Definition 4.1.1 Let f be a function on $[0, \infty)$. The Laplace transform of f is defined by the integral operator

$$\mathcal{L}[f] = \int_0^{\infty} f(t)e^{-st} dt = F(s).$$

◇

Before we get into when we can use the Laplace transform and why it is so powerful, we'll start with some examples of computing the transforms of some basic functions.

Example 4.1.2 $f(t) = 1$. We'll start by computing the Laplace transform of $f(t) = 1$. Recall that when we're dealing with improper integrals (with limits at ∞ in this case), we need to consider the integral as shorthand for a limit.

$$\begin{aligned} L[1] &= \int_0^{\infty} 1 \cdot e^{-st} dt = \lim_{B \rightarrow \infty} \int_0^B e^{-st} dt \\ &= \lim_{B \rightarrow \infty} \left(\frac{-1}{s} e^{-st} \right) \Big|_0^B \\ &= \frac{-1}{s} \lim_{B \rightarrow \infty} e^{-sB} - 1 \\ &= \frac{1}{s}. \end{aligned}$$

So $\mathcal{L}[1] = \frac{1}{s}$, as long as $s > 0$ (or the improper integral fails to converge, a condition that will usually apply). □

Example 4.1.3 $f(t) = t$. Now let's look at $f(t) = t$, which will require integration by parts.

$$\mathcal{L}t = \int_0^{\infty} t \cdot e^{-st} dt = \lim_{B \rightarrow \infty} \int_0^B te^{-st} dt$$

$$\begin{aligned}
&= \lim_{B \rightarrow \infty} \left. \frac{-t}{s} e^{-st} \right|_0^B - \int_0^B \frac{-1}{s} e^{-st} dt \\
&= \lim_{B \rightarrow \infty} \frac{-B}{s} e^{-sB} - 0 + \frac{1}{s} \lim_{B \rightarrow \infty} \int_0^B e^{-st} dt \\
&= 0 - 0 + \frac{1}{s} \mathcal{L}[1] = \frac{1}{s^2}
\end{aligned}$$

so long as $s > 0$. (Note that it is very easy to show that as $B \rightarrow \infty$, we get $\frac{-B}{s} e^{-sB} \rightarrow 0$ by L'Hospital's rule.) \square

Checkpoint 4.1.4 Use the same idea as the previous example to show that

$$\mathcal{L}[t^n] = \frac{n!}{s^{n+1}}, \quad s > 0$$

for any positive integer n .

Example 4.1.5 $f(t) = e^{at}$. Now consider the exponential function e^{at} for some constant a .

$$\begin{aligned}
\mathcal{L}[e^{at}] &= \int_0^\infty e^{at} \cdot e^{-st} dt = \lim_{B \rightarrow \infty} \int_0^B e^{-(s-a)t} dt \\
&= \lim_{B \rightarrow \infty} \left. \frac{-1}{s-a} e^{-(s-a)t} \right|_0^B \\
&= \frac{-1}{s-a} \lim_{B \rightarrow \infty} e^{-(s-a)B} - 1 = \frac{1}{s-a}
\end{aligned}$$

as long as $s > a$. \square

Example 4.1.6 $f(t) = \cos bt$. Our final introductory example concerns the basic trig function $f(t) = \cos bt$. (We can use parts here or just note that

$$\int e^{at} \cos bt dt = \frac{e^{at}}{a^2 + b^2} (a \cos bt + b \sin bt)$$

which is a standard table integral.) Then,

$$\begin{aligned}
\mathcal{L}[\cos bt] &= \int_0^\infty \cos bt \cdot e^{-st} dt = \lim_{B \rightarrow \infty} \int_0^B e^{-st} \cos bt dt \\
&= \lim_{B \rightarrow \infty} \left[\frac{e^{-st}}{s^2 + b^2} (-s \cos bt + b \sin bt) \right] \Big|_0^B \\
&= \frac{s}{s^2 + b^2}
\end{aligned}$$

\square

Checkpoint 4.1.7 Use a standard integral and the argument in the previous example to show that

$$\mathcal{L}[\sin bt] = \frac{b}{s^2 + b^2}, \quad s > 0.$$

Notice that at this point, we have constructed the Laplace transforms of the basic forms of forcing functions that we consider in the method of undetermined coefficients (which is no accident!).

4.1.2 Basic properties of the Laplace transform

One of the most important properties of the integral is that it acts linearly on the integrand. That is, for constants a, b and integrable functions f, g , we have

$$\int af + bg = a \int f + b \int g.$$

Because the integral is linear, often operations that are defined in terms of integrals are also linear. This is the case with the Laplace transform as long as it converges, since

$$\begin{aligned} \mathcal{L}[af(t) + bg(t)] &= \int_0^\infty e^{-st}(af(t) + bg(t)) dt \\ &= a \int_0^\infty e^{-st}f(t) dt + b \int_0^\infty e^{-st}g(t) dt \\ &= a\mathcal{L}[f(t)] + b\mathcal{L}[g(t)] \end{aligned}$$

This often gets recorded as the equivalent conditions

$$\mathcal{L}[f + g] = \mathcal{L}[f] + \mathcal{L}[g] \quad (4.1.1)$$

and

$$\mathcal{L}[cf] = c\mathcal{L}[f]. \quad (4.1.2)$$

Example 4.1.8 Laplace transform of a combination of functions. Compute the Laplace transform of $f(t) = t^2 + 3e^{2t} + \sin 3t$, including the domain.

Linearity makes this easy, since

$$\begin{aligned} \mathcal{L}[2t^2 + 3e^{2t} + \sin 3t] &= \mathcal{L}[2t^2] + \mathcal{L}[3e^{2t}] + \mathcal{L}[\sin 3t] \\ &= 2\mathcal{L}[t^2] + 3\mathcal{L}[e^{2t}] + \mathcal{L}[\sin 3t] \\ &= 2\frac{1}{s^2} + 3\frac{1}{s-2} + \frac{3}{s^2+9}, \quad s > 2. \end{aligned}$$

□

Combined with the results from the previous discussion, we can now compute the transforms of a lot of the most important forcing functions that we consider when we learn constant coefficient linear equations. So far, it seems like we haven't gained much of an advantage, but the next section should make clear that the Laplace transform can handle functions far beyond what can be done with earlier methods. Also still open is the big question:

Question 4.1.9 How does the Laplace transform apply to differential equations?

□

4.1.3 Piecewise continuous functions

The most important quality of the Laplace transform is that it can be used on functions that are piecewise continuous. In physical situations, this corresponds to forcing functions that change at points of time. For example, a continuously applied constant force could suddenly become a harmonic force modeled by a trig function at some time, and then at a later time, the force could disappear. Functions that change definitions, but which remain continuous in between those points of change are called **piecewise continuous functions**.

Definition 4.1.10 A function $f(t)$ is a piecewise continuous function on an interval $[a, b]$ if the interval can be divided up into finitely many subintervals

$[a_i, a_{i+1}]$ so that

1. f is a continuous function when restricted to each $[a_i, a_{i+1}]$, and
2. f does not possess a vertical asymptote at any of the points a_i (that is, f approaches a finite limit at the end of each subinterval)

If f is piecewise continuous on every interval of the form $[0, b]$ for any constant b , then f is piecewise continuous on $[0, \infty)$. \diamond

4.2 Periodic functions

4.3 The Laplace transform of derivatives

4.4 The shifting theorems

4.5 Heaviside functions

4.6 Impulses - the Dirac delta “function”

4.7 Convolution