Chapter 3

Fourier series

3.1 The Fourier series hypothesis

ries-hypothesis

At the end of the previous chapter we saw that addressing the initial value problem using eigensolutions depending on answering the following question:

> Given a function f defined on $[0, L]$, can we choose constants a_k so that the functions f_n, defined by
>
> $$f_n(x) = \sum_{k=1}^{n} \alpha_k \psi_k(x) \tag{3.1}$$ Fourier-series-hypothesis
>
> converge to f as $n \to \infty$?

The mathematics needed to address these questions is based on the work pioneered by Jean-Baptiste Joseph Fourier. In 1822 Fourier published his book *Théorie analytique de la chaleur* (Analytic theory of heat), in which he claimed that any "reasonable" function can be constructed as the infinite sum of sine functions.

There is a fair amount of mathematical theory that we need to develop before we are able to understand Fourier's claim in detail. Before we do that, however, it might be helpful to look at an example from Fourier himself. In §228 of his book[1], Fourier presents the example of a function $\varphi(x)$ defined for $0 \leq x \leq \pi$ by

$$\varphi(x) = \begin{cases} x & \text{for } 0 \leq x \leq \alpha, \\ \alpha & \text{for } \alpha \leq x \leq \pi - \alpha, \\ \pi - x & \text{for } \pi - \alpha \leq x \leq \pi. \end{cases} \tag{3.2}$$ Fourier-phi

[1] A digital copy of Fourier's book is freely available at Google books. I encourage you to take a look!

Fourier then states that

$$\varphi(x) = \frac{4}{\pi}\left(\sin(\alpha)\sin(x) + \frac{\sin(3\alpha)}{3^2}\sin(3x) + \frac{\sin(5\alpha)}{5^2}\sin(5x) + \frac{\sin(7\alpha)}{7^2}\sin(7x) + \dots\right) \quad (3.3)$$

Fourier-phi-ser

The following Sage code produces a plot of the function above with $\alpha = \pi/3$ and the sum truncated at the $\sin(7x)$ term.

```
var('x,k,n')
n=3
a = pi/3

f(x)=(4/pi)*sum(sin((2*k+1)*a)*sin((2*k+1)*x)/((2*k+1)^2),k
    ,0,n)

plot(f,(x,0,pi),figsize=[4,2])
```

If we take n to be a larger and larger, then we see that the function

$$\varphi_n(x) = \sum_{k=0}^{n} \frac{\sin((2k+1)\alpha)}{(2k+1)^2}\sin((2k+1)x)$$

does seem to approach the function $\varphi(x)$.

• revise?

Exercise 3.1.1. *Take α very close to zero. You should see some weird wiggles appearing in the function φ_n. This is called "Gibbs phenomenon." Look it up on Wikipedia.*

• needed

Exercise 3.1.2. $\alpha = \pi/2$ *This is the initial shape of a plucked string. What is the corresponding solution to the wave equation?*

3.2 Vector spaces

The key to understanding, and using, Fourier's claim is to take the perspective that functions "behave like functions." In order to make this precise, let's first review the properties of vectors that you learned in the vector calculus course.

The key properties of vectors is that they can be *added* and *scaled*:

- If $\mathbf{v}, \mathbf{w}$ are vectors, then we can add them to obtain a new vector $\mathbf{v}+\mathbf{w}$.
- If $\mathbf{v}$ is a vector and α is a number, then we can rescale $\mathbf{v}$ by α to obtain a new vector $\alpha\mathbf{v}$.

The two operations, adding and scaling, satisfy a long list of properties (see Exercise 3.2.1 below) that basically mean that adding and scaling are compatible in the way you think they should be. For example, $\alpha(\mathbf{v}+\mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$, etc. A collection of objects that can be added and scaled in a compatible way is called a ***vector space***. The numbers that are used to rescale the objects in the vector space are called ***scalars***. In this class we consider vector spaces with both real scalars and, in §3.6 below, complex scalars.

Example 3.1. *The collection of all continuous functions with domain* $[0, L]$ *forms a vector space. To see this note that if* f *and* g *are continuous functions, then the function* $f + g$*, defined by* $(f + g)(x) = f(x) + g(x)$*, is a continuous function. Similarly, if* f *is a continuous function and* α *is a real number, then* αf *is also a continuous function.*

It is common to give vector spaces a mathematical symbol as a name. The vector space of vectors in n-dimensional space is given the symbol $\mathbb{R}^n$. The vector space of continuous functions with domain $[0, L]$ is given the symbol $C([0, L])$.

In the vector calculus course, you used the scaling property of vector to parametrize lines in $\mathbb{R}^n$. For example, the line in $\mathbb{R}^2$ passing through $\mathbf{p} = (1, 3)$ and parallel to $\mathbf{v} = \langle 4, -2\rangle$ can be parametrized by the path

$$\mathbf{x}(t) = \mathbf{p} + t\mathbf{v} = (1 + 4t, 3 - 2t). \tag{3.4}$$

If instead we wanted a path that oscillated about the point $\mathbf{p}$ in the direction of $\mathbf{v}$ with frequency of one oscillation per unit time, then the desired path is

$$\mathbf{x}(t) = \mathbf{p} + \sin(2\pi t)\mathbf{v} = (1 + 4\sin(2\pi t), 3 - 2\sin(14\pi t)).$$

We can animate this oscillating path with the following Sage code.

```
c = animate([point((1+4*sin(2*pi*t),3-2*sin(2*pi*t))) for t
    in srange(0,1,0.05)],
             xmin=-4,ymin=-2,xmax=6,ymax=6,figsize=[4,3])
c.show(delay=25)
```

An analogous construction can be done in the vector space $C([0, L])$. Remember, though that "points" and "vectors" in this setting are actually functions. Thus a "path" is really a function $u(t, x)$ that for each fixed t is a vector (ie, function) in $C([0, L])$. Thus we can parametrize the line passing through the "point" $f(x) = e^x$ that is parallel to the "vector" $\psi(x) =$

$\cos\left(\frac{3\pi}{L}x\right)$ with the path

$$u(t,x) = f(x) + t\psi(x) = e^x + t\cos\left(\frac{3\pi}{L}x\right).$$

If instead we wanted a path that oscillated about the function f in the direction of ψ with frequency of one oscillation per unit time, then the desired path is

$$u(t,x) = f(x) + \sin(2\pi t)\psi(x) = e^x + \sin(2\pi t)\cos\left(\frac{3\pi}{L}x\right).$$

We can animate this oscillating path with the following Sage code, where we have set $L = \pi$.

```
var('x')
u = animate([exp(x) + sin(2*pi*t)*cos(3*x) for t in srange
    (0,1,0.05)],
             xmin=0,ymin=-2,xmax=pi,ymax=6,figsize=[4,3])
u.show(delay=25)
```

From this last example we see that we can interpret the standing wave solutions to the wave equation as oscillating paths in the vector space of functions. In particular, a standing wave of the form $u(t,x) = A(t)\psi(x)$ is interpreted as being an oscillating path whose "direction" is determined by $\psi(x)$.

We now introduce two important concepts in the study of vector spaces, "independence" and "basis." We motivate these concepts with the following example.

example:3D-linear-algebra

Example 3.2. *Consider the vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ *given by*

$$\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

We make two claims:

1. *Any vector* $\mathbf{x} \in \mathbb{R}^3$ *can be constructed by adding together multiples of* $\mathbf{u}, \mathbf{v}, \mathbf{w}$*, and*

2. *there is only one combination of* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *that will add up to form* $\mathbf{x}$*.*

To verify the first claim means to find scalars α, β, γ *such that*

$$\mathbf{x} = \alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{w}.$$

This means that we want to solve the equation

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

We can write this equation as a system of three equations with three unknowns (α, β, γ)*:*

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

We can systematically solve this by adding/subtracting multiples of these equations to eliminate the variables on the right. The result is that we must have

$$\begin{aligned} \alpha &= \frac{1}{2}x_1 - \frac{1}{4}x_2 + \frac{1}{2}x_3 \\ \beta &= \frac{1}{2}x_1 - \frac{1}{4}x_2 - \frac{1}{2}x_3 \\ \gamma &= \frac{1}{2}x_2. \end{aligned}$$

(Here is some Sage code that does the computation using matrix algebra.) This computation shows that for any possible vector $\mathbf{x}$ *we can choose scalars* α, β, γ *in such a way that we can build* $\mathbf{x}$ *from* $\mathbf{u}, \mathbf{v}, \mathbf{w}$*. This addresses the first claim.*

The second claim is also verified by the computation above. Note that we don't have any choice in what α, β, γ *are. Thus for any given* $\mathbf{x}$*, there is a unique way to construct it from* $\mathbf{u}, \mathbf{v}, \mathbf{w}$*.*

It is useful, however, to consider the second claim without the first claim. This is because there are circumstances where we have a list of vectors that can't necessarily be used to build any vector, but for which any construction might necessarily be unique. So, suppose that there is a vector $\mathbf{x}$ *for which there are two possible combinations, one given by numbers* $\check{\alpha}, \check{\beta}, \check{\gamma}$ *and another given by* $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$*:*

$$\begin{aligned} \mathbf{x} &= \hat{\alpha}\mathbf{u} + \hat{\beta}\mathbf{v} + \hat{\gamma}\mathbf{w} \\ \mathbf{x} &= \check{\alpha}\mathbf{u} + \check{\beta}\mathbf{v} + \check{\gamma}\mathbf{w} \end{aligned}$$

Subtracting these two equations gives

$$\mathbf{0} = \underbrace{(\hat{\alpha} - \check{\alpha})}_{\alpha}\mathbf{u} + \underbrace{(\hat{\beta} - \check{\beta})}_{\beta}\mathbf{v} + \underbrace{(\hat{\gamma} - \check{\gamma})}_{\gamma}\mathbf{w}.$$

In other words, we have a combination of the vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$*, given by the numbers* α, β, γ*, that adds up to zero. This means that* α, β, γ *must satisfy the system of equations*

$$\begin{aligned}0 &= \alpha + \beta + \gamma \\ 0 &= 2\gamma \\ 0 &= \alpha - \beta.\end{aligned}$$

It is easy to deduce that the only solution to this system is $\alpha = 0, \beta = 0, \gamma = 0$ *which means that the two combinations of* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *are not, in fact, different.*

Notice that in this alternate approach to the second claim, it suffices to consider which combinations $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *that make the zero vector.*

Example 3.2 motivates a number of definitions. We define a ***linear combination*** of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ to be any sum of the form

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k,$$

where $\alpha_1, \ldots, \alpha_k$ are scalars. Thus in Example 3.2, we found that any vector $\mathbf{x} \in \mathbb{R}^3$ can be expressed as a linear combination of $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

A collection of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is called ***(linearly) independent*** if the only linear combination of those vectors that yields the zero vector is the combination where each of the scalars is zero. In Example 3.2, we found that the vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are linearly independent.

mple:degree-2-polynomials

Example 3.3. *Consider the functions* $p_0(x) = 1$*,* $p_1(x) = x$*,* $p_2(x) = x^2$ *as elements of the vector space* $C([0, 1])$*. Linear combinations of* p_0, p_1, p_2 *are polynomials of degree* 2*.*

In this vector space, the zero vector is the function that is identically zero. It is easy to see that the only linear combination of p_0, p_1, p_2 *that yields this zero function is*

$$(0)p_0 + (0)p_1 + (0)p_2.$$

Thus p_0, p_1, p_2 *are linearly independent.*

We end this section by addressing a technical, but important, point. Notice that the definition of linear combination only deals with finite sums of vectors. This is important because with infinite sums of vectors we need to worry about convergence of the sum. For example, we might want to extend the list of polynomials in Example 3.3 to the infinite list

$$p_0(x) = 1, \quad p_1(x) = x, \quad \ldots p_k(x) = x^k, \ldots$$

of functions in $C([0,1])$. We also know that the exponential function $\exp(x) = e^x$ is in $C([0,1])$ and that it has the Taylor series

$$\begin{aligned}\exp(x) &= 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \dots \\ &= p_0(x) + p_1(x) + \frac{1}{2}p_2(x) + \frac{1}{3!}p_3(x) + \dots.\end{aligned}$$

It is tempting to view this series expansion of the exponential function as a sort of "infinite linear combination" of the functions $p_0, p_1, p_2, \dots$. However, it is easy to construct infinite sums of these functions that do not converge to a continuous function. For example, the geometric series

$$\begin{aligned}\frac{1}{1-x} &= 1 + x + x^2 + x^3 + \dots \\ & p_0(x) + p_1(x) + p_2(x) + p_3(x) + \dots\end{aligned}$$

does not converge to a function in $C([0,1])$. Consequently, we do not use the phrase "linear combination" when we are talking about an infinite sum. Instead, we use the word "series" to when considering an infinite sum.

In the exercises you show that the various collections of functions are vector spaces. Before stating the exercises we introduce a bit of notation and terminology.

Functions that can be differentiated as many times as we want are called ***smooth***. The collection of all smooth functions with domain Ω is given the symbol $C^\infty(\Omega)$. For example, the exponential function is in the collection $C^\infty(\mathbb{R})$, but the absolute value function is not in $C^\infty(\mathbb{R})$.

The absolute value function, however, only has "trouble" at one spot, namely at $x = 0$. Functions that are smooth except at a finite number of points are called ***piecewise smooth***. The collection of piecewise smooth functions with domain Ω is given the symbol $C^\infty_{\mathrm{pw}}(\Omega)$. You might wonder what the derivative of the absolute value function is. At most places, of course, the derivative is simply the slope. But at $x = 0$ the derivative not defined. For convenience, when there is a "jump" in a piecewise smooth function, we define the value at the jump to be zero.

Finally, we say that a piecewise smooth function f with domain Ω is ***square integrable*** if

$$\int_\Omega |f|^2 dV$$

is finite; here dV represents the length/area/volume element that is appropriate for the region Ω. For example, the absolute value function is square integrable on the domain $\Omega = [-1, 1]$, in which case $dV = dx$. The collection of piecewise smooth functions with domain Ω that are square-integrable is given the symbol $\mathcal{L}^2(\Omega)$[2].

HW:vector-space-defn

Exercise 3.2.1. *Look up the definition of vector space online. What are all the properties that adding and scaling must satisfy?*

★**Exercise 3.2.2.**

1. *For each real number p define the function $f_p(x) = x^p$ on the domain $(0, 1)$. For which p is f_p in $\mathcal{L}^2((0, 1))$?*

2. *Use the fact that $(f - g)^2 \geq 0$ to show that $2fg \leq f^2 + g^2$ for all f, g. Subsequently show that $\mathcal{L}^2(\Omega)$ is a vector space.*

★**Exercise 3.2.3.**

1. *Let $\mathcal{L}^2_0(\mathbb{R})$ be the collection of all square integrable functions with domain $\mathbb{R}$ and satisfying the infinite-string boundary condition (2.15). Show that $\mathcal{L}^2_0(\mathbb{R})$ is a vector space.*

2. *Let $\mathcal{L}^2_P([-L, L])$ be the collection of all square integrable functions with domain $[-L, L]$ that satisfy the periodic boundary condition (2.16). Explain why $\mathcal{L}^2_P([0, L])$ is a vector space.*

3. *Let $\mathcal{L}^2([0, L])$ be the collection of all square integrable functions with domain $[0, L]$ that satisfy the Dirichlet boundary condition (2.17). Explain why $\mathcal{L}^2_D([0, L])$ is a vector space.*

4. *Let $\mathcal{L}^2_N([0, L])$ be the collection of all square integrable functions with domain $[0, L]$ that satisfy the Neumann boundary condition (2.18). Explain why $\mathcal{L}^2_N([0, L])$ is a vector space.*

lutions-form-vector-space

Exercise 3.2.4. *Show that the collection of solutions $u(t)$ to the differential equation*

$$\frac{d^2u}{dt^2} + 14u = 0$$

forms a vector space. How many linearly independent solutions can you find?

[2] The spaces $\mathcal{L}^2(\Omega)$ can be extended to the Lebesgue space $L^2(\Omega)$. However, this requires the development of the Lebesgue theory of integration, which is beyond the scope of this course.

Exercise 3.2.5. *We say that a vector space is **finite dimensional** is there is a maximum number vectors in any collection of linearly independent vectors. If there is no maximum, then we say that the vector space is **infinite dimensional**.*

1. *Explain why $\mathbb{R}^2$ is finite dimensional.*
2. *Explain why $C([0,1])$ is infinite dimensional.*
3. *Is the collection of solutions to the ODE in Exercise 3.2.4 finite dimensional or infinite dimensional?*

3.3 Inner product spaces

In the vector calculus course, lengths and angles defined by vectors in $\mathbb{R}^n$ were determined using the dot product. In this section we generalize the notion of a dot product to other vector spaces. Before we do, let's recall the basic properties of the dot product for vectors in $\mathbb{R}^n$:

- *Symmetry property.* We have $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$ for all vectors $\mathbf{v}, \mathbf{w}$.
- *Positivity property.* We have $\mathbf{v} \cdot \mathbf{v} \geq 0$ for all vectors $\mathbf{v}$.
- *Definite property.* We have $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

These last two properties are often combined in to a single "positive definite property."

It is convenient to think of the dot product as a function that takes in two vectors and returns a scalar. Using this perspective, we define an ***inner product*** on a real vector space to be a function that takes in two vectors u, v and returns a number, which we give the symbol $\langle u, v\rangle$, having the following properties:

- *Linearity property.* We have $\langle \alpha_1 v_1 + \alpha_2 v_2, w\rangle = \alpha_1\langle v_1, w\rangle + \alpha_2\langle v_2, w\rangle$ for all scalars α_1, α_2 and for all vectors v_1, v_2, w in V.
- *Symmetry property.* We have $\langle v, w\rangle = \langle w, v\rangle$ for all v, w in V.
- *Positivity property.* We have $\langle v, v\rangle \geq 0$ for all v in V.
- *Definite property.* We have $\langle v, v\rangle = 0$ if and only if $v = 0$.

A (real) vector space together with an inner product is called an ***inner product space***.

The following example is very important.

Example 3.4. *For the vector space $\mathcal{L}^2(\Omega)$ we define the L^2* ***inner product*** *by*

$$\langle u, v\rangle = \int_\Omega u\, v\, dV. \tag{3.5}$$ standard-L2-inn

It is easy to see that the properties of integration imply the linearity and symmetry properties. The positivity property follows from the fact that $u^2 \geq 0$ for any function u. The definite property follows from considering each smooth piece of the function u independently, and using our convention that functions in $\mathcal{L}^2(\Omega)$ are zero along "jumps".

The inner product (3.5) is referred to in the rest of these notes as the ***standard inner product***

For vectors in $\mathbb{R}^n$ we define the norm (or magnitude) of the vector using the dot product. We can similarly define the ***norm*** of vector u, which we denote by $\|u\|$, by

$$\|u\| = \langle u, u\rangle^{1/2}.$$

example:first-norm-in-L2

Example 3.5. *Consider the functions $u(x) = 1 - x^2$ and $v(x) = \cos(\pi x)$ in $\mathcal{L}^2([0,1])$. We have*

$$\begin{aligned}\|u\| &= \langle u, u\rangle^{1/2} \\ &= \left(\int_0^1 (1-x^2)(1-x^2)\, dx\right)^{1/2} \\ &= \sqrt{\frac{8}{15}}\end{aligned}$$

and

$$\begin{aligned}\|v\| &= \left(\int_0^1 \cos^2(\pi x)\, dx\right)^{1/2} \\ &= \frac{1}{\sqrt{2}}.\end{aligned}$$

We can also use the inner product to define the angle between two vectors. If both u and v are nonzero vectors in an inner product space, then the ***angle*** between them is defined to be the number θ such that

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\,\|v\|}.$$

Example 3.6. *The angle between the functions u, v in Example 3.5 must satisfy*

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\,\|v\|} = \frac{\sqrt{15}}{2}\int_0^1 (1-x^2)\cos(\pi x)\, dx = \frac{\sqrt{30}}{2\pi}.$$

Thus $\theta \approx 0.5$ radians.

Two non-zero vectors are said to be ***orthogonal*** if the angle between them is $\pi/2$. In other words, two (non-zero) vectors u, v are orthogonal if $\langle u, v\rangle = 0$.

Example 3.7. *Consider the functions u, v from Example 3.5 as well as the function $w(x) = x - 3/8$. Direct computation shows that u and w are orthogonal, but u and v are not.*

The standard inner product (3.5) is not the only possible inner product one can define for vector spaces of functions. Consider functions defined on the domain Ω. If w is a function on Ω such that $w(x) \geq 0$, with the value zero (if present) only permitted along the boundary of Ω, then

$$\langle u, v\rangle_w = \int_\Omega u\, v\, w\, dV$$

is an inner product for functions with domain Ω. We refer to $\langle\cdot,\cdot\rangle_w$ as a ***weighted inner product*** and refer to w as the ***weight function***. The corresponding norm is denoted $\|\cdot\|_w$. The collection of piecewise smooth functions u with domain Ω such that $\|u\|_w$ finite is given the symbol $\mathcal{L}^2_w(\Omega)$, and is referred to as a ***weighted $\mathcal{L}^2$ space***.

e-inner-product

Example 3.8 (Hermite inner product)**.** *Consider the domain $\Omega = \mathbb{R}$ and the function $w_H(x) = e^{-x^2}$. We call the resulting inner product the **Hermite inner product**[3], and give it the symbol $\langle\cdot,\cdot\rangle_H$. In particular, we have*

$$\langle u, v\rangle_H = \int_{-\infty}^{\infty} u(x)v(x)e^{-x^2}\, dx.$$

Note that the function $u(x) = 1$ has finite norm with respect to the Hermite inner product. In fact,

$$\|u\|_H = \left(\int_{-\infty}^{\infty} e^{-x^2}\, dx\right)^{1/2} = \pi^{1/4}$$

The function $v(x) = x$ also has finite norm. Furthermore, $\langle u, v\rangle = 0$, meaning that these two functions are orthogonal. (The easiest way to see this is that u is even, v is odd, and the weight function is even. Since we are integrating an odd function over a symmetric domain the integral is zero.)

Exercise 3.3.1. *Let Ω be the unit disk in $\mathbb{R}^2$. The functions $u(x, y) = x$, $v(x, y) = x^2 + y^2$, $w(x, y) = e^{x^2+y^2}$ are elements in $\mathcal{L}^2(\Omega)$.*

• would be better to replace these random functions with something more meaningful

[3]The inner product is named after mathematician Charles Hermite.

1. *Compute the norms of* u, v, *and* w.
2. *Compute* $\langle u, v\rangle$.
3. *Compute* $\langle v, w\rangle$.

• needed

★**Exercise 3.3.2.** *sequences* l^2 *"square summable" lists*

• needed

Exercise 3.3.3. *Cauchy-Schwartz inequality*

nfunctions-are-orthogonal ★**Exercise 3.3.4.**

1. *Show that the eigenfunctions* ψ_k *defined by* (2.23) *are orthogonal to one another with respect to the standard inner product in* $\mathcal{L}^2([0, L])$.
2. *Compute* $\|\psi_k\|$. *Does the answer depend on* k?

olynomials-are-orthogonal **Exercise 3.3.5.** *In this exercise we work in* $\mathcal{L}^2_H(\mathbb{R})$, *defined using the Hermite inner product from Example 3.8. We define the following functions:*

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2, \quad H_3(x) = 8x^3 - 12x.$$

1. *Show that each of the functions* H_k *are orthogonal to one another.*
2. *Compute the norm of each of the functions* H_k.

3.4 The best approximation problem

est-approximation-problem

A list of vectors in an inner product space is called an ***orthogonal collection*** if each vector is orthogonal to all of the other vectors. For example:

- The vectors
$$\mathbf{i} = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \mathbf{j} = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix}0\\0\\1\end{pmatrix}$$
form an orthogonal collection in $\mathbb{R}^3$.
- In Exercise 3.3.4 you showed that the Dirichlet eigenfunctions (2.23) are orthogonal in $\mathcal{L}^2([0, L])$.

We now introduce the ***best approximation problem***:

> Suppose $v_1, \ldots, v_n$ is a orthogonal collection of vectors in an inner product space, and let u be some other vector. What linear combination of the vectors $v_1, \ldots, v_n$ best approximates u?

More precisely, how can we choose scalars $\alpha_1, \ldots, \alpha_n$ so that the vector

$$\tilde{u} = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

best approximates u in the sense that

$$\|u - \tilde{u}\|$$

is as small as possible.

To understand this problem a bit better, let's look at two simple examples in $\mathbb{R}^3$.

Example 3.9. *Consider the orthogonal collection that contains only one vector*

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Let $\mathbf{u}$ *be the vector*

$$\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

In this setting, the best approximation problem asks us to find which vector $\alpha\mathbf{v}$ *is closest to* $\mathbf{u}$.

We can easily visualize what is happening here. Vectors of the form $\alpha\mathbf{v}$ *all lie in the line passing through* $(0,0,0)$ *and with direction given by* $\mathbf{v}$. *As we move along this line we can compute the distance to the point* $(1,2,3)$. *The best approximation problem asks us to find the point with the smallest distance.*

We can solve this using methods from first semester calculus. Define the function f *by*

$$f(\alpha) = \|\alpha\mathbf{v} - \mathbf{u}\|^2.$$

If we have a minimizer for f *then we have a value of* α *where the square of the norm is smallest, and hence where the norm is smallest. We compute*

$$f(\alpha) = \begin{pmatrix} \alpha - 1 \\ \alpha - 2 \\ \alpha - 3 \end{pmatrix} \cdot \begin{pmatrix} \alpha - 1 \\ \alpha - 2 \\ \alpha - 3 \end{pmatrix} = (\alpha-1)^2 + (\alpha-2)^2 + (\alpha-3)^2.$$

It is easy to compute $f'(\alpha)$ *and then solve* $0 = f'(\alpha)$ *for* α. *The result is* $\alpha = 2$. *Thus the vector of the form* $\alpha\mathbf{v}$ *that best approximates* $\mathbf{u}$ *is the vector*

$$\tilde{\mathbf{u}} = 2\mathbf{v} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

Example 3.10. *Consider now the orthogonal collection*

$$\mathbf{v}_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix}1\\-1\\0\end{pmatrix}.$$

(You should quickly check that this is an orthogonal collection!) Again, let $\mathbf{u}$ *be the vector*

$$\mathbf{u} = \begin{pmatrix}1\\2\\3\end{pmatrix}$$

We would like to find scalars α_1, α_2 *such that*

$$\tilde{\mathbf{u}} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$$

is such that $\|\tilde{\mathbf{u}} - \mathbf{u}\|$ *is as small as possible.*

Motivated by the previous example, we define a function f *by*

$$f(\alpha_1, \alpha_2) = \|\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 - \mathbf{u}\|^2.$$

We can now use the optimization methods from vector calculus to find the values of α_1, α_2 *that minimize* f. *To do this we compute*

$$f(\alpha_1, \alpha_2) = (\alpha_1 + \alpha_2 - 1)^2 + (\alpha_1 - \alpha_2 - 2)^2 + (\alpha_1 - 3)^2.$$

Thus

$$\operatorname{grad} f = \begin{pmatrix}6\alpha_1 - 12\\4\alpha_2 + 2\end{pmatrix}$$

Setting $\operatorname{grad} f = 0$, *we find that the optimal values are*

$$\alpha_1 = 2, \quad \alpha_2 = -\frac{1}{2}.$$

Consequently the linear combination of $\mathbf{v}_1, \mathbf{v}_2$ *that best approximates* u *is*

$$\tilde{\mathbf{u}} = 2\mathbf{v}_2 - \frac{1}{2}\mathbf{v}_2 = \begin{pmatrix}3/2\\5/2\\2\end{pmatrix}.$$

The two examples above show us how to address the best approximation problem in general. Suppose we have orthogonal vectors $v_1, \ldots, v_n$ and also a vector u that we want to approximate. We define a function f by

$$f(\alpha_1, \ldots, \alpha_k, \ldots, \alpha_n) = \|\alpha_1 v_1 + \cdots + \alpha_k v_k + \cdots + \alpha_n v_n - u\|^2.$$

Since the vectors v_k are orthogonal to one another, we have

$$\begin{aligned} f(&\alpha_1, \ldots, \alpha_k, \ldots, \alpha_n) \\ &= \|u\|^2 + \alpha_1^2\|v_1\|^2 + \cdots + \alpha_k^2\|v_k\|^2 + \cdots + \alpha_n^2\|v_n\|^2 \\ &\qquad - 2\alpha_1\langle v_1, u\rangle - \cdots - 2\alpha_k\langle v_k, u\rangle - \cdots - 2\alpha_n\langle v_n, u\rangle. \end{aligned}$$

Thus the gradient of f is

$$\operatorname{grad} f = \begin{pmatrix} 2\alpha_1\|v_1\|^2 - 2\langle v_1, u\rangle \\ \vdots \\ 2\alpha_k\|v_k\|^2 - 2\langle v_k, u\rangle \\ \vdots \\ 2\alpha_n\|v_n\|^2 - 2\langle v_n, u\rangle \end{pmatrix} \tag{3.6}$$

Consequently it is easy to see that in order to minimize f, and hence in order to minimize $\|u - \tilde{u}\|$, we should choose

$$\alpha_1 = \frac{\langle v_1, u\rangle}{\|v_1\|^2}, \ldots, \alpha_k = \frac{\langle v_k, u\rangle}{\|v_k\|^2}, \ldots, \alpha_n = \frac{\langle v_n, u\rangle}{\|v_n\|^2}.$$

In other words, the linear combination of the $v_1, \ldots, v_n$ that best approximates the vector u is

$$\tilde{u} = \frac{\langle v_1, u\rangle}{\|v_1\|^2}v_1 + \cdots + \frac{\langle v_k, u\rangle}{\|v_k\|^2}v_k + \cdots + \frac{\langle v_n, u\rangle}{\|v_n\|^2}v_n.$$

One should interpret this formula in the following way. The vectors $v_1, \ldots, v_n$ describe n different orthogonal directions inside the vector space. The quantity

$$\frac{\langle v_k, u\rangle}{\|v_k\|^2}v_k$$

represents that part of the vector u that lies in direction v_k. (The technical mathematics term is "the projection of u on to the line spanned by v_k.) Thus the best approximation is accomplished by finding the portion of u lying in the directions $v_1, \ldots, v_n$ and then summing.

Of course, there may be portions of u that are orthogonal to all of the vectors $v_1, \ldots, v_n$. Therefore the vector $\tilde{u}$ might not be equal to u. The difference $u - \tilde{u}$ is precisely that part of u that is orthogonal to all of the vectors $v_1, \ldots, v_n$.

The resolution of the best approximation problem has a very important consquence: it tells us how to choose the constants α_k in (3.1). In Exercise

3.3.4 you showed that the Dirichlet eigenfunctions ψ_k are orthogonal, and have norm $\|\psi_k\|^2 = \frac{\pi}{2}$. If we want the finite sum

$$f_n(x) = \sum_{k=1}^{n} \alpha_k \psi_k(x) \tag{3.7}$$

finite-Fourier-

to best approximate the function f, then we should choose

$$\alpha_k = \frac{\langle \psi_k, f\rangle}{\|\psi_k\|^2} \tag{3.8}$$

Fourier-coeffic

We call these numbers the ***Fourier coefficients*** of the function f.

It is important to note that we have not showed that this choosing α_k according to (3.8) necessarily implies that $f_n \to f$ as $n \to \infty$, as the Fourier hypothesis conjectures, but if there is to be any hope of convergence, then this is how the constants need be chosen. The convergence f_n to f is discussed below.

With this information in hand, let's revisit Fourier's example from §3.1.

example:Fouriers-example

Example 3.11. *The function that we want to approximate is the functio φ given in (3.2). Here we take $L = \pi$. A lengthy, but straightforward, computation shows that*

$$\langle \varphi, \psi_k\rangle = \frac{1-(-1)^k}{k^2}\sin(k\alpha).$$

Recall from Exercise 3.3.4 that $\|\psi_k\|^2 = \frac{\pi}{2}$. Thus we choose

$$\alpha_k = \begin{cases} \frac{4}{k^2\pi}\sin(k\alpha) & \text{if } k \text{ is odd,} \\ 0 & \text{if } k \text{ is even,} \end{cases}$$

which is precisely the values of α_k that yield (3.3).

Here is another simple example.

richlet-constant-function

Example 3.12. *Consider the function $f(x) = 1$ on the domain $[0, \pi]$. We compute*

$$\langle \psi_k, f\rangle = \frac{1-(-1)^k}{k}$$

and thus the Fourier coefficients are

$$\alpha_k = \begin{cases} \frac{4}{k\pi} & \text{if } k \text{ is odd,} \\ 0 & \text{if } k \text{ is even.} \end{cases}$$

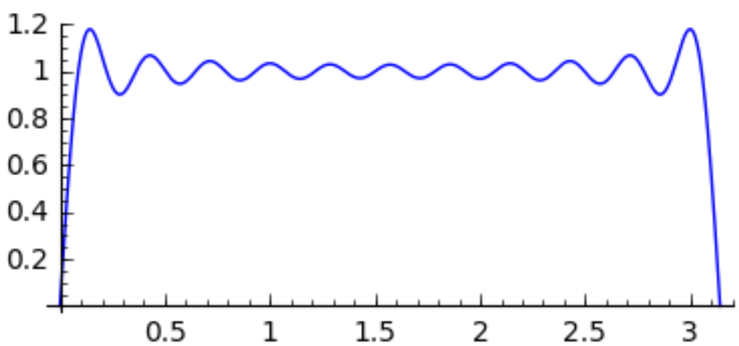
Since we only need to consider odd values of k, set $k = 2l + 1$, where $l = 0, 1, 2, 3, \ldots$. Thus the function f is best approximated by

$$f_n(x) = \sum_{l=0}^{n} \frac{4}{\pi(2l+1)} \sin\left((2l+1)x\right).$$

The Sage code

```
var('x,l,n')
n=10
f(x) =sum((4/(pi*(2*l+1)))*sin((2*l+1)*x),l,0,n)
plot(f,(x,0,pi),figsize=[4,2])
```

yields the following plot of this approximation when $n = 10$:



This does a reasonable job of approximating the constant function $f(x) = 1$. Play around with the code and see what happens when n is large. Does the approximation improve?

Our resolution to the best approximation problem gave us a very simple way to approximate a function f by the Dirichlet eigenfunctions: simply compute the Fourier coefficients α_k and construct (3.7). The limit as $n \to \infty$ of (3.7) is called the ***Fourier sine series*** (because the Dirichlet eigenfunctions ψ_k are sine functions). To indicate that the Fourier sine series does the best job possible at approximating the function f, we write

$$f \sim \sum_{k=1}^{\infty} \alpha_k \psi_k.$$

The examples above, and the exercises below, provide evidence that the Fourier sine series does converge to the function f, although Example 3.12

leads us to suspect that something funny is happening at the endpoints. Details about the convergence are addressed in §?? below. • section reference nee

The best approximation method here works any time we have an inner product space and a collection of orthogonal vectors in that space. The Dirichlet eigenfunctions ψ_k are only one possibility. This raises some questions. Were we lucky that the Dirichlet eigenfunctions were orthogonal? Or is there some "reason" that we ended up with an orthogonal collection? Is there a systematic way to construct orthogonal collections of functions? One approach to generating orthogonal collections is addressed in Exercise 3.4.2. A more general theory appears in Chapter ?? below.

• chapter reference needed

rier-sine-series-examples

Exercise 3.4.1.

1. *Find Fourier coefficients for "square wave" function*

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x < \frac{\pi}{2} \\ -1 & \text{if } \frac{\pi}{2} < x \le \pi. \end{cases}$$

(According to our convention, what should the value of f be in order to have f in $\mathcal{L}^2([0,\pi])$? Does this value affect the Fourier coefficients?) Use Sage to generate a plot of the resulting approximation.

2. *Find Fourier coefficients for the function*

$$f(x) = \begin{cases} x & \text{if } 0 \le x < \frac{\pi}{2} \\ 0 & \text{if } x = \frac{\pi}{2} \\ x - \pi & \text{if } \frac{\pi}{2} < x \le \pi. \end{cases}$$

Use Sage to generate a plot of the resulting approximation.

gendre-using-Gram-Schmidt

Exercise 3.4.2. *In this exercise you explore a method to generate an orthogonal collection of polynomials in $\mathcal{L}^2([-1,1])$. The polynomials $P_0, P_1, P_2, P_3, \ldots$ are recursively constructed according to the following conditions:*

- *P_k is a polynomial of degree k,*
- *P_k is orthogonal to $P_0, \ldots, P_{k-1}$,*
- *$P_k(1) = 1$.*

1. *Explain why we must have $P_0(x) = 1$. Compute $\|P_0\|^2$.*

2. *Now suppose that $P_1(x) = a_0 + a_1 x$. Show that the condition $\langle P_1, P_0\rangle = 0$ implies that $a_0 = 0$. Conclude that $P_1(x) = x$. Compute $\|P_1\|^2$.*

3. *Suppose that $P_2 = a_0 + a_1 x + a_2 x^2$. Show that the conditions*

$$\langle P_2, P_0 \rangle = 0 \quad \langle P_2, P_1 \rangle = 0 \quad P_2(1) = 1$$

imply that $P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$. Compute $\|P_2\|^2$.

4. *Find $P_3(x)$. Show that $\|P_3\|^2 = \frac{2}{7}$.*

Exercise 3.4.3 (Requires Exercise 3.4.2)**.** *Let $f(x) = e^x$ and let P_k be the polynomials from Exercise 3.4.2. Find the constants $\alpha_0, \ldots, \alpha_3$ such that the polynomial*

$$\tilde{f}(x) = \alpha_0 P_0(x) + \alpha_1 P_1(x) + \alpha_2 P_2(x) + \alpha_3 P_3(x)$$

best approximates f in $\mathcal{L}^2([-1,1])$.

Exercise 3.4.4 (Requires Exercise 3.3.5)**.** *In this exercise we work in $\mathcal{L}^2_H(\mathbb{R})$, defined using the Hermite inner product (see Exercise 3.3.5). Let $f(x) = e^x$.*

1. *Show that f is an element of $\mathcal{L}^2_H(\mathbb{R})$.*

2. *Find the constants $\alpha_0, \ldots, \alpha_3$ such that the polynomial*

$$\tilde{f}(x) = \alpha_0 H_0(x) + \alpha_1 H_1(x) + \alpha_2 H_2(x) + \alpha_3 H_3(x)$$

best approximates f in $\mathcal{L}^2_H(\mathbb{R})$.

You might find the following Sage code helpful:

```
var('x')
a=integral((2*x)*exp(x)*exp(-x^2),(x,-Infinity,Infinity))
show(a)
```

3.5 The Dirichlet IBVP

In the previous section we saw that the Fourier sine series gives the best approximation of a function that can be constructed as an infinite sum of the Dirichlet eigenfunctions ψ_k. We postpone discuss of the extent to which the Fourier sine series converges to §?? below. For now, let us assume that the Fourier sine series converges in a "reasonable" manner in order that we might return to our discussion of the Dirichlet initial boundary value problem for the one dimensional wave equation. • **section reference needed**

Let us recall the discussion in §2.5. Suppose we are given functions $u_0(x)$ and $v_0(x)$, defined on $[0, L]$. The corresponding Dirichlet IBVP asks us to find a function $u(t, x)$ such that

- the initial conditions are satisfied, meaning that $u(0,x) = u_0(x)$ and $\frac{\partial u}{\partial t}(0,x)$ for $x \in (0,L)$,
- the Dirichlet boundary condition is satisfied, meaning that $u(t,0) = 0$ and $u(t,L) = 0$ for all $t > 0$, and
- the wave equation
$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$$
is satisfied for $t > 0$ and $x \in (0,L)$.

The discussion in §2.5 shows that we can construct a solution to the IVBP of the form (2.29) provided we can choose constants a_k and b_k with

$$u_0(x) = \sum_{k=1}^{\infty} a_k \psi_k(x) \quad \text{and} \quad v_0(x) = \sum_{k=1}^{\infty} b_k \frac{k\pi}{L} \psi_k(x).$$

We now recognize these as the Fourier sine series for the functions u_0 and v_0. Thus we should choose

$$a_k = \frac{\langle \psi_k, u_0 \rangle}{\|\psi_k\|^2} \quad \text{and} \quad b_k = \frac{L}{k\pi} \frac{\langle \psi_k, v_0 \rangle}{\|\psi_k\|^2},$$

which means that the solution to the initial boundary value problem is

$$\begin{aligned} u(t,x) = \sum_{k=1}^{\infty} \frac{\langle \psi_k, u_0 \rangle}{\|\psi_k\|^2} \cos\left(\frac{k\pi}{L} t\right) \sin\left(\frac{k\pi}{L} x\right) \\ + \sum_{k=1}^{\infty} \frac{L}{k\pi} \frac{\langle \psi_k, v_0 \rangle}{\|\psi_k\|^2} \sin\left(\frac{k\pi}{L} t\right) \sin\left(\frac{k\pi}{L} x\right). \end{aligned} \tag{3.9}$$

1D-Dirichlet-so

Example 3.13. *Suppose we want to solve the Dirichlet initial boundary value problem for the one-dimensional wave equation on the domain $[0,\pi]$ with initial shape given by*

$$u_0(x) = \begin{cases} x & \text{for } 0 \le x \le \frac{\pi}{2}, \\ \pi - x & \text{for } \frac{\pi}{2} \le x \le \pi. \end{cases}$$

and $v_0(x) = 0$. Note that the function u_0 agrees with the function φ in Example 3.11 in the case that $\alpha = \frac{\pi}{2}$. Thus in (3.9) we ahve $b_k = 0$ and

$$a_k = \frac{2(1 - (-1)^k) \sin(\frac{k\pi}{2})}{k^2 \pi}$$

and the solution is

$$u(t,x) = \sum_{k=1}^{\infty} \frac{2(1-(-1)^k)\sin(\frac{k\pi}{2})}{k^2\pi} \cos(kt)\sin(kx).$$

Since the summands are zero when k is even, we can set $k = 2l+1$, which yields

$$u(t,x) = \sum_{l=0}^{\infty} \frac{4(-1)^l}{(2l+1)^2\pi} \cos((2l+1)t)\sin((2l+1)x).$$

The following Sage code provides an animation of the solution.

```
var('x,n,l')
n=10

solplot = animate([sum(((4*(-1)^l)/(pi*(2*l+1)^2))*sin((2*l
    +1)*x)*cos((2*l+1)*t),l,0,n ) for t in srange(0,6,0.1)],
    xmin=0, xmax=pi, ymax=2, ymin = -2, figsize=[4,2])
solplot.show(delay = 20)
```

Exercise 3.5.1 (Requires Exercise 3.4.1)**.**

1. *Solve the Dirichlet IBVP for the one dimensional wave equation on the domain $[0,\pi]$ with initial conditions*

$$u_0(x) = \begin{cases} 1 & \text{if } 0 \le x < \frac{\pi}{2} \\ -1 & \text{if } \frac{\pi}{2} < x \le \pi \end{cases}$$

   and $v_0(x) = 0$. Use Sage to generate an animation.

2. *Solve the Dirichlet IBVP for the one dimensional wave equation on the domain $[0,\pi]$ with initial conditions $u_0(x) = 0$ and*

$$v_0(x) = \begin{cases} x & \text{if } 0 \le x < \frac{\pi}{2} \\ \pi - x & \text{if } \frac{\pi}{2} < x \le \pi. \end{cases}$$

   Use Sage to generate an animation.

3.6 Complex inner product spaces spaces

x-vector-spaces

We now introduce inner product spaces involving complex numbers. Recall that complex numbers take the form $a + ib$, where both a and b are real numbers and where $i^2 = -1$. We now introduce some vocabulary regarding complex numbers. The ***real part*** of complex number $a + ib$ is the real number a; the ***imaginar part*** is the real number b.

The ***complex conjugate*** of a complex number is the complex number formed by everywhere replacing i by $-i$. The complex conjugate of a complex number is denoted with an overbar. Thus the conjugate of $z = a + ib$ is $\overline{z} = a + ib$.

The ***modulus*** of a complex number z is the non-negative real number $(\overline{z}z)^{1/2}$ and is given the same symbol as absolute value. Thus the modulus of $z = a + ib$ is $|z| = \sqrt{a^2 + b^2}$. The modulus represents the extent to which a complex number is zero: we have $|z| = 0$ exactly when $z = 0$.

An important identity of complex numbers is Euler's formula:

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Thus the real part of $e^{i\theta}$ is $\cos\theta$ and the complex part is $\sin\theta$. Note that for any complex number z we have

$$z = |z|e^{i\theta}$$

for some number $\theta \in [0, 2\pi)$.

Example 3.14. *Consider the complex number $z = -\sqrt{3} + 3i$. The real part of z is $-\sqrt{3}$ and the imaginary part is 3. The modulus is $|z| = \sqrt{12}$ and we can write*

$$z = \sqrt{12}e^{i\frac{2\pi}{3}}.$$

A ***complex vector space*** is a vector space where the scalars are complex numbers. Since all of the usual algebraic operations are possible with complex numbers, complex vector spaces behave exactly the same way that real vector spaces do.

Example 3.15. *The vector space $\mathbb{C}^2$, which consists of column vectors having two complex entries, is a complex vector space. For instance, the vector*

$$\mathbf{v} = \begin{pmatrix} 1+i \\ -3i \end{pmatrix}$$

is in $\mathbb{C}^2$. If we scale $\mathbf{v}$ by the scalar $2 - i$ we obtain

$$(2 - i)\mathbf{v} = \begin{pmatrix} 3+i \\ 3-6i \end{pmatrix}.$$

While the definition of complex vector spaces is essentially the same as for real vector spaces, the definition of complex inner product spaces is slightly different. This is because we need to make use of complex conjugates in order to compute the modulus of complex numbers. Thus the definition of complex vector spaces involves complex conjugates. In particular, function $\langle \cdot, \cdot \rangle$ that takes in two elements of a complex vector space and gives out a complex number makes is a ***complex inner product*** if it

- is conjugate linear in the first entry and linear in the second entry [4], meaning

$$\langle \alpha v, w\rangle = \overline{\alpha}\langle v, w\rangle \qquad \langle v, \alpha w\rangle = \alpha\langle v, w\rangle$$
$$\langle v_1 + v_2, w_1 + w_2\rangle = \langle v_1, w_1\rangle + \langle v_1, w_2\rangle + \langle v_2, w_1\rangle + \langle v_2, w_2\rangle;$$

- is conjugate symmetric, meaning that

$$\langle v, w\rangle = \overline{\langle w, v\rangle};$$

- is positive definite, meaning that

$$\langle v, v\rangle \geq 0,$$

  with equality only when $v = 0$. (Technical note: the conjugate-symmetry property implies that $\langle v, v\rangle$ is real, thus it makes sense to discuss this inequality.)

Example 3.16. *We can define a complex inner product on $\mathbb{C}^2$ by*

$$\left\langle \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\rangle = \overline{v}_1 w_1 + \overline{v}_2 w_2.$$

We note that all of the definitions that rely on inner products are exactly the same in the complex case as in the real case. For example, we say that two vectors v, w are ***orthogonal*** if $\langle w, v\rangle = 0$.

We now introduce the example of a complex inner product that is most important for this class. Fix some $L > 0$ and consider all piecewise smooth functions that have domain $[-L, L]$ and give complex numbers as outputs. Functions with complex outputs are called ***complex-valued*** and take the form

$$u(x) = u_{\text{real}}(x) + i\, u_{\text{im}}(x),$$

[4] Note that there is not consensus about whether the conjugate linear slot should be first or second.

where u_{real} and u_{im} are functions with real outputs. A complex-valued function u is called ***piecewise smooth*** if the functions u_{real} and u_{im} are each piecewise smooth real functions. The derivative of such a function is computed by

$$u'(x) = u'_{\text{real}}(x) + i\, u'_{\text{im}}(x).$$

We say that a complex-valued piecewise-smooth function u is ***square integrable***, and thus in $\mathcal{L}^2(\Omega)$, if

$$\int_\Omega |u|^2\, dV < \infty,$$

where $|u|^2 = \overline{u}\, u$ is the modulus of u and where dV is the appropriate length/area/volume element. For functions v, w in $\mathcal{L}^2(\Omega)$ we define the inner product

$$\langle v, w\rangle = \int_\Omega \overline{v}\, w\, dV.$$

Thus, just as in the case of real-valued functions, $\mathcal{L}^2(\Omega)$ is the collection of all functions whose norm (defined by the integral inner product) is finite.

Example 3.17. *Consider the domain $\Omega = [0, \pi]$. The function $v(x) = e^{ix}$ is complex-valued with real part $v_{real}(x) = \cos(x)$ and imaginary part $v_{im}(x) = \sin(x)$. We have $\overline{v(x)} = e^{-ix}$ and*

$$\|v\|^2 = \langle v, v\rangle = \int_0^\pi e^{-ix} e^{ix}\, dx = \int_0^\pi dx = \pi.$$

Thus $\|v\| = \sqrt{\pi}$ and v is in $\mathcal{L}^2([0, \pi])$. Note that

$$v'(x) = ie^{ix} = i\cos(x) - \sin(x).$$

Similarly, we see that the function $w(x) = ix - 5$ is in $\mathcal{L}^2([0, \pi])$. We ask, are these the functions v and w orthogonal? We compute the inner product of v and w to be

$$\langle v, w\rangle = \int_0^\pi e^{-ix}(ix - 5)\, dx = \left[-(x - i)e^{-i\,x} - 5i\, e^{-i\,x}\right]_0^\pi = \pi + 8i.$$

Thus the two functions are not orthogonal.

We conclude this section with the following note about notation.

Remark 3.18.

1. *In physics you sometimes see a star used to indicate complex conjugate, writing* z^* *instead of* $\overline{z}$. *Also, physicists use the so-called "bra-ket" notation* $\langle\cdot|\cdot\rangle$ *for inner products. Finally, physicists sometimes write "*d^3x*" for the volume element* dV*, which they place before the integrand rather than after it. The result of all this is that the following things all describe the same (three dimensional) integral:*

$$\begin{aligned} &\textit{Math} \qquad \langle u, v\rangle = \int_\Omega \overline{u}v\,dV \\ &\textit{Physics} \quad \langle u|v\rangle = \int_\Omega d^3x\,u^*v \end{aligned}$$

2. *In mathematics, one often sees complex inner products with the complex conjugate on the second, rather than the first, entry. Under this convention, complex inner products are conjugate linear in the second entry and linear in the first entry.*

-eigenfunctions ★**Exercise 3.6.1.** *In this exercise you work in* $\mathcal{L}^2([-L, L])$ *for some* $L > 0$. *Consider the functions* $\phi_k(x) = e^{i\frac{k\pi}{L}x}$ *where* $k = \ldots, -2, -1, 0, 1, 2, \ldots$.

1. *Show that the functions* ϕ_k *are in* $\mathcal{L}^2([-L, L])$. *What is* $\|\phi_k\|$*?*

2. *Show that the functions* ϕ_k *satisfy the periodic boundary conditions* (2.16).

3. *Show that the functions* ϕ_k *are orthogonal to one another.*

3.7 Periodic Fourier series

We now consider the best approximation problem in complex vector spaces. Suppose we have an orthogonal collection of complex vectors $v_1, \ldots, v_n$ and want to choose complex numbers $\alpha_1, \ldots, \alpha_n$ so that the sum

$$\tilde{u} = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

best approximates vector u. Proceeding as in §3.4, we consider the function

$$f(\alpha_1, \ldots, \alpha_n) = \|u - (\alpha_1 v_1 + \cdots + \alpha_n v_n)\|^2,$$

which we want to minimize. Note that when we expand this expression because the inner product being used to define the norm is complex conjugate in the first entry. Thus

$$\begin{aligned}f(\alpha_1,\ldots,\alpha_n) = \|u\|^2 &- \overline{\alpha}_1\langle v_1,u\rangle - \cdots - \overline{\alpha}_n\langle v_n u\rangle \\ &- \alpha_1\overline{\langle v_1,u\rangle} - \cdots - \alpha_n\overline{\langle v_n u\rangle} \\ &+ |\alpha_1|^2\|v_1\|^2 + \cdots + |\alpha_n|^2\|v_n\|^2\end{aligned}$$

Now we need to be careful because $\alpha_1,\ldots,\alpha_n$ are complex numbers. Thus in order to use methods of multivariable calculus for optimization we need to convert the problem to a problem of real variables. We can do this by writing $\alpha_1 = a_1 + ib_1$, etc., and then setting

$$\frac{\partial f}{\partial a_k} = 0 \quad \text{and} \quad \frac{\partial f}{\partial b_k} = 0 \tag{3.10}$$

first-complex-o

for each $k = 1,\ldots,n$.

For computational simplicity we compute, using the chain rule, in terms of the variables $\alpha_k = a_k + ib_k$ and $\overline{\alpha}_k = a_k - ib_k$. Using the chain rule we have

$$\begin{aligned}\frac{\partial f}{\partial a_k} &= \frac{\partial f}{\partial \alpha} + \frac{\partial f}{\partial \overline{\alpha}} \\ \frac{\partial f}{\partial b_k} &= i\left(\frac{\partial f}{\partial \alpha} - \frac{\partial f}{\partial \overline{\alpha}}\right)\end{aligned}$$

Thus (3.10) is equivalent to the condition that

$$\frac{\partial f}{\partial \alpha_k} = 0 \quad \text{and} \quad \frac{\partial f}{\partial \overline{\alpha}_k} = 0$$

for each $k = 1,\ldots,n$.

In order to compute the derivatives of f, note that for each k we have $|\alpha_k|^2 = \overline{\alpha}_k\alpha_k$. Thus

$$\frac{\partial f}{\partial \alpha_k} = -\overline{\langle v_k,u\rangle} + \overline{\alpha}_k\|v_k\|^2 \qquad \frac{\partial f}{\partial \overline{\alpha}_k} = -\langle v_k,u\rangle + \alpha_k\|v_k\|^2.$$

Setting each of these equal to zero we conclude that the optimal choice of α_k is

$$\alpha_k = \frac{\langle v_k,u\rangle}{\|v_k\|^2}.$$

Thus the best approximation of u is given by

$$u_n = \frac{\langle v_1,u\rangle}{\|v_1\|^2}v_1 + \cdots + \frac{\langle v_k,u\rangle}{\|v_k\|^2}v_k + \cdots + \frac{\langle v_n,u\rangle}{\|v_n\|^2}v_n.$$

It is important to note the order in the inner product: in the complex setting the order matters because the inner product is conjugate symmetric, not symmetric.

For the purposes of this course, the most important example of orthogonal functions in a complex inner product space are the functions

$$\phi_k(x) = e^{i\frac{k\pi}{L}x} \qquad k = \dots, -2, -1, 0, 1, 2, \dots \tag{3.11}$$ complex-periodic-eigenfun

introduced in Exercise 3.6.1. The approximation of a function determined by the functions (3.11) is called the ***(periodic) Fourier series***[5]. Given a function u in $\mathcal{L}^2([-L.L])$, Fourier series is constructed by first computing the ***Fourier coefficients***

$$c_k = \frac{\langle \phi_k, u\rangle}{\|\phi_k\|^2} = \frac{1}{2L}\int_{-L}^{L} e^{-i\frac{k\pi}{L}y} u(y)\, dy. \tag{3.12}$$ FS-coefficient-formula

These coefficients are used to form the associated Fourier series

$$\tilde{u} = \sum_{k=-\infty}^{\infty} c_k \phi_k.$$

example:FS-line **Example 3.19.** *Consider the function $u(x) = x$. We compute the Fourier coefficients*

$$c_k = \frac{1}{2L}\int_{-L}^{L} e^{-i\frac{k\pi}{L}y} x\, dx = \begin{cases} (-1)^k \frac{L}{k\pi} i & \text{if } k \neq 0 \\ 0 & \text{if } k = 0. \end{cases}$$

Thus the Fourier series for u is

$$\tilde{u}(x) = \sum_{k=-\infty}^{-1} (-1)^k \frac{L}{k\pi} i e^{i\frac{k\pi}{L}x} + \sum_{k=1}^{\infty} (-1)^k \frac{L}{k\pi} i e^{i\frac{k\pi}{L}x}$$

Notice that we can re-index the first sum to obtain an expression that only involves real-valued functions.

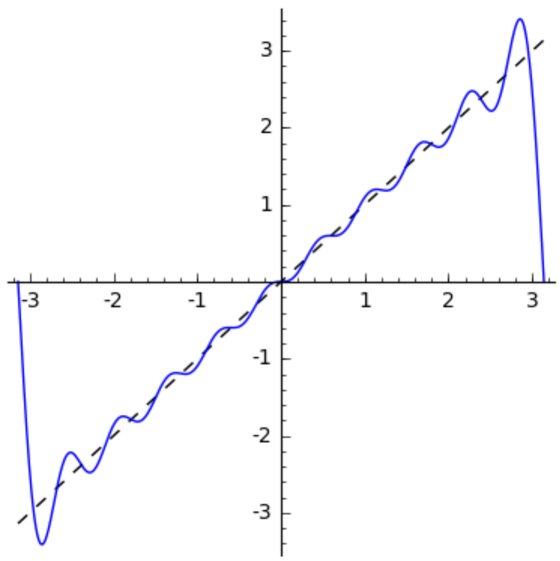
$$\tilde{u}(x) = \sum_{k=1}^{\infty} (-1)^k \frac{L}{k\pi} i \left(e^{i\frac{k\pi}{L}x} - e^{-i\frac{k\pi}{L}x}\right) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{2L}{k\pi} \sin\left(\frac{k\pi}{L}x\right)$$

The following Sage code generates a plot of the function u (dashed black line) and the $n = 10$ partial sum of the Fourier series; here we have set $L = \pi$ for the purposes of plotting.

[5] Here the word periodic is in parentheses because this is the "default" Fourier series. If someone refers to Fourier series without specifying boundary conditions, they mean the complex periodic series.

```
var('x,k,n')
n=10
f(x) = x
fplot = plot(f,(x,-pi,pi),figsize=[4,2],color='black',
    linestyle="dashed")

fs(x) = sum((2*((-1)^(k+1))/k)*sin(k*x),k,1,n)
fsplot = plot(fs,(x,-pi,pi),figsize=[4,4])

show(fplot+fsplot)
```

The resulting graphic is



example:FS-parabola **Example 3.20.** *Consider the function $v(x) = x^2$. Using the Sage code*

```
var('y,k,L')
ck = integrate( exp(-I*k*pi*y/L)*y^2, (y,-L,L))
show(ck)
```

we compute the Fourier coefficients

$$c_k = \frac{1}{2L}\int_{-L}^{L} e^{-i\frac{k\pi}{L}y}\,dy = \begin{cases} \frac{(-1)^k 2L^2}{\pi^2 k^2} & \text{if } k \neq 0 \\ \frac{L^2}{3} & \text{if } k = 0. \end{cases}$$

Thus the Fourier series for u is

$$\tilde{v}(x) = \frac{L^2}{3} + \sum_{k\neq 0} \frac{(-1)^k 2L^2}{\pi^2 k^2} e^{i\frac{k\pi}{L}x}.$$

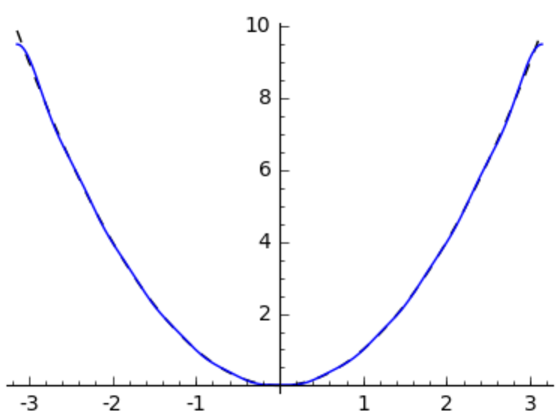
Re-indexing the sum, we can write this as

$$\tilde{v}(x) = \frac{L^2}{3} + \sum_{k=1}^{\infty} \frac{(-1)^k 4L^2}{k^2\pi^2} \cos\left(\frac{k\pi}{L}x\right)$$

The following Sage code generates a plot of the function u (dashed black line) and the $n = 10$ partial sum of the Fourier series; here we have set $L = \pi$.

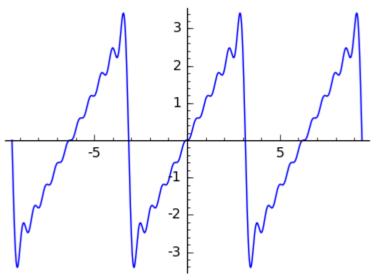
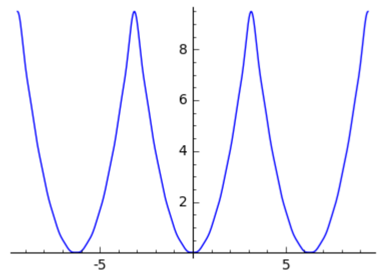
```
var('x,k,n')
n=10
f(x) = x^2
fplot = plot(f,(x,-pi,pi),figsize=[4,2],color='black',
    linestyle="dashed")

fs(x) = pi^2/3 + sum(((4*(-1)^k)/(k^2))*cos(k*x),k,1,n)
fsplot = plot(fs,(x,-pi,pi),figsize=[4,3])

show(fplot+fsplot)
```

The resulting graphic is

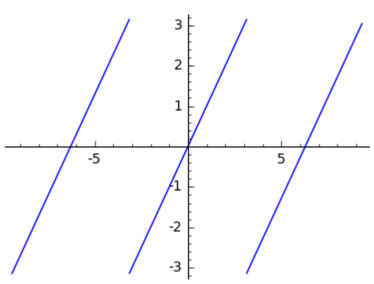
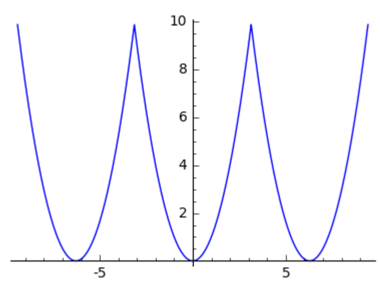


There are two important things to note about Examples 3.19 and 3.20. The first is that in both cases we were able to express the Fourier series as a sum of real functions: in the first case the sum is in terms of sine functions; in the second case it is in terms of cosine functions. In general, if the function u we start with is real-valued, then the Fourier series is able to be expressed in a manner which makes apparent that it, too, is real-valued. See Exercise 3.7.2 for a further exploration of this.

The second important thing to note about Examples 3.19 and 3.20 is the behavior at the endpoints of the domain. In Example 3.20, the Fourier series approximation is quite good, even when we only plot the first 10 terms. In Example 3.19, the behavior at the endpoints is a bit strange. To better understand what's going on here, notice that the functions (3.11) are all $2L$-periodic, meaning that $\phi_k(x+2L) = \phi_k(x)$ for any real number x. Thus the Fourier series approximation of any function is also $2L$-periodic. To see this, consider the following graphic, which contains the plots of the Fourier series approximations from the two examples, now shown on the domain $[-3\pi, 3\pi]$:

The functions that these plots are approximating are the ***periodic extensions*** of the functions $u(x) = x$ and $v(x) = x^2$, which are obtained by extending the domain from $[-L, L]$ to all real numbers under the condition that the functions are $2L$-periodic. The plots of these periodically extended functions are

The Sage code used to generate these plots is:

```
var('x')

f1(x) = piecewise([((-3*pi,-pi),x+2*pi),((-pi,pi), x),((pi
    ,3*pi), x-2*pi)])
f1plot = plot(f1,(x,-3*pi,3*pi),exclude=(-pi, pi))

f2(x) = piecewise([((-3*pi,-pi),(x+2*pi)^2),((-pi,pi), x^2)
    ,((pi,3*pi), (x-2*pi)^2)])
```

```
f2plot = plot(f2,(x,-3*pi,3*pi))

show(f1plot,figsize=[4,3])
#show(f2plot,figsize=[4,3])
```

These latter plots shed light on the behavior at the endpoints seen in Example 3.19. The Fourier series approximation is actually approximating the periodic extension of the function $u(x) = x$, which is not continuous at $x = \pm L$. Rather, the periodic extension jumps at $x = \pm L$, and thus the Fourier series approximation does as well.

-Fourier-series ★**Exercise 3.7.1.** *Find the (periodic) Fourier series for the following functions. Have Sage make a plot of the approximations.*

1. *The "square"*
$$u(x) = \begin{cases} -1 & \text{if } -L < x < 0 \\ 1 & \text{if } 0 < x < L. \end{cases}$$

2. *The "triangle"*
$$u(x) = \begin{cases} L + x & \text{if } -L < x < 0 \\ L - x & \text{if } 0 < x < L. \end{cases}$$

3. *The "sawtooth"*
$$u(x) = \begin{cases} x + L & \text{if } -L < x < 0 \\ x & \text{if } 0 < x < L. \end{cases}$$

4. *The "pulse of size a"*
$$u(x) = \begin{cases} 0 & \text{if } |x| > a \\ \frac{1}{2a} & \text{if } |x| < a, \end{cases}$$

   where a is some positive number less than L.

• needed

S-periodic-trig **Exercise 3.7.2.** *Suppose that u is a real-valued function in $\mathcal{L}^2([-L, L])$ and that*
$$\tilde{u}(x) = \sum_{k=-\infty}^{\infty} c_k e^{i\frac{k\pi}{L}x}$$
is the corresponding Fourier series.

1. *Show that*

$$\tilde{u}(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos\left(\frac{k\pi}{L}x\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{k\pi}{L}x\right),$$

where

$$a_0 = \frac{1}{2L}\int_{-L}^{L} u(y)\,dy$$

and, for $k = 1, 2, 3, \dots,$

$$a_k = \frac{1}{L}\int_{-L}^{L} u(y)\cos\left(\frac{k\pi}{L}x\right)\,dy, \qquad b_k = \frac{1}{L}\int_{-L}^{L} u(y)\sin\left(\frac{k\pi}{L}x\right)\,dy.$$

2. *Find the relationship between the coefficients* a_k, b_k *and the complex Fourier coefficients* c_k.

3.8 The periodic IBVP

We now address the initial boundary value problem on the domain $[-L, L]$ with periodic boundary conditions. Given functions u_0, v_0 defined on the interval $[-L, L]$, we seek a function $u(t, x)$, defined for $t \geq 0$ and $-L \leq x \leq L$ such that:

- the initial conditions $u(0, x) = u_0(x)$ and $\partial_t u(0, x) = v_0(x)$ are satisfied for $-L < x < L$,
- the periodic boundary conditions

$$u(t, -L) = u(t, L) \quad \text{and} \quad \partial_x u(t, -L) = \partial_x u(t, L)$$

  hold for all times $t > 0$, and
- the wave equation $\partial_t^2 u(t, x) = \partial_x^2 u(t, x)$ is satisfied for $t > 0$ and $-L < x < L$.

We begin by looking for standing wave solutions to the wave equation with periodic boundary conditions. Inserting $u(t, x) = A(t)\psi(x)$ in to the wave equation we see that

$$\frac{1}{A}\frac{d^2A}{dt^2} = \frac{1}{\psi}\frac{d^2\psi}{dx^2}.$$

Thus, as in §2.4, we find that A and ψ must satisfy

$$\frac{d^2A}{dt^2} = \lambda A$$

$$\frac{d^2\psi}{dx^2} = \lambda\psi, \quad \psi(-L) = \psi(L) \quad \psi'(-L) = \psi'(L)$$

for some constant λ. From Exercise 2.4.1, we conclude that the solutions to the ψ equation are

$$\psi_k(x) = e^{i\frac{k\pi}{L}x}, \quad k = \cdots - 2, -1, 0, 1, 2, \ldots$$

with $\lambda_k = -\omega_k^2 = -\left(\frac{k\pi}{L}\right)^2$. Thus we obtain complex-valued standing wave solutions

$$u_k(t,x) = \cos\left(\frac{k\pi}{L}t\right)e^{i\frac{k\pi}{L}x} \qquad v_k(t,x) = \sin\left(\frac{k\pi}{L}t\right)e^{i\frac{k\pi}{L}x}.$$

Consequently, the most general solution is

$$u(t,x) = \sum_{k=-\infty}^{\infty} a_k \cos\left(\frac{k\pi}{L}t\right)e^{i\frac{k\pi}{L}x} + \sum_{k=-\infty}^{\infty} b_k \sin\left(\frac{k\pi}{L}t\right)e^{i\frac{k\pi}{L}x}.$$

We can now use Fourier series introduced in Exercise 3.7.1 can be used to address the initial boundary value problem for the one-dimensional wave equation with periodic boundary conditions. In particular, we can choose the constants a_k and b_k to match the initial conditions.

Exercise 3.8.1 (Relies on Exercise 3.7.1)**.** *Find the solution to the wave equation satisfying periodic boundary conditions and with initial condition u_0 being the "square" and $v_0 = 0$. Express the solution both in terms of complex exponentials and as a real-valued function.*

Exercise 3.8.2 (Relies on Exercise 3.7.1)**.** *Find the solution to the wave equation satisfying periodic boundary conditions and with initial condition u_0 being the "triangle" and $v_0 = 0$. Express the solution both in terms of complex exponentials and as a real-valued function.*

Exercise 3.8.3 (Relies on Exercise 3.7.1)**.** *Find the solution to the wave equation satisfying periodic boundary conditions and with initial condition u_0 being the "sawtooth" and $v_0 = 0$. Express the solution both in terms of complex exponentials and as a real-valued function.*

Exercise 3.8.4 (Relies on Exercise 3.7.1)**.** *Find the solution to the wave equation satisfying periodic boundary conditions and with initial condition u_0 being the "pulse of size a" and $v_0 = 0$. Express the solution both in terms of complex exponentials and as a real-valued function.*

3.9 Convergence for Fourier series (optional)

In order to show convergence for Fourier series we first address the case of periodic Fourier series. Then we show how convergence for the periodic Fourier series implies convergence for the Fourier sine series.

The main convergence result is this:

> *Let u be a real-valued function in $\mathcal{L}^2([-L, L])$. For each x in the interval $[-L, L]$ define the left and right limits*
>
> $$u(x^-) = \lim_{y \to x^-} u(y) \qquad u(x^+) = \lim_{y \to x^+} u(y),$$
>
> *where at $x = \pm L$ we use the periodic extension of u to define the limit coming from outside $[-L, L]$. Let c_k be the (complex, periodic) Fourier coefficients of u. Then for each x in $[-L, L]$ we have*
>
> $$\lim_{n\to\infty} \sum_{k=-n}^{n} c_k e^{i\frac{k\pi}{L}x} = \frac{1}{2}\left(u(x^-) + u(x^+)\right).$$
>
> *We express this convergence by the notation*
>
> $$u(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{i\frac{k\pi}{L}x}.$$

In other words, the Fourier series approximation converges to the value of the (periodic extension of the) function u at all points where u is continuous, and converges to the average of the left and right limits of the function at points where the periodic extension of u jumps.

The proof of this convergence result involves the following steps.

1. First we show that for any function u we have $\langle \phi_k, u\rangle \to 0$ as $k \to \pm\infty$. This result is known as the Riemann-Lebesgue lemma[6].
2. We then re-write the Fourier series of u in terms of a special function known as the Dirichlet kernel.
3. We then show that the properties of the Dirichlet kernel allow us to re-express the problem in a way that allows us to use the Riemann-Lebesgue lemma.

[6] This is actually a special case of the more general Riemann-Lebesgue lemma.

To accomplish the first step, let u be a function in $\mathcal{L}^2([-L, L])$ and let u_n be the Fourier approximation

$$u_n = \sum_{k=-n}^{n} c_k \phi_k,$$

where c_k are the Fourier coefficients and ϕ_k are given by (3.11). Let $u_{\text{rem}} = u - u_n$. Using the formula for the coefficients c_k and the orthogonality of the functions ϕ_k we have

$$\begin{aligned}
\langle u_{\text{rem}}, u_n \rangle &= \langle u, u_n \rangle - \langle u_n, u_n \rangle \\
&= \sum_{k=-n}^{n} \left(c_k \langle u, \phi_k \rangle - |c_k|^2 \|\phi_k\|^2 \right) \\
&= \sum_{k=-n}^{n} \left(\frac{\langle \phi_k, u \rangle \overline{\langle \phi_k, u \rangle}}{\|\phi_k\|^2} - \frac{\langle \phi_k, u \rangle \overline{\langle \phi_k, u \rangle}}{\|\phi_k\|^4} \|\phi_k\|^2 \right) \\
&= 0.
\end{aligned}$$

Thus

$$\|u\|^2 = \langle u_n + u_{\text{rem}}, u_n + u_{\text{rem}} \rangle = \|u_n\|^2 + \|u_{\text{rem}}\|^2 \geq \|u_n\|^2.$$

The important part of this last inequality is that the norm of $\|u_n\|^2$ is bounded above by a constant (the square of the norm of u) that is independent of n.

The orthogonality of the ϕ_k, and the fact that $\|\phi_k\|^2 = 2L$, implies that

$$\|u_n\|^2 = 2L \sum_{k=-n}^{n} |c_k|^2 = \sum_{k=-n}^{n} |\langle \phi_k, u \rangle|^2.$$

Thus we have

$$\sum_{k=-n}^{n} |\langle \phi_k, u \rangle|^2 \leq \|u\|^2.$$

This means that the partial sums of the series

$$\sum_{k=-\infty}^{\infty} |\langle \phi_k, u \rangle|^2$$

are uniformly bounded and thus the series converges. Since the series converges, the terms $|\langle \phi_k, u \rangle|^2$ in the series much tend to zero as $k \to \pm\infty$. This establishes the Riemann-Lebesgue lemma.

We now proceed to the second step. Fix a function $u \in \mathcal{L}^2([-L, L])$ and again let u_n be the partial sum of the Fourier series:

$$u_n(x) = \sum_{k=-n}^{n} c_k e^{i\frac{k\pi}{L}x},$$

where

$$c_k = \frac{1}{2L} \int_{-L}^{L} e^{-i\frac{k\pi}{L}y} u(y)\, dy.$$

Thus

$$u_n(x) = \int_{-L}^{L} D_n(x-y) u(y)\, dy, \tag{3.13}$$

FS-converge-fir

where the function D_n is defined by

$$D_n(z) = \frac{1}{2L} \sum_{k=-n}^{n} e^{i\frac{k\pi}{L}z}.$$

The functions D_n are called the ***Dirichlet kernel***.

There are a number of useful properties of the Dirichlet kernel. Note that

$$D_n(z) = \frac{1}{2L} + \sum_{k=1}^{n} \cos\left(\frac{k\pi}{L}z\right).$$

This implies that D_n is an even, real-valued function with

$$\int_{-L}^{0} D_n(z)\, dz = \frac{1}{2} = \int_{0}^{L} D_n(z)\, dz. \tag{3.14}$$

Dirichlet-kerne

It further implies that D_n is $2L$-periodic. We can also use the partial sum formula for geometric series to write the Dirichlet kernel as

$$D_n(z) = \frac{1}{2L} e^{-i\frac{n\pi}{L}z} \sum_{k=0}^{2n} \left(e^{i\frac{\pi}{L}z}\right)^k = \frac{1}{2L} \frac{e^{i\frac{(n+1)\pi}{L}z} - e^{-i\frac{n\pi}{L}z}}{e^{i\frac{\pi}{L}z} - 1}. \tag{3.15}$$

Dirichlet-kerne

We now extend u to a function on all of the real line by requiring that it be $2L$-periodic. Then making the change of variables $y = x + z$ in (3.13) and using the fact that D_n is even yields

$$u_n(x) = \int_{-L}^{L} D_n(z) u(x+z)\, dz.$$

Thus from (3.14) we have

$$u_n(x) - \frac{1}{2}\left(u(x^-) + u(x^+)\right)$$
$$= \int_{-L}^{0} D_n(z)\left(u(x+z) - u(x^-)\right)\,dz$$
$$+ \int_{0}^{L} D_n(z)\left(u(x+z) - u(x^+)\right)\,dz$$

Using (3.15) we may express this as

$$u_n(x) - \frac{1}{2}\left(u(x^-) + u(x^+)\right) = \langle \phi_{-(n+1)}, U\rangle - \langle \phi_n, U\rangle,$$

where the function U is defined by

$$U(x) = \begin{cases} \dfrac{u(x+z) - u(x^-)}{2L(e^{i\frac{\pi}{L}z} - 1)} & \text{if } -L \le x < 0, \\[2ex] \dfrac{u(x+z) - u(x^+)}{2L(e^{i\frac{\pi}{L}z} - 1)} & \text{if } 0 < x \le L. \end{cases}$$

Using l'Hopital's rule we see that U is in $\mathcal{L}^2([-L, L])$. Thus we may apply the Riemann-Lebesgue lemma to conclude that

$$\langle \phi_{-(n+1)}, U\rangle \to 0 \quad \text{and} \quad \langle \phi_n, U\rangle \to 0$$

as $n \to \infty$. This establishes the main convergence result.

To see how the main convergence result implies convergence for the Fourier sine series, let u be a function in $\mathcal{L}^2([0, L])$. We extend the function u to a function in $\mathcal{L}^2([-L, L])$ by requiring that u be an odd function; ie, $u(-x) = -u(x)$. We know from the main convergence result that the complex Fourier series for u converges on the domain $[-L, L]$. Using Exercise 3.7.2, we can express the complex Fourier series in terms of cosines and sines. But the fact that u is an odd function implies that the coefficients a_k in that exercise are all zero. (The integral of an odd function over a symmetric domain is zero.) Thus the Fourier series for u is a series of only sine functions. These sine functions are precisely the Dirichlet eigenfunctions ψ_k, and thus restricting the complex Fourier series to the domain $[0, L]$ gives us the Fourier sine series. Since the complex Fourier series converges, so does the Fourier sine series.