Notes on Fourier Analysis

pqnelson@gmail.com

December 17, 2008

Contents

List of Figures

Part I

Fourier Series

1 Differential Equations: A Motivation for Fourier Series

Example 1.1. (Diffusion of heat in solid body)

The initial motivation behind Fourier series was to solve the heat equation. The specific problem was trying to model the flow of heat in a thin metal rod, such as the one seen in figure (1.1). The equation describing it requires us first to consider several variables. For simplicity, consider only a one dimensional "thing" insulated metal rod with length L.

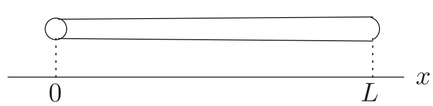

Figure 1.1: A thin insulated metal rod of length L

We can describe the problem with some conditions

$$u(x,t) = \text{temperature at point } x \text{ and time } t \tag{1.1a}$$

$$u(x,0) = f(x) \quad \text{(temperature at time } t=0\text{, initial condition)} \tag{1.1b}$$

$$u(0,t) = u(L,t) = 0 \quad \text{(boundary condition)} \tag{1.1c}$$

Observe that we hold the temperature at the endpoints fixed to be zero at all time. This is weird physically, because we rarely see such a system. So we can describe the heat flow in a thin metal rod by the equation

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2} \tag{1.2}$$

where k is the diffusitivity of the rod (how quickly heat can spread), and this is known as the heat equation. Now that we have described the dynamics of the heat flow for this thin metal rod, solving it should be trivial...right?

Skip work, here's answer

Well, without the boundary conditions and initial conditions, one can struggle for a long long long time to find that

$$u(x,t) = t^{-1/2}\exp\left(\frac{-x^2}{4kt}\right) \tag{1.3}$$

is a solution for $t > 0$. So we would just need to solve the boundary value problem. This is not as trivial as it seems.

seperation of variables

Our toolkit from taking basic differential equations (which is: "Guess really really good!") is inadequate to solve this problem. We need to introduce a new technique which we call **seperation of variables**. What we do is we assume we can write the function $u(x,t)$ as

$$u(x,t) = X(x)T(t) \tag{1.4}$$

This usually works but not always[1]. We will not investigate it here but it should be noted that it may not always work.

The second step to the "seperation of variables" technique is to plug in (1.4) into the differential equation

$$\partial_t u(x,t) = k\partial_x^2 u(x,t) \tag{1.5a}$$

$$\partial_t(X(x)T(t)) = k\partial_x^2(X(x)T(t)) \tag{1.5b}$$

$$X(x)T'(t) = kX''(x)T(t) \tag{1.5c}$$

Don't believe it? Note the left equation is independent of x, the middle equation is independent of t, so take the t derivative of everything and you get the middle equation being 0, similar results for the x derivative imply that they're equal to some constant.

and at first this seems useless. What do we do with this result? Well, we divide both sides by $kX(x)T(t)$ and set it to be equal to some constant

$$\frac{T'(t)}{kT(t)} = \frac{X''(x)}{X(x)} = \begin{pmatrix}\text{some constant}\\ \text{to be determined}\end{pmatrix} \tag{1.6}$$

Observe that we have on the left hand side a function of t only, it is independent of x. Similarly, the middle term is independent of t. That is, we note

$$\underbrace{\frac{T'(t)}{kT(t)}}_{\text{independent of } x} = \underbrace{\frac{X''(x)}{X(x)}}_{\text{independent of } t} = \begin{pmatrix}\text{some constant}\\ \text{to be determined}\end{pmatrix}$$

[1] It should be noted that this is a coordinate-choice, so there are a finite number of coordinate systems we can do this in (e.g. Morse et al. [1] point out that the Hemholtz equation, for example, can only be solved via seperation of variables in 11 coordinate systems).

Now, we can solve for the general solution of $T(t)$

$$\frac{T'(t)}{T(t)} = kA \Rightarrow \log|T(t)| = kAt + c_1 \Rightarrow T(t) = T_0 e^{kAt} \tag{1.7}$$

where $T_0 = T(0)$ is some constant.

Now, for solving $X(x)$, we have a bit more difficult of a problem. The differential equation of relevance is

$$\frac{X''(x)}{X(x)} = A \Rightarrow X''(x) - AX(x) = 0 \tag{1.8}$$

Recall for general second order differential equations

$$y'' + ay' + by = 0 \tag{1.9}$$

we have the **characteristic equation**

$$r^2 + ar + b = 0. \tag{1.10}$$

It has two roots, by the fundamental theorem of Algebra. So we have the two roots r_1 and r_2 which gives us that general solution

Fundamental Thm Of Algebra: A polynomial of degree n has exactly n complex roots

$$y(x) = c_1 e^{r_1 x} + c_2 e^{r_2 x}. \tag{1.11}$$

For us, we have a simple characteristic equation

$$r^2 - A = 0 \Rightarrow r = \pm\sqrt{A} \tag{1.12}$$

thus the general solution is

$$X(x) = c_1 e^{x\sqrt{A}} + c_2 e^{-x\sqrt{A}} \tag{1.13}$$

So now we have a new problem: what values of A give nonzero solutions? That is, which ones satisfy the boundary conditions

What's A?

$$X(0) = X(L) = 0 \tag{1.14}$$

without $X(x)$ being trivial (i.e. $X(x) = 0$).

If $A > 0$, then we have by one boundary condition

$$X(0) = c_1 + c_2 = 0 \Rightarrow c_1 = -c_2. \tag{1.15}$$

And by the other boundary condition

$$X(L) = c_1 e^{L\sqrt{A}} + c_2 e^{-L\sqrt{A}} = 0 \Rightarrow c_1 e^{2L\sqrt{A}} + c_2 = 0 \tag{1.16}$$

and by putting these two together we find

$$c_1(1 - e^{2L\sqrt{A}}) = 0 \tag{1.17}$$

which requires either $c_1 = 0$ or $\exp(2L\sqrt{A}) = 1$. The latter implies either $L = 0$ or $A = 0$. We assumed that $L > 0$ and $A > 0$. So that implies $c_1 = 0$. Thus the general solution is

$$X(x) = c_2 e^{-x\sqrt{A}} \tag{1.18}$$

which is really just exponential decay.

For the $A < 0$ case, the general solution can be written as

$$X(x) = c_1 \cos(\lambda x) + c_2 \sin(\lambda x) \tag{1.19}$$

where $\lambda = \sqrt{-A}$. We plug in the boundary conditions

$$X(0) = c_1 = 0 \tag{1.20}$$

and

$$X(L) = c_2 \sin(\lambda L) = 0 \Rightarrow \lambda L = n\pi \Rightarrow \lambda = \frac{n\pi}{L} \tag{1.21}$$

where $n \in \mathbb{Z}$. So here we have $A = -n^2\pi^2/L^2$.

Putting this all together, the general solution to the boundary value problem is

$$u_n(x,t) = \left(c_0 \exp(-k\frac{n^2\pi^2}{L^2}t)\right)\left(c_1 \sin(\frac{n\pi}{L}x\right) \tag{1.22}$$

But also observe that *any linear combinations* of these solutions *are also solutions* to the boundary value problem. So the general solution is

$$u(x,t) = \sum_{n=1}^{\infty} a_n e^{-kn^2\pi^2 t/L^2} \sin\left(\frac{n\pi}{L}x\right) \tag{1.23}$$

where a_n are arbitrary constants.

"*How to make $u(x,t)$ satisfy the initial condition $u(x,0) = f(x)$?*" Well, we have a small substitution to make, we just set (1.23) equal to the given boundary condition

$$\sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi}{L}x\right) = f(x) \tag{1.24}$$

Point to ponder

This is where Fourier works his magic, expanding a function in terms of sines and cosines. Is it possible to find an expansion for *any* heat distribution?

2 A First Look at Fourier Series

Remember Euler's formula

$$e^{i\theta} = \cos(\theta) + i\sin(\theta). \tag{2.1}$$

Observe that

$$e^{-i\theta} = e^{i(-\theta)} = \cos(-\theta) + i\sin(-\theta) \tag{2.2}$$

and

$$\overline{e^{i\theta}} = e^{-i\theta} = \cos(\theta) - i\sin(\theta). \tag{2.3}$$

We set $e^{-i\theta} = e^{-i\theta}$ and find from the real part

$$\cos(-\theta) = \cos(\theta) \tag{2.4}$$

and from the imaginary part

$$\sin(-\theta) = -\sin(\theta). \tag{2.5}$$

We call $\cos(\theta) \approx 1 - \theta^2/2 + ...$ an **even function** and $\sin(\theta) \approx \theta - \theta^3/6 + ...$ an **odd function**.

Definition 2.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a function, then f is **periodic with period** p (or p-**periodic**) if

$$f(x+p) = f(x) \tag{2.6}$$

holds.

Example 2.1. The functions $\cos(\theta)$, $\sin(\theta)$ and $\exp(i\theta)$ are all 2π-periodic.

Lemma 2.1. If we have a p-periodic function f, then

$$\int_a^{a+p} f(x)dx \tag{2.7}$$

is independent of a.

Proof. If we change a by some c (so $a \to a + c$) then

$$\int_{a+c}^{a+c+p} f(x)dx = \int_{a+c}^{a+p} f(x)dx + \int_{a+p}^{a+c+p} f(x)dx \quad \text{(by linearity of integral)} \tag{2.8a}$$

$$= \int_{a+c}^{a+p} f(x)dx + \int_{a}^{a+c} f(x)dx \quad \text{(by periodicity of } f(x)) \tag{2.8b}$$

$$= \int_{a}^{a+p} f(x)dx \quad \text{(collecting terms)} \tag{2.8c}$$

where we justify the second step (the one justified in a handwavy way with the "periodicity of f") by using the Riemann sum and noting such reasoning works for that particular definition of the integral and we use it here too. ☐

Definition 2.2. Let $f : \mathbb{R} \to \mathbb{R}$ be 2π-periodic and integrable on $[-\pi, \pi]$. The **Fourier Series** f is

$$f(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} \tag{2.9}$$

where

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-i\theta}d\theta \tag{2.10}$$

is called the **Fourier Coefficient** of f.

Remark 2.1. This is called the "**exponential form**" of the Fourier Series. There are other forms of the Fourier series.

We can equivalently write the Fourier series in a slightly different approach,

$$f(\theta) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n\cos(n\theta) + b_n\sin(n\theta)) \tag{2.11}$$

where

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)d\theta \tag{2.12a}$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos(n\theta)d\theta \tag{2.12b}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin(n\theta)d\theta \tag{2.12c}$$

Observe that

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\left(\frac{e^{in\theta}+e^{-in\theta}}{2}\right)d\theta \tag{2.13a}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)(e^{in\theta}+e^{-in\theta})d\theta \tag{2.13b}$$

$$= c_n + c_{-n} \tag{2.13c}$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\left(\frac{e^{in\theta}-e^{-in\theta}}{2i}\right)d\theta \tag{2.14a}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)(e^{in\theta}-e^{-in\theta})d\theta \tag{2.14b}$$

$$= i(c_n - c_{-n}) \tag{2.14c}$$

thus putting it all together we get

$$c_n = \frac{a_n - ib_n}{2}. \tag{2.15}$$

So really, we are starting to see a connection between the two forms.

The equivalence of the two forms can be explicity written

$$\sum_{n=-\infty}^{\infty} c_n e^{in\theta} = c_0 + \sum_{n=1}^{\infty} c_n e^{in\theta} + c_{-n} e^{-in\theta} \tag{2.16a}$$

$$= c_0 + \sum_{n=1}^{\infty} c_n(\cos(n\theta) + i\sin(n\theta)) + c_{-n}(\cos(n\theta) - i\sin(n\theta) \tag{2.16b}$$

$$= c_0 + \sum_{n=1}^{\infty} (c_n + c_{-n})\cos(n\theta) + i(c_n - c_{-n})\sin(n\theta) \tag{2.16c}$$

$$= c_0 + \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta) \tag{2.16d}$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta) \tag{2.16e}$$

which explicitly demonstrates the equivalence between the two forms.

2.1 Derivation of Definition of Fourier Series

We want to expand our periodic function in terms of sines and cosines

$$f(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} \tag{2.17}$$

So...what is c_n? Well, first we should probably remember that

$$\int_{-\pi}^{\pi} e^{inx} e^{-inx} dx = \int_{-\pi}^{\pi} dx = 2\pi \tag{2.18}$$

and if $m \in \mathbb{Z}$ such that $m \neq 0$ then

$$\int_{-\pi}^{\pi} e^{inx} e^{-i(n+m)x} dx = \int_{-\pi}^{\pi} e^{-imx} dx \tag{2.19a}$$

$$= \frac{1}{-im}(e^{-im\pi} - e^{im\pi}) \tag{2.19b}$$

$$= \frac{i}{m}((e^{-i\pi})^m - (e^{i\pi})^m) \tag{2.19c}$$

$$= \frac{i}{m}((-1)^m - (-1)^m) \tag{2.19d}$$

$$= \frac{i}{m}(0) = 0. \tag{2.19e}$$

So we found in general

$$\int_{-\pi}^{\pi} e^{inx} e^{-imx} dx = 2\pi\delta_{mn} \tag{2.20}$$

where δ_{mn} is the kronecker delta (it is 1 if $m = n$ and 0 otherwise).

So, if we then take

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{2.21}$$

then

$$\int_{-\pi}^{\pi} f(x)e^{-ikx}dx = \sum_{n=-\infty}^{\infty} c_n \int_{-\pi}^{\pi} e^{inx}e^{-ikx}dx \tag{2.22a}$$

$$= \sum_{n=-\infty}^{\infty} (2\pi\delta_{kn})c_n \tag{2.22b}$$

$$= (2\pi \cdot 1)c_k + \sum_{n\neq k}(2\pi \cdot 0)c_n \tag{2.22c}$$

$$= (2\pi)c_k \tag{2.22d}$$

so we conclude that

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}dx \tag{2.23}$$

which we already knew.

Remark 2.2. We can also define Fourier coefficients in a slightly different way. Instead of

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx \tag{2.24}$$

we can equivalently write

$$c_n = \int_0^1 f(x)e^{i2\pi nx}dx \tag{2.25}$$

where f becomes periodic with period 1.

Remark 2.3. The Fourier coefficients are constant, and the constant Fourier coefficient

$$c_0 = \frac{1}{2}a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx$$

gives us the average value of f over a 2π interval.

Example 2.2. Let f be a 2π periodic function given by

$$f(\theta) = |\theta|, \qquad -\pi \leq \theta \leq \pi \tag{2.26}$$

which is shown in figure 2.1. We can calculate its Fourier series simply by using

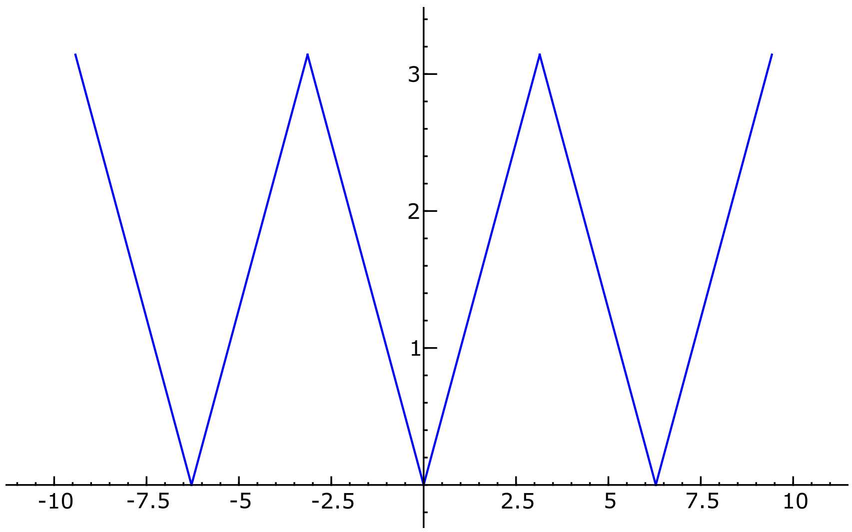

Figure 2.1: The Triangle Wave.

the a_n and b_n cofficients. Observe

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)d\theta \tag{2.27a}$$

$$= \frac{1}{\pi}\left(\int_0^{\pi} \theta d\theta + \int_{-\pi}^{0} -\theta d\theta\right) \tag{2.27b}$$

$$= \frac{2}{\pi}\int_0^{\pi} \theta d\theta \tag{2.27c}$$

$$= \frac{2}{\pi}\left(\frac{\theta^2}{2}\Big|_0^{\pi}\right) \tag{2.27d}$$

$$= \frac{2}{\pi}\left(\frac{\pi^2}{2}\right) \tag{2.27e}$$

$$= \pi \tag{2.27f}$$

Also observe that

$$f(-\theta) = f(\theta) \tag{2.28}$$

so the $b_n = 0$ for all $n \in \mathbb{Z}$. The only coefficients we have to calculate are the

a_n. Observe

$$\begin{aligned}
a_n &= \frac{2}{\pi}\int_0^{\pi} \theta\cos(\theta)d\theta && (2.29a)\\
&= \frac{\pi}{2}\left[\frac{\theta\sin(n\theta)}{n}\Big|_0^{\pi} - \int_0^{\pi}\frac{\sin(n\theta)}{n}d\theta\right] && (2.29b)\\
&= \frac{2}{\pi}\left[(0) + \frac{\cos(n\theta)}{n^2}\Big|_0^{\pi}\right] && (2.29c)\\
&= \frac{2}{\pi}\left(\frac{(-1)^n - 1}{n^2}\right) && (2.29d)\\
&= \begin{cases} 0 & n \text{ is even} \\ -4/(n^2\pi) & n \text{ is odd} \end{cases} && (2.29e)
\end{aligned}$$

Now we can plug these coefficients into the Fourier series to find

$$\begin{aligned}
f(\theta) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos(n\theta) && (2.30a)\\
&= \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos((2n-1)\theta)}{(2n-1)^2} && (2.30b)
\end{aligned}$$

Now that we have expanded it, the question becomes: *does the Fourier series converge?* The answer is yes. We observe that $\cos(\theta) \leq 1$ and

$$\frac{1}{(2n-1)^2} \leq \frac{1}{n^2} \tag{2.31}$$

thus term for term we have the inequality

$$\sum_{n=1}^{\infty}\frac{\cos((2n-1)\theta)}{(2n-1)^2} \leq \sum_{n=1}^{\infty}\frac{1}{n^2}. \tag{2.32}$$

We know that $\sum 1/n^2$ converges (for a handwavy argument if one is unconvinced: $\sum 1/n^2 \leq \int_1^{\infty}(1/n^2)dn < \infty$ so the sum is finite), and our Fourier series is less than this convergent sum, so it must converge (by the Weierstrass M-test).

3 Bessel's Inequality

Bessel Inequality. If f is 2π-periodic and Riemann integrable from $-\pi$ to π, and c_n are the Fourier coefficients of f, then

$$\sum_{n=-\infty}^{\infty}|c_n|^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(\psi)|^2 d\psi < +\infty \tag{3.1}$$

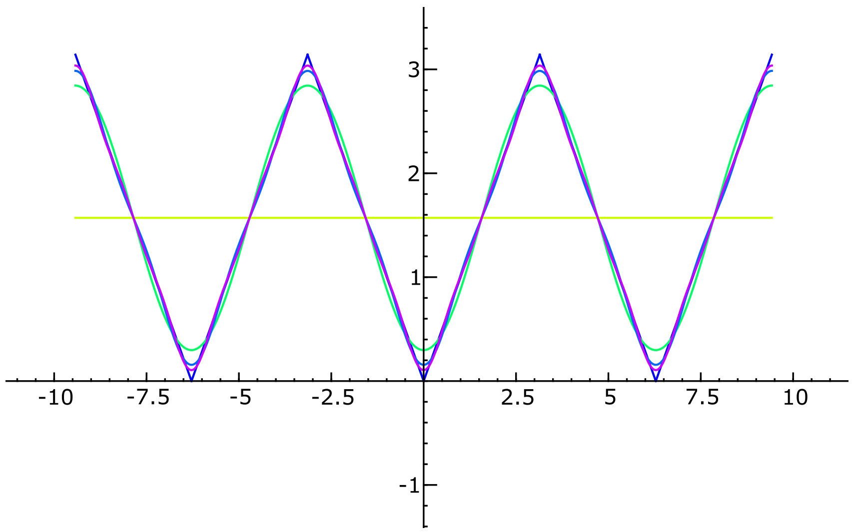

Figure 2.2: The Partial sum to the Triangle wave with 1 through 5 terms. The actual triangle wave is shown in blue.

This is important because it allows us to create a bound for the Fourier series that guarantees convergence. It should be noted that equivalently it can be shown

$$\frac{1}{4}|a_0|^2 + \frac{1}{2}\sum_{n=1}^{\infty}|a_n|^2 + |b_n|^2 = \sum_{n=-\infty}^{\infty}|c_n|^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(\psi)|^2 d\psi < +\infty. \tag{3.2}$$

Why? Well, recall

$$a_0 = 2c_0 \tag{3.3a}$$
$$a_n = c_n + c_{-n} \tag{3.3b}$$
$$b_n = i(c_n - c_{-n}) \tag{3.3c}$$

and also recall in complex analysis, $|z|^2 = z\bar{z}$ for $z \in \mathbb{C}$. We would like

$$|a_n|^2 \to 0 \qquad \text{as } |n| \to \infty \tag{3.4}$$

or

$$c_n \to 0 \quad \text{as } |n| \to \infty. \tag{3.5}$$

The question that arises is "What kind of functions are Riemann integrable and 2π-periodic?" *Riemann integrable and periodic?*

Let us first review a number of definitions which will simplify our life quite a bit.

Definition 3.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then the **left limit** of f at x is

$$f(x^-) = \lim_{y \to x^-} f(y) = \lim_{y \uparrow x} f(y) \tag{3.6}$$

(y is increasing to x). The **right limit** of f at x is

$$f(x^+) = \lim_{y \to x^+} f(y) = \lim_{y \downarrow x} f(y) \tag{3.7}$$

(y decreases to x).

Definition 3.2. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then f is **continuous** at x if

$$\lim_{y \to x^+} f(y) = \lim_{y \to x^-} f(y) < \infty. \tag{3.8}$$

More rigorously, for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|y - x| < \delta \Rightarrow |f(y) - f(x)| < \varepsilon. \tag{3.9}$$

Let us examine an example of the $\varepsilon - \delta$ definition of the limit.

Example 3.1. Let $f(x) = x^3$ and $x_0 \in \mathbb{R}$. We have

$$|x - x_0| < \delta \tag{3.10}$$

and in our function we have

$$|x^3 - x_0^3| < \varepsilon. \tag{3.11}$$

We need to determine δ in terms of ε. So we have

$$(x^2 + x_0 x + x_0^2)(x - x_0) = x^3 - x_0^3 \tag{3.12}$$

so

$$|x^3 - x_0^3| \leq |x^2 + x_0 x + x_0^2||x - x_0| \tag{3.13a}$$
$$\leq |x^2 + x_0 x + x_0^2|\delta \tag{3.13b}$$
$$< \varepsilon. \tag{3.13c}$$

Thus we set

$$\delta < \frac{\varepsilon}{|x^2 + x_0 x + x_0^2|} < \frac{\varepsilon}{x_0^2} \tag{3.14}$$

where the right hand side can be justified since we want the denominator to be as small as possible, so we have $x = x_0$ and the denominator becomes $3x_0^2$ but this is less than x_0^2 so we have a good bound. Thus we set

$$\delta = \frac{\varepsilon}{x_0^2} \tag{3.15}$$

if $x_0 \neq 0$. If $x_0 = 0$, then we have

$$\delta = \frac{\varepsilon}{x^2} \tag{3.16}$$

which is a *better* solution since it's *more general* than the first solution. That is, Eq (3.16) contains the specific case of Eq (3.15). But we have found a $\delta > 0$ in terms of $\varepsilon > 0$ that demonstrates continuity for arbitrary x_0.

Definition 3.3. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then f has a **jump discontinuity** at x_0 if f is discontinuous at x_0 but $f(x^+)$ and $f(x^-)$ both exist and are finite.

Definition 3.4. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then we say f is **piecewise continuous** on $\mathbb{R}$ if on any bounded interval $[a, b]$, where $-\infty < a < b < +\infty$, f has only jump discontinuities at finitely many points.

Definition 3.5. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We say that f is **piecewise smooth** on $\mathbb{R}$ (denoted $PS(\mathbb{R})$ if both f and f' are piecewise continuous on $\mathbb{R}$ (i.e. on $[a, b]$ both f and f' have only finitely many jump discontinuities).

These definitions are nothing really of interest, they are weaker than just being continuous, or just being smooth. Consequently, we will not review them in great detail.

3.1 Dirichlet Kernel

We want to consider partial sums of the Fourier series of a given function. To do this, we will introduce a mathematical object called the N^{th} **Dirichlet Kernel**

$$D_N(\phi) = \frac{1}{2\pi}\sum_{n=-N}^{N} e^{in\phi} \tag{3.17}$$

Observe that

$$\begin{aligned}
\int_{-\pi}^{0} D_N(\phi)d\phi &= \int_{-\pi}^{0}\left(\frac{1}{2\pi}\sum_{n=-N}^{N} e^{in\phi}\right)d\phi && \text{(3.18a)}\\
&= \frac{1}{2\pi}\phi + \frac{1}{\pi}\sum_{n=1}^{N}\frac{\sin(n\phi)}{n}\Big|_{-\pi}^{0} && \text{(3.18b)}\\
&= \frac{1}{2} + \frac{1}{\pi}\sum_{n=1}^{N}\frac{\sin(n\pi)}{n} && \text{(3.18c)}\\
&= \frac{1}{2} + 0 && \text{(3.18d)}
\end{aligned}$$

Similarly, we have

$$\int_{0}^{\pi} D_N(\phi)d\phi = \frac{1}{2} \tag{3.19}$$

But let us now see something quite miraculous, we can evaluate the partial sums of the Fourier series using the Dirichlet kernel. Observe

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)D_N(\theta-\phi)d\theta = \frac{1}{2\pi}\sum_{n=-N}^{N} e^{-in\phi}\int_{-\pi}^{\pi} f(\theta)e^{in\theta}d\theta \tag{3.20a}$$

$$= \sum_{n=-N}^{N} e^{-in\phi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{in\theta}d\theta\right) \tag{3.20b}$$

$$= \sum_{n=-N}^{N} e^{-in\phi}c_{-n} \tag{3.20c}$$

$$= \sum_{n=-N}^{N} e^{in\phi}c_n \tag{3.20d}$$

which is precisely the N^{th} partial sum of the Fourier series.

4 Convergence of Fourier Series, Termwise Integration and Differentiation

4.1 Convergence of Fourier Series

Now let us examine another definition of piecewise continuous [2]:

Definition 4.1. If f is piecewise continuous on (a,b), then f is continuous except "perhaps" at finitely many jumps ("discontinuities"); if f is piecewise continuous on $\mathbb{R}$, then it is piecewise continuous for all $-\infty < a < b < \infty$.

But it is worthy of note that if f is continuous, it has "nicer" Fourier coefficients than if f were merely piecewise continuous. What are "nice" and "mean" Fourier coefficients? Well, how quickly they decay determines how many terms we need to get a good approximation. If the fourier coefficients decay quickly (so $c_n \to 0$ quickly), then the Fourier coefficients are "nicer".

Proposition 4.1. If f is continuous and piecewise smooth on $\mathbb{R}$, then f is continuous and f' is piecewise continuous.

Proof. How can we make this claim? Well recall that we defined piecewise smooth to be when both f and f' are piecewise continuous. But we have the additional property that f is continuous, which is stronger than piecewise continuous. So if f is continuous and piecewise smooth, it is equivalent to saying that f' is piecewise continuous and f is continuous (since continuous is stronger than piecewise continuous, we call it the stronger of the two). All continuous functions are also piecewise continuous, so it is more strict a condition to be continuous. □

Now we should probably review a few definitions.

Definition 4.2. Consider a series of functions

$$\sum_n g_n(x) = g(x).$$

If for each $x \in D \subseteq \mathbb{R}$ (each x in the domain which could be the entire real line), we say the series **converges absolutely** if $\sum |g_n(x)|$ converges.

Definition 4.3. Consider a series of functions

$$\sum_n g_n(x) = g(x).$$

We say that the series **converges uniformly** if for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that

$$|g(x) - \sum_{n=0}^{M} g_n(x)| < \varepsilon, \qquad \forall M \geq N. \tag{4.1}$$

Weierstrass M-test

There is an excellent test to see if a series converges or not. The idea is to take a series that we know converges, but is term for term bigger than the series we are examining. If the series we are testing is term for term smaller than something that converges, then we know the series we are testing converges.

The standard "tricks" to do this is to use bounds for the series. Typically we want to make the numerator (the top part of the fraction) as big as possible, and the denominator as small as possible. So we typically bound $\sin(x) \leq 1$ and $\cos(x) \leq 1$. There are other tricks that perhaps we will see later on.

This test which says "If term for term a series in question is less than a series known to converge, then the series in question converges" is known in the mathematical community as **Weierstrass M-test**. More specifically, if we have a series that is known to converge

$$\sum_n M_n = M \tag{4.2}$$

and if we need to test a given series

$$\sum_n g_n(x)$$

M-test guarantees both uniform *and* absolute *convergence*

and also suppose that term for term $|g_n(x)| \leq M_n$ for all $x \in D$ (where D is the domain of the functions), then

$$\sum g_n(x) \leq M \tag{4.3}$$

converges absolutely and uniformly.

Example 4.1. Recall our Fourier series expansion of the triangle wave

$$f(\theta) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)\theta]}{(2n-1)^2} \tag{4.4}$$

By the Weierstrass M-test, we have

$$\left|\frac{\cos[(2n-1)\theta]}{(2n-1)^2}\right| \leq \frac{1}{(2n-1)^2} \tag{4.5}$$

for all θ. If we are also uncertain, we know

$$\frac{1}{(2n-1)^2} \leq \frac{1}{n^2} \tag{4.6}$$

for all $n \in \mathbb{N}$ and we know that the series

$$\sum_{n=1}\frac{1}{n^2} = \frac{\pi^2}{9}. \tag{4.7}$$

Thus by the Weierstrass M-test, the Fourier series converges uniformly and absolutely.

Theorem 4.1. If f is a 2π-periodic, continuous and piecewise smooth function, then the Fourier series of f converges to f absolutely and uniformly.

The consequence of Theorem (4.1) is that if f is equal to its Fourier series, then

$$\int_a^b f(\theta)d\theta = \sum_n \int_a^b c_n e^{in\theta}d\theta \tag{4.8}$$

for any $-\infty < a < b < \infty$.

Sketch Of Proof. We need to show that

$$\sum_{n=-\infty}^{\infty}|c_n| < \infty, \tag{4.9}$$

then apply the Weierstrass M-test to get absolute and uniform convergence. We need to show three things: 1) Fourier coefficient of f' (denoted c'_n) is $c_n = -ic'_n/n$; 2) the Bessel inequality for f' says $\sum|c'_n|^2 \leq (1/2\pi)\int_{-\pi}^{\pi}|f'(\theta)|^2d\theta$; 3) the Cauchy-Schwarz inequality, which says $\sum|\alpha_n\beta_n| \leq (\sum|\alpha_n|)^{1/2}(\sum|\beta_n|)^{1/2}$. Thus we find

$$\sum|c_n| = \sum_{n=-\infty}^{\infty}\left|\frac{c'_n}{n}\right| + |c_0| \tag{4.10a}$$

$$\leq \left(\sum_{n\neq 0}\frac{1}{n^2}\right)^{1/2}\left(\sum_{n\neq 0}|c'_n|^2\right)^{1/2} + |c_0| \tag{4.10b}$$

$$< +\infty \tag{4.10c}$$

We can plug in the Bessel inequality to find

$$\sum |c_n| \leq \left(\frac{\pi}{\sqrt{6}}\right)\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|f'(\theta)|^2 d\theta\right) + |c_0| \tag{4.11}$$

On that note, we will end our sketch of the proof.

We will need to examine two ideas before we can perform our proof. One is the Fourier coefficients of the derivative of a function. The other is the Cauchy-Schwarz inequality, which proves itself useful time and time again in mathematics.

4.2 Termwise Integration of Fourier Series

We still haven't really answered the question "When does the Fourier series converge, and to what?" We saw for 2π-periodic, continuous and piecewise smooth functions that the Fourier series converges absolutely and uniformly (and in a remarkably handwavy way, anything that is C^1 would converge absolutely and uniformly too). So let us now try to examine all the details of when and what does the Fourier series converge to. We will let

$$F(\theta) = \int_0^\theta f(\phi)d\phi. \tag{4.12}$$

Assumption. Let f be 2π-periodic and piecewise continuous on $\mathbb{R}$.

Now if f has an antiderivative $\widetilde{F}$, then

$$F(\theta) = \widetilde{F}(\theta) - \widetilde{F}(0) \tag{4.13}$$

by the fundamental theorem of calculus. So we can make a number of additional claims

a) F is continuous, i.e. $F(\theta^+) = F(\theta^-)$ for all θ

b) F is piecewise smooth on $\mathbb{R}$, and f is piecewise continuous. We have $F'(\theta) = (d/d\theta)\int_0^\theta f(\phi)d\phi = f(\theta)$ where $f(\theta)$ is continuous.

c) (Assumption for this step: $c_0 = a_0/2 = (1/2\pi)\int_{-\pi}^{\pi} f(\theta)d\theta = 0$) F is 2π-periodic (i.e. $F(\theta+2\pi) - F(\theta) = \int_\theta^{\theta+2\pi} f(\phi)d\phi = \int_{-\pi}^{\pi} f(\phi)d\phi = 0$ by our assumption.

Let the Fourier coefficients of F be C_n, A_n and B_n, and the coefficients for $F' = f$ be c_n, a_n and b_n. So we know

$$c_n = C_n' = inC_n \Rightarrow C_n = \frac{c_n}{in} \tag{4.14}$$

and similarly

$$A_n = -\frac{b_n}{n}, \qquad B_n = \frac{a_n}{n} \tag{4.15}$$

Thus the Fourier series of F is

$$F(\theta) = C_0 + \sum_{n\neq 0} \frac{c_n}{in} e^{in\theta} \tag{4.16a}$$

$$= \frac{1}{2}A_0 + \sum_{n=1} \left(\frac{-b_n}{n}\cos(n\theta) + \frac{a_n}{n}\sin(n\theta) \right) \tag{4.16b}$$

where $C_0 = A_0/2 = (1/2\pi)\int_{-\pi}^{\pi} F(\theta)d\theta$.

For any 2π-periodic , piecewise continuous function f, then it can be expanded into a Fourier Series

$$f(\theta) = c_0 + \sum_{n\neq 0} c_n e^{in\theta} \tag{4.17}$$

where c_0 may be zero. We may define $g(\theta) = f(\theta) - c_0$ which satisfy conditions (a), (b), and (c) outlined above. So

$$\int_0^\theta (f(\phi) - c_0)d\phi = C_0 + \sum_{n\neq 0} C_n e^{in\theta} \tag{4.18a}$$

$$= C_0 + \sum_{n\neq 0} \frac{c_n}{in} e^{in\theta} \tag{4.18b}$$

So $\int_0^\theta (f(\phi) - c_0)d\phi = \int_0^\theta f(\phi)d\phi - c_0\theta = F(\theta) - c_0\theta$ and

$$C_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} (F(\theta) - c_0\theta)d\theta = \frac{1}{2\pi}F(\theta)d\theta \tag{4.19}$$

since $\int_{-a}^{a} \theta d\theta = 0$.

Example 4.2. We will consider making the step function (figure (4.1)) periodic

$$f(\theta) = \begin{cases} 1 & 0 < \theta < \pi \\ -1 & -\pi < \theta < 0. \end{cases} \tag{4.20}$$

We would like to compute the Fourier series of this step function, but observe that

$$\int f(\theta)d\theta = |\theta| \tag{4.21}$$

and we previously calculated the Fourier series of this function! Recall from Eq (2.30) that

$$|\theta| = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty} \frac{\cos((2n-1)\theta)}{(2n-1)^2}. \tag{4.22}$$

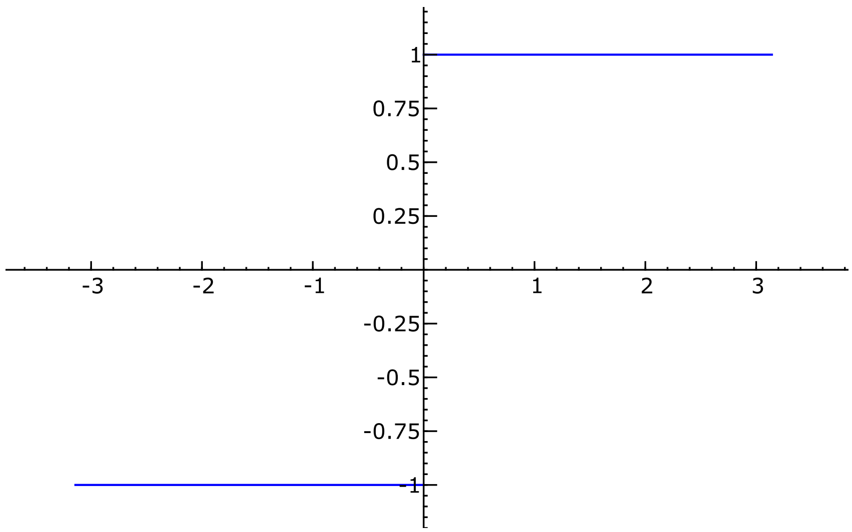

Figure 4.1: Step function that we make periodic.

We take its derivative to find

$$\frac{d}{d\theta}|\theta| = \frac{4}{\pi}\sum_{n=1}\frac{\sin((2n-1)\theta)}{2n-1} \tag{4.23}$$

and we identify it with the Fourier series of the step function

$$f(\theta) = \frac{4}{\pi}\sum_{n=1}\frac{\sin((2n-1)\theta)}{2n-1}. \tag{4.24}$$

5 Smoothness and Decay of Fourier Coefficients, Fourier Expansion on Arbitrary Intervals

Here is a deep concept that we will re-examine in Lie Groups (perhaps if we get there): any periodic function can be considered a function on a circle. (In a hand-wavy way, we can change the period to be from p to 2π by a change of coordinates, then we've got from the Fourier Series a linear combination of functions on the circle $\exp(inx)$ which are all linearly independent.)

Definition 5.1. Let $k \in \mathbb{N}$. A function $f : \mathbb{R} \to \mathbb{R}$ is of class C^k, $f \in C^k$, if f has derivatives up to order k and (in addition to the derivatives existing) they are continuous. When we discuss **smoothness**, we are discussing the existence of derivatives.

C^0 *are continuous functions*

At first it may seem useless to introduce such a definition, but the strength in having such a system is that all C^{k+1} functions are also C^k functions. It provides a clear hierarchy of "nice" functions.But also note, if a function is continuous, then it is necessarily C^0.

But let us reiterate what we have done before we continue. Let f be a 2π periodic, piecewise smooth and continuous. This means $f \in C^0$ and $f' \in PC(\mathbb{R})$. This means that $f(\theta) = \sum c_n e^{in\theta}$ $\forall\theta$. This series *converges absolutely and uniformly*. Also, the derivative of f is equal to the termwise derivatives of its Fourier series

$$f'(\theta) = \sum_{n=-\infty}^{\infty} inc_n e^{in\theta} = \sum n(-a_n \sin(n\theta) + b_n \cos(n\theta)) \tag{5.1}$$

at θ where f' is continuous. If we apply the Bessel Inequality, we have the following inequality on f':

$$\sum |nc_n|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f'(\theta)|^2 d\theta < \infty \tag{5.2}$$

So it follows that

$$nc_n \to 0 \quad \text{as } n \to \infty \tag{5.3}$$

Similarly we can argue that $\sum |na_n|^2$ and $\sum |nb_n|^2$ are finite. So $na_n \to 0$ and $nb_n \to 0$ as $n \to \infty$.

So this implies $a_n, b_n \sim 1/n^k$ where $k > 1$. If $f \in C^1$ and $f'' \in PC^1(\mathbb{R})$, then we can apply the same result again

$$\sum |nc_n|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f''(\theta)|^2 d\theta < \infty \tag{5.4}$$

So $n^2c_n,\ n^2a_n,\ n^2b_n \to 0$ as $n \to \infty$. There is a general pattern which is beginning to emerge here.

Theorem 5.1. Let f be 2π-periodic, if $f \in C^{k-1}$ and $f^{(k)} \in PC(\mathbb{R})$, then

$$\sum |c_n n^k|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f^{(k)}(\theta)|^2 d\theta < \infty \tag{5.5}$$

which implies $n^k c_n \to 0$ as $n \to \infty$. Thus

$$|c_n| \le \frac{(\text{constant})}{n^k + \varepsilon_n}, \qquad \text{where } \varepsilon_n > 0. \tag{5.6}$$

On the other hand, if we compute the Fourier coefficients of a function and $|c_n| \le c/|n|^{k+\alpha}$ where $c > 0$, $\alpha > 0$, we may say that f is C^{k-1}.

5.1 Fourier Series on Arbitrary Interval

Fourier series on an interval $[a, b]$ where $a < b$. How is this done? Well, usually, as a rule of thumb, we don't work all of $\mathbb{R}$ (so $-\infty < a < b < +\infty$). For instance, the heat distribution

$$\begin{aligned} f(x) &= \text{temperature distribution along a rod of length } L \\ & f : [0, L] \to \mathbb{R} \end{aligned}$$

More generally, let $s(t)$ be a signal generated over a time interval $[0, T]$.

The strategy is to first change variables to make the domain from $[0, \pi]$; then we extend our (odd or even) Fourier coefficients to make f defined on $[-\pi, \pi]$. Then we can just make it periodic to extend it to all of $\mathbb{R}$. If one examines any Fourier expansion table, the functions are usually defined on the interval $[-\pi, \pi]$ (or $[0, 2\pi]$). They are periodic extensions of functions defined on $[-\pi, \pi]$ to $\mathbb{R}$.

Definition 5.2. Let $f : [-\pi, \pi) \to \mathbb{C}$, then $f_{ext} : \mathbb{R} \to \mathbb{C}$ is defined by $f_{ext}(\theta + 2\pi) = f(\theta)$ for all $\theta \in [-\pi, \pi]$ is the **periodic extension of f to** $\mathbb{R}$.

CAUTION: f_{ext} may be discontinuous and nondifferentiable at $\theta = (2n - 1)\pi$ where $n \in \mathbb{Z}$. So f_{ext} is continuous at $\theta = (2n - 1)\pi$ if and only if $f(-\pi) = f(\pi^-)$. Similarly, $f^{(k)}_{ext}((2n - 1)\pi)$ exists if and only if $f^{(j)}(\pi^-) = f^{(j)}(-\pi^+)$ where $j = 0, 1, 2, ..., k$.

Example 5.1. Consider $f(\theta) = \theta^2$ where $\theta \in [-\pi, \pi)$. We have $f : [a, b] \to \mathbb{C}$, then we change variable

$$\theta = \pi \left(\frac{x - a}{b - a} \right), \qquad 0 \le \theta \le \pi \tag{5.7}$$

Observe what is going on, we rearrange these terms given arbitrary a and b to find some change of variables from the interval $[a, b]$ to a new interval. It is seen to be

$$g(\theta) = f(x) = f\left(\frac{b - a}{\pi}\theta + a \right), \quad g : [0, \pi] \to \mathbb{C} \tag{5.8}$$

So we can see a fairly clean and orderly scheme of how to change the interval.

Definition 5.3. Let $f : [0, \pi] \to \mathbb{C}$, then the **odd extension** of f to $[-\pi, \pi]$ is

$$f_{odd}(-\theta) = -f(\theta), \quad 0 < \theta < \pi \tag{5.9a}$$

$$f_{odd}(0) = 0 \tag{5.9b}$$

and the **even extension** of f to $[-\pi, \pi]$ is

$$f_{even}(-\theta) = f(\theta), \qquad 0 \le \theta \le \pi \tag{5.10}$$

(note the possible values for θ differ for even and odd extensions).

So what about the Fourier Series of Periodically Extended functions? It's fairly straightforward to find for odd extensions:

$$a_n = \frac{1}{\pi}\int_0^\pi f_{odd}(\theta)\cos(n\theta)d\theta = 0 \tag{5.11a}$$

$$b_n = \frac{1}{\pi}\int_0^\pi f_{odd}(\theta)\sin(n\theta)d\theta \tag{5.11b}$$

$$\Rightarrow \sum_{n=1}^{\infty} b_n \sin(n\theta) = \text{“Fourier Sine Series”} \tag{5.11c}$$

and for even extensions:

$$a_n = \frac{1}{\pi}\int_0^\pi f_{even}(\theta)\cos(n\theta)d\theta \tag{5.12a}$$

$$b_n = \frac{1}{\pi}\int_0^\pi f_{even}(\theta)\sin(n\theta)d\theta = 0 \tag{5.12b}$$

$$\Rightarrow \sum_{n=1}^{\infty} a_n \cos(n\theta) = \text{“Fourier Cosine Series”} \tag{5.12c}$$

6 A Review of Vector Spaces

Remember that the real numbers (denoted as $\mathbb{R}$) and the complex numbers (denoted as $\mathbb{C}$) are usually the scalars we work with in linear algebra (there are other neurotic examples that mathematicians love, but since this is for physicists, we won't examine them). We will generally call this set of scalars, equipped with scalar multiplication and scalar addition, a "**field**". We will denote a generic field of scalars as $\mathbb{F}$.

Now, a **vector space** V is always given as being "over" a field of scalars $\mathbb{F}$. It is equipped with only two operations:

1. Vector Addition, $+ : V \times V \to V$ which is denoted as $\mathbf{v}+\mathbf{w}$, where $\mathbf{v}, \mathbf{w} \in V$,
2. Scalar Multiplication, $* : \mathbb{F} \times V \to V$, denoted usually by $\alpha\mathbf{v}$ for all $\alpha \in \mathbb{F}$ and $\mathbf{v} \in V$

It *DOES NOT* have a cross product, or a dot product, or any other sort of fun stuff. These *ARE NOT* properties of a vector space. The *ONLY* properties are scalar multiplication and vector addition.

We need to specify several properties of vector addition and scalar multiplication. There are four axioms of vector addition

1. Vector addition is associative: for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, we have $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$;

2. Vector Addition is commutative: for all $\mathbf{v}, \mathbf{w} \in V$, we have $\mathbf{v}+\mathbf{w} = \mathbf{w}+\mathbf{v}$;

3. Vector addition has an identity element: there exists an element $\mathbf{0} \in V$ (dubbed the **zero vector**) such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$;

4. Vector Addition has inverse elements: For all $\mathbf{v} \in V$ there exists an element $\mathbf{w} \in V$, called the **additive inverse** of $\mathbf{v}$, such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$.

Now, it is important to note that the zero vector is unique. Observe a short proof: suppose we have two zero vectors $\mathbf{0}$ and $\widetilde{\mathbf{0}}$, then for some $\mathbf{v}$ and its additive inverse $\mathbf{w}$ we have

$$\mathbf{v} + \mathbf{w} = \mathbf{0} \tag{6.1a}$$

$$= \widetilde{\mathbf{0}} \tag{6.1b}$$

$$\Rightarrow \mathbf{0} = \widetilde{\mathbf{0}} \tag{6.1c}$$

Which implies the uniqueness of the zero vector.

There are similarly four axioms of scalar multiplication:

1. Distributivity holds forscalar multiplication over vector addition: $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$ for all $\alpha \in \mathbb{F}$, $\mathbf{v}, \mathbf{w} \in V$;

2. Distributivity holds for scalar multiplication over field addition (ie. adding two scalars and multiplying by their sum): $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{w}$ for all $\alpha, \beta \in \mathbb{F}$ and $\mathbf{v} \in V$.

3. Scalar multipication is compatible with multipocation in the sense of the field of scalars: $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$ for all $\alpha, \beta \in \mathbb{F}$ and $\mathbf{v} \in V$.

4. Scalar multiplication has an identity element: $\mathbf{1v} = \mathbf{v}$ for all $\mathbf{v} \in V$ and the multiplicative identity of $\mathbb{F}$: $\mathbf{1}$.

6.1 Inner Products

Now we wish to speak of certain "extra-structure" on top of a vector space. What is this "extra-structure" of which I speak? Well, things like the norm, the inner product, these sort of "extra goodies" are the "extra-structure" that I am going to discuss. Let us first consider the table of similarities between the k dimensional vector space with complex entries (i.e. a vector space over the field $\mathbb{C}$ with k dimensions denoted as $\mathbb{C}^k$ which consists of an ordered "k-tuple" consisting of k components, with vector addition and scalar multiplication defined componentwise) with the piecewise continuous functions on (a, b) (which we denote as $PC(a, b)$).

$\mathbb{C}^k$	$PC(a,b)$
k dimensional	infinite dimensional
$\vec{a}=(a_1,...,a_k)$ with $a_j\in\mathbb{C}$	elements are functions that are piecewise continuous on $[a,b]$.
Addition is done componentwise $\vec{a}+\vec{b}=(a_1+b_1,...,a_n+b_n)$	The "components" are continuous, so we do addition "point-wise" $f(x)+g(x)$
Scalar multiplication done componentwise $\alpha\vec{a}=(\alpha a_1,...,\alpha a_k)$	Scalar multiplication is done in the naive way by merely multiplying by a constant $\alpha f(x)$

Now we would like to consider the various "extra structure" that we can add to a vector space, and whether this can be generalized in the $PC(a,b)$ vector space.

$\mathbb{C}^k$	$PC(a,b)$
We sum up the entries componentwise for the inner product[2] Mathematicians use the notation $\langle a,b\rangle=\sum_j^k a_j\bar{b}_j$ but physicists use Bra-Ket notation $\langle b\vert a\rangle = \langle a,b\rangle$ is the translation between the two notations	We similarly sum up all the points for two functions $\langle f,g\rangle=\int_a^b f(x)\bar{g(x)}dx$
We have $\langle\vec{0},\vec{u}\rangle=0$ for all $\vec{u}\in\mathbb{C}^k$	$\langle 0,f\rangle=0$ for all $f\in PC(a,b)$.

Observe the consequences of defining the inner product for $PC(a,b)$ to be

$$\langle f,g\rangle=\int_a^b f(x)\overline{g(x)}dx \tag{6.2}$$

we have linearity in the first slot

$$\langle\alpha f+\beta g,h\rangle=\alpha\langle f,h\rangle+\beta\langle g,h\rangle \tag{6.3}$$

Linearity in first slot

for arbitrary $\alpha,\beta\in\mathbb{C}$. We can see this by going back to the definition of the inner product and plugging it in

$$\begin{aligned}
\langle\alpha f+\beta g,h\rangle &= \int_a^b(\alpha f(x)+\beta g(x))\overline{h(x)}dx && (6.4a)\\
&= \int_a^b\left(\alpha f(x)\overline{h(x)}+\beta g(x)\overline{h(x)}\right)dx && (6.4b)\\
&= \int_a^b \alpha f(x)\overline{h(x)}dx+\int_a^b\beta g(x)\overline{h(x)}dx && (6.4c)\\
&= \alpha\int_a^b f(x)\overline{h(x)}dx+\beta\int_a^b g(x)\overline{h(x)}dx && (6.4d)\\
&= \alpha\langle f,h\rangle+\beta\langle g,h\rangle. && (6.4e)
\end{aligned}$$

All we used was the distributivity of multiplication, the linearity of the integral (twice actually), and then we concluded with using the definition of the inner product on $PC(a, b)$.

There is also a similar property for the second slot, but it is called **antilinear** This sounds bizarre, but using the same song and dance we see that

$$\langle f, \alpha g\rangle = \bar{\alpha}\langle f, g\rangle \tag{6.5}$$

where $\bar{\alpha}$ is the complex conjugate of α. However, we see that

$$\begin{aligned}
\langle f, g+h\rangle &= \int_a^b f(x)(\overline{g(x)} + \overline{h(x)})dx && \text{(6.6a)}\\
&= \int_a^b (f(x)\overline{g(x)} + f(x)\overline{h(x)})dx && \text{(6.6b)}\\
&= \int_a^b f(x)\overline{g(x)}dx + \int_a^b f(x)\overline{h(x)}dx && \text{(6.6c)}\\
&= \langle f, g\rangle + \langle f, h\rangle. && \text{(6.6d)}
\end{aligned}$$

Thus antilinearity in the second slot can be summed up as

Antilinearity in second slot

$$\langle f, \alpha g + \beta h\rangle = \bar{\alpha}\langle f, g\rangle + \bar{\beta}\langle f, h\rangle \tag{6.7}$$

for $\alpha, \beta \in \mathbb{C}$.

The last property of the inner product of interest is its Hermitian Symmetry. That is to say that we have

$$\langle f, g\rangle = \langle g, f\rangle^* \tag{6.8}$$

which is shorthand for

$$\langle g, f\rangle^* = \overline{\langle g, f\rangle}. \tag{6.9}$$

For finite dimensional vector spaces, we need to also take the transpose. One of the peculiarities of the infinite dimensional vector spaces like $PC(a, b)$ is that we don't need (or even have well defined) a transpose operation.

6.2 Induced Norm

Let us review the case of the properties of the norm for finite dimensional vector spaces V over $\mathbb{C}$. Once we have an inner product, we can define (invent, or "induce") a norm

$$\|n\| = \sqrt{\langle n, n\rangle}. \tag{6.10}$$

So in $\mathbb{C}^k$, this would be

$$\|\vec{a}\| = \sqrt{\sum_{j=1}^{k} |a_j|^2} \tag{6.11}$$

but in $PC(a,b)$ we have

$$\|f\| = \left(\int_a^b |f(x)|^2 dx \right)^{1/2} \tag{6.12}$$

So the question arises: what are the properties of the norm?

The norm has the property of homogeneity which means

$$\|\alpha \vec{u}\| = |\alpha| \|\vec{u}\| \tag{6.13}$$

for all $\alpha \in \mathbb{C}$ and $\vec{u} \in V$.

Additionally, there is the property of positivity

$$\|\vec{a}\| = 0 \iff \vec{a} = \vec{0}, \quad \text{and } \|\vec{a}\| > 0 \quad \forall \vec{a} \neq \vec{0} \tag{6.14}$$

or in other words, the norm of a vector is zero if and only if it is the zero vector. Otherwise, the norm of a vector is positive.

Now, the condition of positivity is not strictly true for the vector space $PC(a,b)$. Consider, working on $PC(0,1)$, the function

$$f(x) = \begin{cases} x & \text{if } x = (1/4), (1/2), (3/4) \\ 0 & \text{otherwise} \end{cases} \tag{6.15}$$

Then we see that

$$\|f\|^2 = \int_0^{1/4} (0) dx + \int_{1/4}^{1/2} (0) dx + \int_{1/2}^{3/4} (0) dx + \int_{3/4}^{1} (0) dx \tag{6.16a}$$

$$= 0 \tag{6.16b}$$

but $f(x) \neq 0$.

The convention we use is that in $PC(a,b)$ two functions are considered to be equal $f(x) = g(x)$ if they agree except at finitely many points. The motivation for this is, in finite dimensional vector space, if

$$\vec{u} = \vec{v} \tag{6.17}$$

then

$$\vec{u} - \vec{v} = \vec{0} \Rightarrow \|\vec{u} - \vec{v}\| = 0. \tag{6.18}$$

So, this would mean that the two are equal. If there are only a finite number of points where $f(x)$ disagrees with $g(x)$, then the norm of the difference would vanish. So by the above scratch work, it would imply that $f(x) = g(x)$.

Lemma 6.1. For any $u, v \in V$ (of arbitrary dimension), then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2\,\mathrm{Re}\left(\langle u, v \rangle\right). \tag{6.19}$$

Proof. It trivially follows from the Hermitian symmetry of the inner product (that is no proof, it should be noted). Observe that

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle && (6.20a)\\ &= \langle u,v\rangle + \langle v,u\rangle + \langle u,u\rangle + \langle v,v\rangle && (6.20b)\\ &= \mathrm{Re}\,(\langle u,v\rangle) + \|u\|^2 + \|v\|^2 && (6.20c)\end{aligned}$$

which is an explicit calculation. □

Cauchy-Schwarz Inequality

Lemma 6.2. (Cauchy-Schwarz Inequality) Let $u, v \in V$ then

$$|\langle u,v\rangle| \leq \|u\|\|v\| \tag{6.21}$$

and it is an equality if and only if $u = \alpha v$ where $\alpha \in \mathbb{C}$.

Proof. For any real constant $t \in \mathbb{R}$,

$$\begin{aligned}0 &\leq \|u+tv\|^2 && (6.22a)\\ &\leq \|u\|^2 + |t|^2\|v\|^2 + 2t\,\mathrm{Re}\,(\langle u,v\rangle) && (6.22b)\end{aligned}$$

We can assume without loss of generality that $\|v\| \neq 0$ (otherwise the inequality holds trivially, a sixth grader could show it). We can form a quadratic equation in t

$$0 \leq q(t) = \frac{\|u\|^2}{\|v\|^2} + 2\frac{\mathrm{Re}(\langle u,v\rangle)}{\|v\|^2}t + t^2 \tag{6.23}$$

We see that $q(t)$ is quadratic, so it has a minimum somewhere (since its coefficients are all positive). We take its derivative and set it to zero:

$$q'(t) = 2\left(\frac{\mathrm{Re}(\langle u,v\rangle)}{\|v\|^2} + t\right) \tag{6.24}$$

We see when setting this to be zero that the minimum occurs when

$$t = -\frac{\mathrm{Re}(\langle u,v\rangle)}{\|v\|^2}. \tag{6.25}$$

We plug this value of t back into $q(t)$ to find

$$0 \leq q\left(-\frac{\mathrm{Re}(\langle u,v\rangle)}{\|v\|^2}\right) = \|u\|^2 - \frac{[\mathrm{Re}(\langle u,v\rangle)]^2}{\|v\|^2} \tag{6.26}$$

Without loss of generality, we may assume that $\langle u,v\rangle \in \mathbb{R}$. Suppose it's really complex, then we can write it in polar form as

$$\langle u,v\rangle = re^{i\theta} \tag{6.27}$$

for $r, \theta \in \mathbb{R}$ such that $r > 0$ and $0 \leq \theta < 2\pi$. Then we find that

$$|\langle u, v\rangle| = r \tag{6.28}$$

which is basic complex analysis. Thus plugging this into Eq (6.26) we find that

$$0 \leq \|u\|^2\|v\|^2 - |\langle u, v\rangle|^2 \tag{6.29}$$

which implies

$$|\langle u, v\rangle|^2 \leq \|u\|^2\|v\|^2 \tag{6.30}$$

which is precisely the inequality! □

Lemma 6.3. (Triangle Inequality) Let $u, v \in V$, then

$$\|u\| + \|v\| \geq \|u + v\| \tag{6.31}$$

and it is equal if $u = tv$ where $t \in \mathbb{R}$.

Definition 6.1. Two vectors $u, v \in V$ are said to be **orthogonal** if $\langle u, v\rangle = 0$.

Lemma 6.4. (Pythagorean Theorem) Let $u_1, \ldots, u_n$ be mutually orthogonal vectors, then

$$\|u_1 + \cdots + u_n\|^2 = \|u_1\|^2 + \cdots + \|u_n\|^2. \tag{6.32}$$

Now, we can think of the Fourier series of $f \in PC(a, b)$ as a sort of expansion with respect to a given basis.

We know that $\{e^{in\theta}\}$ forms an orthogonal basis for all $n \in \mathbb{Z}$. Observe that

$$\langle \exp(in\theta), \exp(in\theta)\rangle = \int_{-\pi}^{\pi} e^{in\theta}e^{-in\theta}d\theta = 2\pi \tag{6.33}$$

and

$$\langle \exp(i(n+m)\theta), \exp(in\theta)\rangle = \int_{-\pi}^{\pi} e^{im\theta}d\theta = \frac{1}{im}\left(e^{im\pi} - e^{-im\pi}\right) = 0. \tag{6.34}$$

So to form an orthonormal basis, we need to use the basis $\{\exp(in\theta)/\sqrt{2\pi}\}$ for all $n \in \mathbb{Z}$.

So wait, why do we care so much about a basis? A vector is a vector is a vector, right? Well, no, it's time to review why basis vectors matter.

6.3 Basis Vectors Matter

In general, let V be a vector space over $\mathbb{C}$. Given n vectors $v_1, ..., v_n \in V$, we have this set of all possible linear combinations of these vectors called the **linear space**

$$\operatorname{span}(v_1, ..., v_n) = \{a_1v_1 + \cdots + a_nv_n : \forall a_1, ..., a_n \in \mathbb{C}\}. \tag{6.35}$$

Now, we can call a given collection of vectors $v_1, ..., v_n$ **linearly independent** if the only scalars $a_1, ..., a_n \in \mathbb{C}$ that satisfy

$$a_1 v_1 + \cdots + a_n v_n = 0 \tag{6.36}$$

are $a_1 = \cdots = a_n = 0$. This means that we cannot write any single vector in terms of a linear combination of the other $n-1$ vectors.

It would be lovely if we could find some collection of linearly independent vectors whose span is equal to the entire vector space V, wouldn't it? (Yes, it would.) There is a special name we give to such collection of vectors, we call the set a **basis** and the vectors are called (appropriately enough) **basis vectors**.

There is also a way to write any given vector as a linear combination of a given basis. What we do is "project" the given vector x onto the basis vector v_j, which is some scalar $\langle x, v_j\rangle$ and then multiply the basis vector by this quantity. So we end up with a sum

$$x = \sum_j \langle x, v_j\rangle v_j. \tag{6.37}$$

Let us consider a few examples.

Example 6.1. Consider the vector space $\mathbb{R}^2$ over the scalar field $\mathbb{R}$ with the usual dot produt. Let $e_1 = (1,1)/\sqrt{2}$, $e_2 = (1,-1)/\sqrt{2}$ be an orthonormal basis (we see this because the dot product of a given basis vector with itself is 1, but dot them to each other results in 0). Let $x = (3,7)$. Then we can write it as a linear combination of the basis

$$x = \langle x, e_1\rangle e_1 + \langle x, e_2\rangle e_2 \tag{6.38}$$

Observe

$$\langle x, e_1\rangle = (3+7)/\sqrt{2} = 5\sqrt{2}, \qquad \text{and } \langle x, e_2\rangle = (3-7)/\sqrt{2} = -2\sqrt{2} \tag{6.39}$$

thus

$$x = (5\sqrt{2})e_1 + (-2\sqrt{2})e_2 = (5e_1 - 2e_2)\sqrt{2} \tag{6.40}$$

which concludes this example. If one doesn't believe it, just plug it all back in

$$(5e_1 - 2e_2)\sqrt{2} = (5,5) + (-2,2) = (3,7). \tag{6.41}$$

Woah, that's exactly x!

Remark 6.1. If we have a basis that is not composed of unit vectors, we have to make them unit vectors *BEFORE* trying to expand a given vector in terms of them.

Remark 6.2. The Fourier series is just an expansion in the basis $\{\exp(in\theta)\}$, and it is normalized silently by making the cofficients be

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-in\theta}d\theta \tag{6.42}$$

The basis is normalized as $\exp(in\theta)/\sqrt{2\pi}$ but then when we take the dot product of the new coefficient

$$\widetilde{c}_n = \int_{-\pi}^{\pi} f(\theta)\frac{e^{-in\theta}}{\sqrt{2\pi}}d\theta \tag{6.43}$$

and there is an extra factor of $1/\sqrt{2\pi}$ in the basis vectors in the expansion, which yields the series being

$$f(\theta) = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi}e^{in\theta}\int_{-\pi}^{\pi} f(\theta)e^{-in\theta}d\theta \tag{6.44}$$

which is precisely what we have with our original Fourier coefficients and our original basis.

7 L^2 spaces

Before beginning just a few definitions on the various types of convergences.

Definition 7.1. Suppose $\{f_n\}$ is a sequence of functions in $L^2(a,b)$.

1. We say that f_n converges to f **pointwise** if for each $x \in [a,b]$,

$$f_n(x) \to f(x) \text{ as } n \to \infty \tag{7.1}$$

(.e. $|f_n(x) - f(x)| \to 0$ as $n \to \infty$). This means for any $\varepsilon > 0$ and any $x \in [a,b]$ there is an integers $N_{\varepsilon,x} > 0$ which depends on ε and x such that

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for all } n > N_{\varepsilon,x} \tag{7.2}$$

2. We say that f_n converges to f **uniformly** if

$$\sup_{a \le x \le b} |f_n(x) - f(x)| \to 0 \text{ as } n \to \infty \tag{7.3}$$

which means for each $\varepsilon > 0$ there is an integer $N > 0$ such that

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for all } n > N \tag{7.4}$$

which is a stronger condition than pointwise convergence since N depends only on the choice of ε.

3. We say that f_n converges to f **in norm** if

$$\|f_n - f\| \to 0 \quad \text{as } n \to \infty \tag{7.5}$$

where $\|\cdot\|$ is the norm on the vector space $L^2(a,b)$. By the definition of the norm this means

$$\int_a^b |f_n(x) - f(x)|^2 dx \to 0 \quad \text{as } n \to \infty \tag{7.6}$$

Motivation: We know that

$$\langle \frac{1}{\sqrt{2\pi}} e^{imx}, \frac{1}{\sqrt{2\pi}} e^{inx} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)x} dx$$
$$= \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

So the set $\{\exp(inx)/\sqrt{2\pi}\}_{-\infty}^{\infty}$ is an orthonormal set. We would like to make it an orthonormal basis.

Recall in $\mathbb{C}^k$: any set $\{u_1, \dots, u_k\}$ that is orthonormal forms a basis for $\mathbb{C}^k$. It has two properties of relevance

1. For any $v \in \mathbb{C}^k$, $v = \sum_j \langle v, u_j \rangle u_k$. In other words, any vector can be written as a linear combination of the vectors u_j, $j = 1, \dots, k$.
2. the sum $\sum_j \alpha_j u_j$ (for all $\alpha_j \in \mathbb{C}$) forms some vector in $\mathbb{C}^k$.

We want to mirror this for our $L^2(a, b)$ space; however, the second property is not true for any orthonormal set because $PC(a, b)$ is not "complete".

So we move to a bigger space, L^2, which is defined as the set of all functions satisfying

$$L^2(a, b) = \{f \text{ on } [a, b] \text{ s.t. } \int_a^b |f(x)|^2 dx < \infty\} \tag{7.7}$$

where we use the Lebesgue integral. What the hell's the Lebesgue integral, and why do we use it? The simplest explanation is that it's the area between the x-axis and the curve (one may ask "But isn't that a Riemann integral?" And that's kind of true, but one typically has vertical rectangles in the Riemann sum, whereas one would have e.g. vertical rectangles in the Lebesgue integral).

Don't Stress over the Lebesgue integral, it really is nothing special...just integration done in the way taught in math 21 B, but the only time we run into problems is with wacky functions and bizarre domains (e.g. domains with only a single point).

If $f \in PC(a, b)$ then $\int_a^b |f(x)|^2 dx < \infty$...so $f \in L^2(a, b)$. So this implies $PC(a, b) \subset L^2(a, b)$. Improper Riemann integrable functions are also in $L^2(a, b)$. Lebesgue integration is, to a physics undergrad, just "really clever integration" (or integration done in the usual way being careful about infinities).

Example 7.1. Let $f(x) = x^{-1/3}$ on $[-1, 1]$. This is not piecewise continuous on $[-1, 1]$ since it blows up around 0. But observe

$$\begin{aligned} \int_{-1}^{1} |f(x)|^2 dx &= \int_{-1}^{1} x^{-2/3} dx \\ &= \lim_{\alpha \to 0^-} \int_{-1}^{\alpha} x^{-2/3} dx + \lim_{\beta \to 0^+} \int_{\beta}^{1} x^{-2/3} dx \\ &= \lim_{\alpha \to 0^-} (3\alpha^{1/3} + 3) + \lim_{\beta \to 0^+} (3 - 3\beta^{1/3}) \\ &= 6 \end{aligned}$$

So $f \in L^2(-1, 1)$.

Definition 7.2. Let $E \subset \mathbb{R}$, E is **(Lebesgue) measure zero** if for any $\varepsilon > 0$, there exists a set of intervals $\{I_1, \ldots\}$ of length $\ell_1, \ldots$ such that

$$E \subset \cup_{j=1} I_j \quad \text{and} \quad \sum_{j=1}^{\infty} \ell_j < \infty \tag{7.8}$$

Let $\Delta > 0$, then $\int_0^\Delta 1dx \neq 0$. For E of measure zero, $\int_E 1dx = 0$.

Remark 7.1. Some functions with no Riemann integral have Lebesgue integral.

Example 7.2. Let $\{r_1, \ldots\}$ be enumeration of the rational numbers, then it has Lebesgue measure zero. Let

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is irrational} \\ 0 & \text{otherwise} \end{cases} \tag{7.9}$$

be defined on $[0,1]$. Then $f = 0$ on a set of measure zero. THe Riemann integral of f does not exist, the upper sum of f is

$$\sum 1\Delta x \tag{7.10}$$

where Δx is the width of the intervals, and the lower sums are

$$\sum 0\Delta x = 0 \tag{7.11}$$

Thus

$$\lim_{\Delta x \to 0} \sum 1\Delta x = \sum 0\Delta x \Rightarrow 1 = 0 \tag{7.12}$$

which is known to be wrong! However, for the Lebesgue integral

$$\int_0^1 |f(x)|^2 dx = 1 \Rightarrow f(x) \in L^2(0,1). \tag{7.13}$$

Question: *Can we extend the definition of the inner product* $\langle \cdot, \cdot \rangle$ *from* $PC(a,b)$ *to* $L^2(a,b)$*?"*

Well, we very generally defined the inner product to be

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}dx \tag{7.14}$$

If $\int f(x)\overline{g(x)}dx$ exists, for all $f, g \in L^2(a,b)$, then $\langle \cdot, \cdot \rangle$ is well defined.

Let f, g be two functions in $L^2(a,b)$. Then

$$|\langle f, g \rangle| \leq \int_a^b |f(x)\overline{g(x)}|dx \tag{7.15}$$

So $|f(x)\overline{g(x)}| \leq |f(x)||g(x)|$ we know for $s, t \in \mathbb{R}$

$$0 \leq (s-t)^2 = s^2 + t^2 - 2st$$

if and only if

$$\frac{1}{2}(s^2+t^2) \geq st \quad \forall s,t \in \mathbb{R}$$

Since $|f(x)|$ and $|g(x)|$ are both real (even if $f(x)$ and $g(x)$ are complex, their absolute value (aka their modulus) returns the "radial length" which is a real number), we see that

$$\frac{1}{2}\left(|f(x)|^2+|g(x)|^2\right) \geq |f(x)||g(x)| \tag{7.16}$$

so

$$|\langle f,g\rangle| \leq \frac{1}{2}\int |f(x)|^2+|g(x)|^2 dx < \infty. \tag{7.17}$$

Thus we may extend the inner product $\langle\cdot,\cdot\rangle$ from $PC(a,b)$ to $L^2(a,b)$. All properties of the inner product and the norm still hold.

Three is one more property that also holds that we have been trying to prove.

Bessel Inequality. If $\{\phi_n\}_{n=1}^{\infty}$ is an orthonormal set in $L^2(a,b)$, and if $f \in L^2(a,b)$, then

$$\sum_{n=1}^{\infty} |\langle f,\phi_n\rangle|^2 \leq \|f\|^2 \qquad \text{in } L^2(a,b) \tag{7.18}$$

Proof. Let N be any positive integer,

$$\begin{aligned}
0 &\leq \|f-\sum_{n=1}^{N}\langle f,\phi_n\rangle\phi_n\|^2 \\
&= \|f\|^2 - 2\operatorname{Re}(\langle f, \sum_{1}^{N}\langle f,\phi_n\rangle\phi_n\rangle + \|\sum_{n=1}^{N}\langle f,\phi_n\rangle\phi_n\|^2
\end{aligned}$$

We can expand out

$$\begin{aligned}
\langle f, \sum_{n=1}^{N}\langle f,\phi_n\rangle\phi_n\rangle &= \sum_{n=1}^{N}\langle f, \langle f,\phi_n\rangle\phi_n\rangle \\
&= \sum_{n=1}^{N}\overline{\langle f,\phi_n\rangle}\langle f,\phi_n\rangle \\
&= \sum_{n=1}^{N}|\langle f,\phi_n\rangle|^2
\end{aligned}$$

We put all this together to get

$$\begin{aligned}
0 &\leq \|f\| - \sum_{n=1}^{N}\langle f,\phi_n\rangle\phi_n\|^2 \\
&\leq \|f\|^2 - \sum_{n=1}^{N}|\langle f,\phi_n\rangle|^2
\end{aligned}$$

for all N. $\square$

Some intuition of distance

Now, for something completely different. Recall back in one real dimension, we measure distance by taking the absolute value of the difference between two numbers. We generalize this in multiple dimensions, from the Pythagorean theorem, to be the norm of the difference between two vectors. We generalize this still to be the norm of the difference between two functions for having some intuition of distance in L^2 (i.e. for $f, g \in L^2(a,b)$, $\|f-g\|$ measures the "distance" between f and g), we also say that f_n converges to $f \in L^2(a,b)$ **in norm** if

$$\|f_n - f\| \to 0 \quad \text{as } n \to \infty. \tag{7.19}$$

8 Orthonormal Basis for L^2

Definition 8.1. Let $f_n, f \in L^2(a,b)$, $f_n \to f$ **converges in norm** $\|f_n - f\| \to 0$ as $n \to \infty$ if and only if $\int_a^b |f_n(x) - f(x)|^2 dx \to 0$ as $n \to \infty$.

Remark 8.1. If one has the norm of two functions $\|f - g\| = 0$, then $f = g$ in $L^2(a,b)$ if and only if $f(x) = g(x)$ for x (outside of sets of measure zero) in $[a,b]$.

Definition 8.2. We say $f = g$ **almost everywhere** if $f(x) = g(x)$ for $x \in [a,b] - E$ where E is measure 0.

Example 8.1. Let

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1] \text{ is irrational} \\ 1 & \text{if } x \in \mathbb{Q} \cap [0,1] \end{cases} \tag{8.1}$$

then $f(x) = 1$ almost everywhere.

Remark 8.2. The pointwise convergence and convergence in norm do not imply each other.

Theorem 8.1. If $f_n \to f$ uniformly on an interval $[a,b]$, then $f_n \to f$ in norm.

Proof. Let $M_n = \sup_{x \in [a,b]} |f_n(x) - f(x)| \to 0$ as $n \to \infty$. Now $\|f_n - f\|^2 = \int_a^b |f_n(x) - f(x)|^2 dx \leq \int_a^b M_n^2 dx = M_n^2 (b-a)$. But $M_n \to 0$ so $M_n^2(b-a) \to 0$ too. $\square$

Definition 8.3. A sequence $\{a_n\}$ in a normed vector space V (i.e. a vector space with a norm $\|\cdot\|$) is called a **Cauchy Sequence** if

$$\|a_m - a_n\| \to 0 \quad \text{as } m, n \to \infty \tag{8.2}$$

V is **complete** if every Cauchy sequence converges to a vector in V.

Example 8.2. In $\mathbb{R}$, $a_n = 1/n$, then

$$\begin{aligned}\|a_m - a_n\| &= \left|\frac{1}{m} - \frac{1}{n}\right| \\ &= \left|\frac{n-m}{mn}\right| \to 0\end{aligned}$$

as $m, n \to \infty$. So a_n is Cauchy and goes to zero.

One can observe that $PC(a,b)$ is not complete by this counter-example to the claim. Consider $PC(0,1)$ and let

$PC(a,b)$ is not complete

$$f_n(x) = \begin{cases} x^{-1/4} & \text{if } x > 1/n \\ 0 & \text{otherwise} \end{cases} \tag{8.3}$$

If $m > n$, $f_m(x) - f_n(x)$ equals $x^{-1/4}$ when $m^{-1} < x \leq n^{-1}$ and equals 0 otherwise, so

$$\|f_m - f_n\|^2 = \int_{1/m}^{1/n} x^{-1/2} dx = 2x^{1/2}|_{1/m}^{1/n} = 2(n^{-1/2} - m^{-1/2}), \tag{8.4}$$

which tends to zero as $m, n \to \infty$. Thus the sequence $\{f_n\}$ is Cauchy; but clearly its limit (either pointwise or in norm) is

$$f(x) = \begin{cases} x^{-1/4} & \text{if } x > 0 \\ 0 \text{if } x = 0 \end{cases} \tag{8.5}$$

and this function does not belong to $PC(0,1)$ because it becomes unbounded as $x \to 0$.

However, it is worthy of note that $L^2(a,b)$ is not only complete but the completion of $PC(a,b)$.

Rudin's notion of completeness [3]

Theorem 8.2. (Rudin 11.42) If $\{f_n\} \in L^2(a,b)$ (n=1,2,...) is a Cauchy sequence, then there exists some function $f \in L^2(a,b)$ such that $\{f_n\}$ converges to $f \in L^2(a,b)$.

This says,in other words, that $L^2(a,b)$ is a *complete* metric space.

Theorem 8.3. (Rudin's Theorem 11.38) For all $f \in L^2(a,b)$, and for each $\varepsilon > 0$, there is a $(b-a)$-periodic function and infinitely smooth $\widetilde{f} \in C^\infty(a,b)$ such that $\|f - \widetilde{f}\| < \varepsilon$.

Theorem 8.4. If $\{\phi_n\}$ is an orthonormal set in $L^2(a,b)$, $f \in L^2(a,b)$ the Bessel inequality states:

$$\sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2 \leq \|f\|^2. \tag{8.6}$$

Thus

$$\sum_{N}^{M} |\langle f, \phi_n \rangle|^2 \to 0 \quad \text{as } M, N \to \infty. \tag{8.7}$$

Further

$$\| \sum_{1}^{N} \langle f, \phi_n \rangle \phi_n \|^2 = \sum_{n=1}^{N} |\langle f, \phi_n \rangle|^2 \tag{8.8}$$

by the Pythagorean theorem. It suffices to show that $\langle f, \phi_n \rangle \phi_n$ is a Cauchys series

$$\| \sum_{M}^{N} \langle f, \phi_n \rangle \phi_n \|^2 \to 0. \tag{8.9}$$

Lemma 8.1. If $f \in L^2(a,b)$, $\{\phi_n\}$ is an orthonormal set in $L^2(a,b)$, then

$$\sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n \tag{8.10}$$

converges in norm; moreover

$$\| \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n \| \leq \|f\|. \tag{8.11}$$

8.1 Criteria for an Orthonormal Basis in $L^2(a,b)$

Theorem 8.5. Let $\{\phi_n\}_{n=1}^{\infty}$ be an orthonormal set in $L^2(a,b)$, the following conditions are equivalent

1. if $\langle f, \phi_n \rangle = 0$ for all $n \in \mathbb{Z}$, then $f = 0$
2. for all $f \in L^2(a,b)$, $f = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n$ (convergence in norm)
3. for all $f \in L^2(a,b)$, we have the **Parseval equality**

$$\|f\|^2 = \sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2. \tag{8.12}$$

Proof. We'll prove that (1) implies (2), (2) implies (3), and (3) implies (1).

(1)⇒(2) Suppose $\langle f, \phi_n \rangle = 0$ for all n. Let

$$\begin{aligned}
g &= f - \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n \\
\langle g, \phi_m \rangle &= \langle f, \phi_m \rangle - \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \langle \phi_n, \phi_m \rangle \\
&= \langle f, \phi_m \rangle - \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \delta_{nm} \\
&= \langle f, \phi_m \rangle - \langle f, \phi_m \rangle \\
&= 0
\end{aligned}$$

for any m. Therefore $g = 0$.

(2)⇒(3) Suppose for any f, we have

$$f = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n \tag{8.13}$$

So

$$\begin{aligned}
\|f\|^2 = \langle f, f \rangle &= \|\sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n\|^2 \\
&= \lim_{N\to\infty} \|\sum_{n=1}^{N} \langle f, \phi_n \rangle \phi_n\|^2 \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} |\langle f, \phi_n \rangle|^2 \quad \text{by Pythagorean thm} \\
&= \sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2.
\end{aligned}$$

(3)⇒(1) Suppose

$$\|f\|^2 = \sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2 \tag{8.14}$$

Suppose $\langle f, \phi_n \rangle = 0$ for all n. Then $\|f\|^2 = 0$ which implies $f = 0$ in $L^2(a, b)$. ☐

Remark 8.3. In $\mathbb{C}^k$ where $\vec{v} = (v_1, ..., v_k)$, then

$$\|\vec{v}\|^2 = \sum_{j=1}^{k} |\langle v, e_j \rangle|^2 \tag{8.15}$$

is the discrete version of Parseval equality.

9 More on Orthonormal Basis, Spectral Theorem

Theorem 9.1. The set $\{\exp(inx)/\sqrt{2\pi}\}_{-\infty}^{\infty}$ is an orthonormal basis in $L^2(-\pi,\pi)$. (Equivalently,

$$\left\{\frac{1}{\sqrt{2\pi}}\right\} \cup \left\{\frac{\cos(nx)}{\sqrt{\pi}}\right\}_1^{\infty} \cup \left\{\frac{\sin(nx)}{\sqrt{\pi}}\right\}_1^{\infty}$$

is an orthonormal basis in $L^2(-\pi,\pi)$.)

Proof. Fix an arbitrary $f \in L^2(a,b)$. Let $\psi_n(x) = \exp(inx)/\sqrt{2\pi}$. $\alpha_n = \langle f, \psi_n \rangle$. We want to show that

$$\|f - \sum_{n=-N}^{N} \alpha_n \psi_n\| \to 0 \text{ as } n \to \infty$$

(Note that $\sum \alpha_n \psi_n$ is the Fourier series of f). Let $\varepsilon > 0$ be given. We will show that there exists an $N \in \mathbb{N}$ such that for all $M \geq N$ the norm

$$\|f - \sum_{n=-m}^{M} \alpha_n \psi_n\| < \varepsilon \tag{9.1}$$

1. There is an $\widetilde{f}$ 2π-periodic, infinitely smooth function such that $\|f - \widetilde{f}\| < \varepsilon/3$

2. The fourier series $\sum_{-\infty}^{\infty} \widetilde{\alpha}_n \psi_n \to \widetilde{f}$ uniformly, hence it converges in norm, so there is an $N \in \mathbb{N}$ such that for all $M \geq N$, $\|\widetilde{f} - \sum_{-M}^{M} \widetilde{\alpha}_n \psi_n\| < \varepsilon/3$

3. So

$$\begin{aligned}
\|\sum_{-M}^{M} (\widetilde{\alpha}_n - \alpha_n)\psi\|^2 &= \sum_{-M}^{M} |\widetilde{\alpha}_n - \alpha_n\|^2 \text{ (Pythagorean thm)} \\
&\leq \sum_{n=-\infty}^{\infty} |\widetilde{\alpha}_n - \alpha_n|^2 \\
&\leq |\widetilde{f} - f|^2 \text{ (Bessel Inequality)}
\end{aligned}$$

Because $\widetilde{\alpha}_n - \alpha_n = \langle \widetilde{f} - f, \psi_n \rangle$.

Finally for any $M \geq N$,

$$\begin{aligned}
\|f - \sum_{-M}^{M} \alpha_n \psi_n\| &= \|f - \widetilde{f} + \widetilde{f} - \sum_{-M}^{M} \widetilde{\alpha}_n \psi_n + \sum_{-M}^{M} \widetilde{\alpha}_n \psi_n - \sum_{-M}^{M} \alpha_n \psi_n\| \\
&\leq \|f - \widetilde{f}\| + \|\widetilde{f} - \sum_{-M}^{M} \widetilde{\alpha}_n \psi_n\| + \|\sum_{-M}^{M} (\widetilde{\alpha}_n - \alpha_n)\psi_n\| \\
&< (\varepsilon/3) + (\varepsilon/3) + (\varepsilon/3) \\
&< \varepsilon.
\end{aligned}$$

□

9.1 Summary on Convergence of Fourier Series

1. The series converges unifrmly and absolutely if f is continuous and piecewise smooth.
2. The series converges pointwise and in norm if f is piecewise continuous and piecewise smooth.
3. It converges in norm if $f \in L^2(a, b)$

Parseval Equality. Let $\{\psi_n\}_{-\infty}^{\infty}$ be an orthonormal basis in $L^2(a, b)$, for any $f \in L^2(a, b)$, then

$$\|f\|^2 = \sum_{-\infty}^{\infty} |\langle f, \psi_n \rangle|^2. \tag{9.2}$$

9.2 Some Advanced Linear Algebra

Definition 9.1. A vector space V (over some field of scalars $\mathbb{F}$) with an inner product $\langle \cdot, \cdot \rangle$ and an induced norm $\| \cdot \|$ if it is complete under convergence in norm (meaning that every series of vectors converges in norm to some vector in V) is a **Hilbert Space**.

Example 9.1. Just a few examples, $L^2(a, b)$ is a Hilbert space, $\mathbb{C}^k$ is a Hilbert space too, and $\mathbb{R}^n$ is a Hilbert space as well.

Definition 9.2. Let W_1, W_2 be subspaces of a Hilbert space. A **linear transformation** $T : W_1 \to W_2$ is a linear map from W_1 to W_2. The **adjoint** of T is defined by

$$\langle T\vec{a}, \vec{b} \rangle = \langle \vec{a}, T^{\dagger}\vec{b} \rangle \tag{9.3}$$

And T is **self-adjoint** if $T = T^{\dagger}$.

Example 9.2. In $\mathbb{C}^k$, we can represent operators by square matrices $T = [t_{ij}]$ $(i, j = 1, ..., k)$. The adjoint is $T^{\dagger} = [\overline{t}_{ji}]$. Or if one prefers, this is the complex conjugate transpose of the matrix. In $\mathbb{R}^k$, self-adjoint is equivalent to being symmetric.

Now one should recall that the eigenvalues/eigenvectors of T are

$$T\vec{a} = \lambda\vec{a} \tag{9.4}$$

where λ is the eigenvalue, and $\vec{a}$ is the eigenvector. The way to think of this is that the eigenvectors are only "stretched" by the operator by a factor of λ.

For self-adjoint operators, the eigenvalues are necessarily real. To see this, let λ, $\vec{a}$ be the eigenvalue and eigenvector of a self-adjoint operator T. So

$$\begin{aligned}
\langle T\vec{a}, \vec{a}\rangle &= \langle \vec{a}, T\vec{a}\rangle \\
&= \langle \lambda\vec{a}, \vec{a}\rangle \\
&= \langle \vec{a}, \lambda\vec{a}\rangle \\
&= \bar{\lambda}\langle \vec{a}, \vec{a}\rangle = \lambda\langle \vec{a}, \vec{a}\rangle \\
\Rightarrow \lambda &= \bar{\lambda} \\
\Rightarrow &\lambda \in \mathbb{R}
\end{aligned}$$

Eigenvectors of self-adjoint operator form orthonormal basis

Theorem 9.2. (Spectral Theorem) (For the finite dimensional case, specifically $\mathbb{C}^k$ or $\mathbb{R}^k$) For any self-adjoint operator T, there is an orthonormal basis of $\mathbb{C}^k$ (or $\mathbb{R}^k$) consisting of eigenvectors of T.

Now one can ask "What the hell good is this?" Well, to solve $T\vec{x} = \vec{b}$ where $\vec{b}$ is a given vector. Suppose $\{u_1, \ldots, u_k\}$ are eigenvectors of T forming an orthonormal basis of $\mathbb{C}^k$ and the corresponding eigenvalues are $\lambda_1, \ldots, \lambda_k$. We can expand

$$\vec{b} = \sum_{j=1}^{k} \beta_j \vec{u_j} \tag{9.5}$$

where $\beta_j = \langle \vec{b}, \vec{u}_j\rangle$. Let $x_j = \langle \vec{x}, \vec{u}_j\rangle$ then $\vec{x} = \sum x_j \vec{u}_j$. Then just by associating each eigenvector to each eigenvector, we have $x_j = \beta_j$.

10 The Regular Sturm-Liouville Problem

Let us first consider an example of a regular Sturm-Liouville problem:

Example 10.1. Heat diffusion through an inhomogeneous thin rod is described by

$$\omega(x)\partial_t u(x,t) = \partial_x[r(x)\partial_x u(x,t)] + p(x)u(x,t) \tag{10.1}$$

with non-uniform material properties, where $r(x)$ is the speed of propagation of heat, $p(x)$ is the source of heat at the point x, but specifically we require $r(x)$ and $\omega(x)$ be positive, and $u(x,t)$ is the temperature at point x and time t.

We can solve this using the seperation of variables so we have $u(x,t) = X(x)T(t)$ and plug this back into the equation to find

$$\omega(x)X(x)T'(t) = T(t)\partial_x[r(x)X'(x)] + p(x)X(x)T(t) \tag{10.2}$$

we divide both sides by $\omega(x)X(x)T(t)$ and set it equal to some constant $-\lambda$:

$$\frac{T'(t)}{T(t)} = \frac{\partial_x[r(x)X'(x)]}{\omega(x)X(x)} + \frac{p(x)}{X(x)} = -\lambda. \tag{10.3}$$

We get two equations

$$T' + \lambda T = 0 \Rightarrow T(t) = T_0 e^{-\lambda t} \tag{10.4}$$

where T_0 is a constant dependent on the initial conditions, and

$$\partial_x[r(x)X'(x)] + p(x)X(x) + \lambda\omega(x)X(x) = 0. \tag{10.5}$$

Solving this equation is equivalent to solving the Sturm-Liouville problem. Similarly, if we try to solve the wave equation we get to boil it down to the same problem.

Definition 10.1. Let L be a map

$$L : C^2(a,b) \to L^2(a,b)$$

be a **linear differential operator** given by

$$\begin{aligned} L(f) &= [rf']' + pf \\ &= rf'' + r'f' + pf \end{aligned}$$

where $r(x)$, $r'(x)$, $p(x)$ are all continuous and real-valued on $[a,b]$, and $r(x) > 0$ on $[a,b]$.

Remark 10.1. From (10.5) it is really equivalent to

$$L(f) + \lambda\omega(x)f(x) = 0 \tag{10.6}$$

We can verify that L is linear, or in other words

$$L(c_1 f + c_2 g) = c_1 L(f) + c_2 L(g) \tag{10.7}$$

where c_1, c_2 are real constants, $f, g \in L^2(a,b)$. So if f, g satisfy the equation, then so does $c_1 f + c_2 g$.

So really we have $L(f) = -\lambda\omega(x)f(x)$. This looks remarkably like an eigenvalue problem from Linear Algebra (ahem hint hint, wink wink, nudge nudge). If we can show that L is adjoint, then we can show that the eigenfunctions f form an orthonormal basis.

Question: What is the adjoint of L? To find out we need to find what $\langle L(f), g\rangle$ is, and see what $\langle f, L(g)\rangle$ is.

Let $f, g \in C^2(a, b)$, then

$$\langle L(f), g\rangle = \int_a^b \left(\partial_x[r(x)f'(x)] + p(x)f(x)\right)\overline{g(x)}dx \tag{10.8a}$$

$$= \int_a^b \partial_x[r(x)f'(x)]\overline{g(x)}dx + \int_a^b p(x)f(x)\overline{g(x)}dx \tag{10.8b}$$

Now we can integrate by parts to find

$$\int_a^b \partial x[r(x)f'(x)]\overline{g(x)}dx = r(x)f'(x)g(x)|_a^b - \int_a^b r(x)f'(x)\overline{g'(x)}dx \tag{10.9}$$

So we can put this back in to find

$$\langle L(f), g\rangle = r(x)f'(x)g(x)|_a^b - \int_a^b r(x)f'(x)\overline{g'(x)}dx + \int_a^b p(x)f(x)\overline{g(x)}dx \tag{10.10}$$

By integration by parts again we find

$$\begin{aligned} &r(x)f'(x)\overline{g(x)}|_a^b - \int_a^b r(x)f'(x)\overline{g'(x)}dx + \int_a^b p(x)f(x)\overline{g(x)}dx \\ &= r(x)f'(x)\overline{g(x)}|_a^b - r(x)\overline{g'(x)}f(x)|_a^b - \int_a^b [r(x)\overline{g'(x)}]'f(x)dx + \int_a^b p(x)f(x)\overline{g(x)}dx \end{aligned} \tag{10.11}$$

We can clean this up a little to find

$$\langle L(f), g\rangle = \int_a^b f(x)\left(\partial_x[r(x)\overline{g'(x)}] + g(x)\overline{g(x)}\right)dx + \underbrace{\left[r(x)f'(x)\overline{g(x)} - r(x)f(x)\overline{g'(x)}\right]}_{\text{boundary terms!}} \tag{10.12}$$

We can write now

$$\langle L(f), g\rangle = \int_a^b f(x)\overline{\left(\partial_x[r(x)\overline{g'(x)}] + g(x)\overline{g(x)}\right)}dx + \text{bdry terms} \tag{10.13}$$

where "bdry terms" are the boundary terms, or more simply

$$\langle L(f), g\rangle = \langle f, L(g)\rangle + \text{bdry terms}. \tag{10.14}$$

Definition 10.2. L is called **formally self-adjoint** because

$$\langle L(f), g\rangle = \langle f, L(g)\rangle + \text{boundary terms} \tag{10.15}$$

this is sometimes called the "Lagrange identity".

Remark 10.2. If the boundary conditions make

$$[r(x)(f'(x)\overline{g(x)} - f(x)\overline{g'(x)})]_a^b = 0 \tag{10.16}$$

then L is self-adjoint. Such boundary conditions we call self-adjoint boundary conditions (we will formally define it below).

If we expand this out we find

$$[r(f'\bar{g}-f\bar{g}')]_a^b = r(b)(f'(b)\bar{g}(b)-f(b)\bar{g}(b)')-r(a)(f'(a)\bar{g}(a)-f(a)\bar{g}(a)') \tag{10.17}$$

This gives us a set of two boundary conditions

$$\begin{aligned} B_1(f) &= \alpha_1 f(a) + \alpha_2 f'(a) + \beta_1 f(b) + \beta_2 f'(b) = 0 \\ B_2(f) &= \alpha_3 f(a) + \alpha_4 f'(a) + \beta_3 f(b) + \beta_4 f'(b) = 0 \end{aligned}$$

where the α's and β's are constants.

Definition 10.3. The boundary conditions B_1, B_2 are called **self-adjoint** with respect to the operator L if

$$[r(x)(f'(x)\overline{g(x)} - f(x)\overline{g'(x)})]_a^b = 0 \tag{10.18}$$

for f, g satisfy $B_1(f) = 0$, $B_2(f) = 0$, $B_1(g) = 0$, $B_2(g) = 0$.

Example 10.2. There are several famous boundary conditions used all the time:

1. The Dirichlet boundary conditions $f(a) = f(b) = 0$;
2. The Neumann boundary conditions $f'(a) = f'(b) = 0$;
3. The periodic boundary conditions $f(a) = f(b)$ and $f'(a) = f'(b)$.

Definition 10.4. (Regular Sturm Liouville Problem) Find all solutions of the boundary value problem

$$\begin{gathered} L(f) + \lambda\omega(x)f(x) = 0 \\ B_1(f) = 0,\ B_2(f) = 0 \end{gathered}$$

where λ is an arbitrary constant and

i) $L(f) = [rf']' + pf$ where $r(x)$, $r'(x)$, $p(x)$ are real and continuous on $[a, b]$, and $r(x) > 0$ on $[a, b]$.

ii) $B_1(f) = B_2(f) = 0$ are self-adjoint with respect to L

iii) $\omega(x)$ is positive and continuous on $[a, b]$.

Remark 10.3. If $r(x) = 1$ and $p(x) = 0$ and $\omega(x) = 1$, then we get

$$f''(x) + \lambda f(x) = 0 \tag{10.19}$$

or in other words, we have the wave equation on an elastic string.

Definition 10.5. For ϕ a solution of Sturm-Liouville problem corresponding to some constant λ, with $\phi \neq 0$m we call ϕ the **eigenvector** (or more often **eigenfunction**) corresponding to the **eigenvalue** λ

Note this definition is because

$$L(f) = -\omega(x)\lambda f(x) \Rightarrow \frac{1}{\omega(x)} L(f) = -\lambda f(x)$$

so λ is an eigenvalue of the linear operator $M(f) = L(f)/\omega(x)$ and eigenfunction $f(x)$.

Definition 10.6. Suppose that $\omega(x)$ is integrable on $[a, b]$ and $\omega(x) > 0$ almost everywhere on $[a, b]$, then the **weighted L^2-Space** is defined as the set

$$L^2_\omega(a, b) = \{f : \int_a^b |f(x)|^2 \omega(x) dx < \infty\} \tag{10.20}$$

11 Properties of Weighted Inner Products

Recall we just introduced $L^2_\omega(a, b) = \{f : (a, b) \to \mathbb{R} | \int_a^b |f(x)|^2 \omega(x) dx < \infty\}$. For us, in practice, we'll often have $\omega(x) > 0$ and it's continuous on $[a, b]$. We have weighted inner products and weighted norms on $L^2_\omega(a, b)$ is defined as follows:

$$\langle f, g \rangle_\omega = \int_a^b f(x)\overline{g(x)}\omega(x) dx \tag{11.1}$$

$$\|f\|_\omega = \sqrt{\langle f, f \rangle_\omega} = \left(\int_a^b |f(x)|^2 \omega(x) dx \right)^{1/2} \tag{11.2}$$

Let demonstrate the weighted inner product and weighted norm satisfy the fundamental properties of the inner product and norm (respectively).

(i) $\langle f, 0 \rangle_\omega = 0$, this is trivial (it really is, unlike all those other times I said something was trivial)

Linearity in First Slot

(ii) Linear in the first slot

$$\begin{aligned} \langle \alpha f + \beta g, h \rangle_\omega &= \int_a^b (\alpha f(x) + \beta g(x))\overline{h(x)}\omega(x) dx \\ &= \int_a^b \alpha f(x)\overline{h(x)}\omega(x) dx + \int_a^b \beta g(x)\overline{h(x)}\omega(x) dx \\ &= \alpha \langle f, g \rangle_\omega + \beta \langle g, h \rangle_\omega \end{aligned}$$

(iii) Antilinear in the second slot *Antilinear in Second Slot*

$$
\begin{aligned}
\langle f, \alpha g + \beta h\rangle_\omega &= \int_a^b f(x)\overline{[\alpha g(x) + \beta h(x)]}\omega(x)dx \\
&= \int_a^b f(x)\overline{\alpha h(x)}\omega(x)dx + \int_a^b g(x)\overline{\beta h(x)}\omega(x)dx \\
&= \bar{\alpha}\langle f, g\rangle_\omega + \bar{\beta}\langle f, g\rangle_\omega
\end{aligned}
$$

(iv) Hermitian symmetric

$$
\begin{aligned}
\langle f, g\rangle_\omega &= \int_a^b f(x)\overline{g(x)}\omega(x)dx \\
&= \overline{\int_a^b \overline{f(x)}g(x)\omega x dx}, \qquad \omega : \mathbb{C} \to \mathbb{R} \\
&= \overline{\langle g, f\rangle}_\omega
\end{aligned}
$$

Those are the fundamental properties of the inner product, so let us now consider the fundamental properties of the norm.

(a) (Homogeneity) $\|\alpha f\|_\omega = |\alpha|\|f\|_\omega$ (this is trivial).

(b) (Positivity) $\|f\|_\omega \geq 0$ and $\|f\|_\omega = 0$ if and only if $f = 0$.

Observe

$$
\|f\|_\omega^2 = \int_a^b \underbrace{|f(x)|^2\omega(x)}_{\text{always positive } \Rightarrow \text{ nonzero integral}} dx \tag{11.3}
$$

If $\omega(x) = 0$ on a small interval from $[\alpha, \beta] \subset [a, b]$, then if $f(x) \neq 0$ on $f[\alpha, \beta]$ and $f(x) = 0$ otherwise, but the norm vanishes

$$
\|f\|_\omega = \int_a^b |f(x)|^2\omega(x)dx = \int_\alpha^\beta |f(x)|^2\omega(x)dx = 0. \tag{11.4}
$$

So

$$
\omega(x) > 0 \qquad \forall x \in [a, b] \tag{11.5}
$$

is needed for the positivity of the norm. Therefore all properties of the norm also hold.

Remark 11.1. The weighted inner product

$$
\langle f, g\rangle_\omega = \int_a^b f(x)\overline{g(x)}\omega(x)dx \tag{11.6}
$$

can be translated into the standard inner product

$$\langle f\omega, g\rangle = \langle f, \omega g\rangle = \langle f, g\rangle_\omega \tag{11.7}$$

because ω is real valued, i.e. $\overline{\omega(x)} = \omega(x)$ for all $x \in [a, b]$. We see that

$$0 < m = \min_{x\in[a,b]} \omega(x) \leq \omega(x) \leq M = \max_{x\in[a,b]} \omega(x) \tag{11.8}$$

which implies

$$0 < m|f(x)|^2 \leq \omega(x)|f(x)|^2 \leq M|f(x)|^2 \tag{11.9}$$

and

$$\overbrace{m\int_a^b |f(x)|^2dx \leq \int_a^b |f(x)|^2\omega(x)dx}^{\text{if } f\in L^2_\omega\text{, then the } L^2 \text{ norm is bounded by this}} \leq M\int_a^b |f(x)|^2dx \tag{11.10}$$

$$\underbrace{\int_a^b |f(x)|^2\omega(x)dx \leq M\int_a^b |f(x)|^2dx}_{\text{if } f\in L^2_\omega\text{, then the norm is finite}}$$

Thus $L^2_\omega(a, b) = L^2(a, b)$. In fact, if $\{\phi_n\}_1^\infty$ is an orthonormal basis for $L^2(a, b)$, then we may obtain a basis $\{\psi_n\}_1^\infty$ for $L^2_\omega(a, b)$ where

$$\psi_n(x) = \frac{1}{\sqrt{\omega(x)}}\phi_n(x). \tag{11.11}$$

The other way holds too, given ψ_n we find

$$\phi_n = \sqrt{\omega(x)}\psi_n \tag{11.12}$$

as an orthonormal basis for L^2.

We can use this to solve the regular Sturm-Liouville problem. We want to prove the Sturm-Liouville problme has its eigenfunctions form an orthonormal basis of $L^2(a, b)$. Recall the linear operator we're working with is

$$L(f) + \lambda\underbrace{\omega(x)}_{\uparrow}f(x) = [r(x)f'(x)]' + p(x)f(x) + \lambda\underbrace{\omega(x)}_{\uparrow}f(x) = 0 \tag{11.13}$$

$\omega(x)$ used in the weighted inner product

The self-adjoint boundary conditions:

$$B_1(f) = \alpha_1 f(a) + \alpha_2 f'(a) + \beta_1 f(b) + \beta_2 f'(b) = 0 \tag{11.14a}$$
$$B_2(f) = \alpha_3 f(a) + \alpha_4 f'(a) + \beta_3 f(b) + \beta_4 f'(b) = 0 \tag{11.14b}$$

For any λ, f is a solution (a nontrivial solution, read: nonzero solution) of the Sturm-Liouville problem, then λ is called an **eigenvalue** and f is called the **eigenfunction** of the Sturm-Liouville problem.

Remark 11.2. Note that we have

$$\langle L(f), f\rangle = \langle f, L(f)\rangle \tag{11.15}$$

The boundary conditions make all other terms vanish. If both f and g satisfy the boundary conditions, then

$$\langle L(f), g\rangle = \langle f, L(g)\rangle. \tag{11.16}$$

11.1 Some Properties of Eigenvalues and Eigenfunctions of the Sturm-Liouville Problem

Theorem 11.1. For the regular Sturm-Liouville problem, there are three properties that hold:

1. All eigenvalues are real
2. Eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the weighted inner product
3. For any eigenvalue λ, the eigenspace is at most two-dimensional (there are at most 2 linearly independent eigenfunctions associated with λ).

Moreover, if the boundary conditions are seperated (i.e. $B_1(f)$ focuses on one endpoint and $B_2(f)$ focuses on the other, e.g.

$$\begin{aligned} B_1(f) &= \alpha_1 f(a) + \alpha_2 f'(a) \\ B_2(f) &= \beta_1 f(a) + \beta_2 f'(a) \end{aligned}$$

then for any eigenvalue λ, the eigenspace associated to λ is one dimensional.

Proof. 1. We showed for Hermitian matrices, eigenvalues are real, so lets take the same steps. Let λ, f be an eigenvalue/eigenfunction pair, then

$$\begin{aligned} \lambda\|f\|_\omega^2 &= \langle \lambda f, f\rangle_\omega \quad \lambda \text{ is real} \\ &= \langle \lambda\omega f, f\rangle \quad \text{by relation of } L^2 \text{ to } L^2_\omega \\ &= \langle -L(f), f\rangle \quad \text{by Sturm-Liouville equation} \\ &= \langle -f, L(f)\rangle \quad \text{by Self-Adjointness of } L \\ &= \langle f, -L(f)\rangle \\ &= \langle f, \lambda\omega f\rangle \quad \text{by S-L eigenproblem} \\ &= \langle f, \lambda f\rangle_\omega \quad \text{by relation of } L^2 \text{ to } L^2_\omega \\ &= \bar{\lambda}\|f\|_\omega^2 \quad \text{by Antilinearity} \\ \lambda = \bar{\lambda} &\Rightarrow \lambda \in \mathbb{R} \end{aligned}$$

2. Suppose f, g are eigenfunctions corresponding to eigenvalues λ, μ where $\lambda \neq \mu$. Then we take the inner product

$$\begin{aligned}
\lambda\langle f, g\rangle_\omega &= \langle \lambda\omega f, g\rangle \\
&= \langle -L(f), g\rangle \\
&= \langle f, -L(g)\rangle \quad \text{both are eigenfunctions, boundary and terms vanish} \\
&= \langle f, \mu\omega g\rangle \quad \text{by S-L problem} \\
&= \mu\langle f, g\rangle_\omega
\end{aligned}$$

$$\Rightarrow (\lambda - \mu)\langle f, g\rangle_\omega = 0$$

but $\lambda - \mu \neq 0$ so it implies $\langle f, g\rangle_\omega = 0$.

3. Apply the existence of solution theorem to the Sturm-Liouville problem (For $L(f) + \lambda\omega f = 0$ is a second order ODE with boundary conditions $f(a) = c_1$, $f(b) = c_2$, then there exists a unique solution.) For us,

$$\begin{aligned}
B_1(f) &= \alpha_1 f(a) + \alpha_2 f'(a) + \beta_1 f(b) + \beta_2 f'(b) = 0 \\
B_2(f) &= \alpha_3 f(a) + \alpha_4 f'(a) + \beta_3 f(b) + \beta_4 f'(b) = 0
\end{aligned}$$

with two free variables (i.e. two degrees of freedom). There are 4 unknowns and two equations imposed on them, so we have 2 degrees of freedom. There are then at most 2 linearly independent eigenfunctions. When the boundary conditions are set:

$$\left.\begin{aligned}
B_1(f) &= \alpha_1 f(a) + \alpha_2 f'(a) \\
B_2(f) &= \beta_1 f(b) + \beta_2 f'(b)
\end{aligned}\right\} \Rightarrow \text{only one degree of freedom} \tag{11.17}$$

Which concludes our proof. □

12 Solving the Heat Equation using Sturm-Liouville Problem

12.1 Existence of Eigenfunctions for the Regular Sturm-Liouville Problem

Theorem 12.1. For the regular Sturm-Liouville problem,

a) There is ano orthonormal basis $\{\phi_n\}_1^\infty$ of $L^2_\omega(a, b)$ consisting of the eigenfunctions of the regular Sturm-Liouville problem.

b) The eigenvalues satisfy $\lim_{n\to\infty} \lambda_n = \infty$

c) If $f \in C^2(a, b)$ and f satisfies the boundary conditions $B_1(f) = B_2(f) = 0$, then

$$\sum \langle f, \phi_n\rangle_\omega \phi_n \to f \text{ uniformly}$$

Implications of Theorems (11.1) **and** (12.1)**:** Regular Sturm-Liouville problem has a solution ϕ if and only if $\lambda \in \mathbb{R}$ (it has real eigenvalues) and for only countably many λ_n.

Remark 12.1. The eigenvalues are determined by the boundary condition

Applications: Strategy for solving the heat diffusion problem is as follows

$$\omega(x)\partial_t u(x,t) = \partial_x[r(x)\partial_x u(x,t)] + p(x)u(x,t) + \underbrace{0}_{\text{no external force or source}} \tag{12.1}$$

with boundary conditions $B_1(u) = B_2(u) = 0$. This is a boundary value problem; the initial value problem has $u(x,0) = f(x)$ which determines the initial distribution of heat. With these conditions we have an initial value problem.

Step 1) Apply seperation of variables $u(x,t) = X(x)T(t)$

$$\Rightarrow T'(t) + \lambda T(t) = 0 \Rightarrow T(t) = T_0 e^{-\lambda t}$$
$$\Rightarrow L(X(x)) + \lambda\omega(x)X(x) = 0 \text{ (Sturm-Liouville Problem!)}$$

Step 2) Find **all** eigenvalues and normalized eigenfunctions for the Sturm-Liouville problem. For each pair (λ_n, ϕ_n) we get one solution to boundary value problem

$$u_n(x,t) = c_n e^{-\lambda_n t}\phi_n(x) \tag{12.2}$$

so $u(x,t) = \sum_{n=1}^{\infty} c_n \exp(-\lambda_n t)\phi_n(x)$ is also a solution to the boundary value problem.

Step 3) Insert the initial value:

$$\sum_{n=1}^{\infty} c_n\phi_n = u(x,0) = f(x) \tag{12.3}$$

So $c_n = \langle f, \phi_n\rangle_\omega$, we conclude the solution to the initial value problem is

$$u(x,t) = \sum_{n=1}^{\infty} \langle f, \phi_n\rangle_\omega e^{-\lambda_n t}\phi_n(x) \tag{12.4}$$

(In practice, finding the eigenvalues λ is always the hardest part!)

Example 12.1. Consider the one-dimensional heat flow The rod is insulated and of uniform material property. So

$$\partial_t u(x,t) = k\partial_x^2 u(x,t) \tag{12.5}$$

where k is the conductance of the material. We have

$$\text{bdry conditions} \begin{cases} \partial_x u(0,t) & = \alpha u(0,t),\ \alpha > 0 \\ \partial_x u(L,t) & = -\alpha u(L,t) \end{cases} \tag{12.6}$$

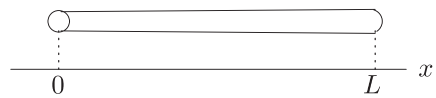

Figure 12.1: A thin insulated metal rod of length L

and our initial value problem is

$$u(x,0) = f(x) \in L^2(0,1) \tag{12.7}$$

What's the eigenvalues and eigenfunctions? Well, the seperation of variables gives us

$$\frac{T'(t)}{T(t)} = \frac{kX''(x)}{X(x)} = \underbrace{-k\nu^2}_{\text{constant}} \tag{12.8}$$

We have two equations

$$T'(t) + k\nu^2 T(t) = 0 \tag{12.9a}$$

$$kX''(x) + k\nu^2 X(x) = 0 \tag{12.9b}$$

with the second equation be such that $X'(0) = \alpha X(0)$, $X'(L) = -\alpha X(L)$. But the second equation is the regular Sturm-Liouville problem with $\omega(x) = r(x) = 1$, $p(x) = 0$. This is the usual $L^2(0,L)$ space.

So what are the eigenvalues $\lambda = \nu^2$ and what are the eigenfunctions? Well, there are two cases:

Case 1 $\nu^2 = 0$ which implies $X''(0) = 0$ and $\alpha X(0) = X'(0)$. So $X(x) = c_1 + c_2$ but $X'(0) = \alpha X(0)$, thus $c_2 = \alpha c_1$. And $X'(L) = -\alpha X(L)$ implies $c_2 = \alpha(c_1 + c_2 L)$. Thus

$$2c_2 + \alpha c_2 L = 0 = (2 + \alpha L)c_2$$

but $2 + \alpha L > 0$ so $c_2 = 0$, which implies trivially $X(x) = 0$.

Case 2 We know what Theorem (11.1) says about eigenvalues, so $\lambda \in \mathbb{R}$. So either 1. $\nu^2 > 0$ which implies $\nu > 0$

$$X''(x) + \nu^2 X(x) = 0 \tag{12.10}$$

the characteristic equation is

$$r^2 + \nu = 0 \Rightarrow \nu = \pm i\mu \tag{12.11}$$

So the general solution is

$$X(x) = c_1 \cos(\nu x) + c_2 \sin(\nu x) \tag{12.12}$$

By the boundary conditions

$$X'(0) = \alpha X(0)$$
$$\Rightarrow c_2\nu = c_1\alpha \Rightarrow c_2 = \frac{\alpha}{\nu}c_1$$

multiply the equation by ν/c_1 we get

$$X(x) = \nu\cos(\nu x) + \alpha\sin(\nu x) \tag{12.13}$$

which corresponds to eigenvalues of ν^2. So all we need to do is find ν, so lets apply the second boundary equation

$$X'(L) = -\alpha X(L) \tag{12.14a}$$
$$\Rightarrow -\nu^2\sin(\nu L) + \alpha\nu\cos(\nu L) = -\alpha\left(\nu\cos(\nu L) + \alpha\nu\sin(\nu L)\right) \tag{12.14b}$$
$$\Rightarrow 2\alpha\nu\cos(\nu L) = (\nu^2 - \alpha^2)\sin(\nu L) \tag{12.14c}$$
$$\Rightarrow \tan(\nu L) = \frac{2\alpha\nu}{\nu^2 - \alpha^2} \tag{12.14d}$$

Remark 12.2. Note that ν cannot possibly equal $i\mu$ because we can plug it back in to what we just found to see that

$$\tan(\nu L) = \tan(i\mu L) = \frac{i2\alpha\mu}{-\mu^2 - \alpha^2} \tag{12.15}$$

but from complex analysis we should remember that

$$\tan(i\theta) = i\tanh(\theta) = i\frac{e^\theta - e^{-\theta}}{e^\theta + e^{-\theta}} \tag{12.16}$$

Thus

$$i\tanh(\mu L) = -i\frac{2\alpha\mu}{\mu^2 + \alpha^2} \tag{12.17}$$

but since $\mu > 0$, $\tanh(\mu L) > 0$ so we have a contradiction.

13 Using Fourier Series to Solve the Heat Equation

We were looking at the Sturm-Liouville problem which came from the heat equation

$$X''(x) + \nu^2 X(x) = 0 \tag{13.1}$$

with the boundary conditions

$$X'(0) = \alpha X(0) \tag{13.2a}$$
$$X'(L) = -\alpha X(L) \tag{13.2b}$$

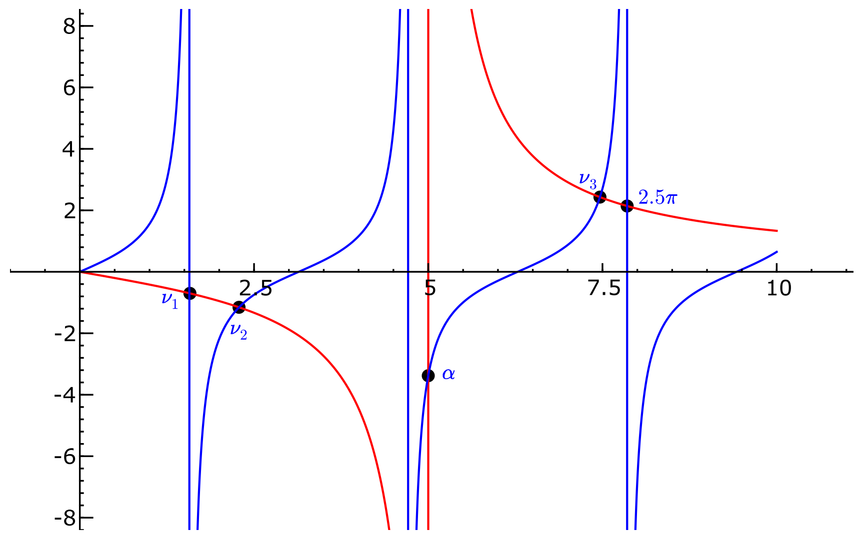

Figure 13.1: A Plot of the Eigenvalue relations for the Heat Equation from Eq (13.3) setting $\alpha = 5$ and $L = 1$. The blue line is $\tan(\nu L)$, the red line is $2\alpha\nu/(\nu^2 - \alpha^2)$. Note that the points labeled α and 2.5π are not eigenvalues but singular points of $2\alpha\nu/(\nu^2 - \alpha^2)$ and $\tan(\nu L)$ respectively that coincidentally overlap in the graph program.

With eigenfunctions $\{\nu\cos(\nu x) + \alpha\sin(\nu x)\}$ where the eigenvalues ν^2 are determined by

$$\tan(\nu L) = \frac{2\alpha\nu}{\nu^2 - \alpha^2} \tag{13.3}$$

But this isn't the most convenient way to determine the eigenvalues, as it's not really convenient to do analytically. The plot of the left hand side in blue and the right hand side in red is given in figure (13.1). We find that $0 < \nu_1 < \nu_2 < \cdots$. We can normalize the eigenfunctions $\phi_\nu(x)$ to find

$$\|\phi_\nu\|^2 = \int_0^L (\nu\cos(\nu x) + \alpha\sin(\nu x))^2 dx \tag{13.4a}$$

$$= \int_0^L \nu^2\cos^2(\nu x) + 2\alpha\nu\sin(\nu x)\cos(\nu x) + \alpha^2\sin^2(\nu x) dx \tag{13.4b}$$

Remember from basic trigonometry

$$\cos^2\theta = \frac{1}{2}(1 + \cos(2\theta))$$
$$\sin^2\theta = \frac{1}{2}(1 - \cos(2\theta))$$

so plugging this in we find

$$\|\phi_\nu(x)\|^2 = \frac{\nu^2}{2}\int_0^L (1+\cos(2\nu x))dx + 2\alpha\int_0^L \sin(\nu x)d\left(\sin(\nu x)\right) + \frac{\alpha^2}{2}\int_0^L (1-\cos(\nu x))dx \tag{13.5a}$$

$$= \left(\frac{\nu^2+\alpha^2}{2}\right)L + \alpha\sin^2(\nu L) + \int_0^L \frac{\nu}{2}\cos(2\nu x) - \frac{\alpha^2}{2}\cos(2\nu x)dx \tag{13.5b}$$

$$= \left(\frac{\nu^2+\alpha^2}{2}\right)L + \alpha\sin^2(\nu L) + \left(\frac{\nu^2-\alpha^2}{4\nu}\right)\sin(2\nu L) \tag{13.5c}$$

We can now take advantage of our eigenrelations

$$\tan(\nu L) = \frac{2\alpha\nu}{\nu^2-\alpha^2} \tag{13.6}$$

to reduce our calculation of the norm of $\phi_\nu(x)$ to be

$$\frac{1}{2}\left(\frac{\nu^2-\alpha^2}{2\nu}\right) = \frac{1}{2}\frac{1}{\tan(\nu L)} = \frac{1}{2}\frac{\cos(\nu L)}{\sin(\nu L)} \tag{13.7}$$

thus

$$\frac{\nu^2-\alpha^2}{2(2\nu)}\sin(\nu L) = \frac{1}{2}\cot(\nu L)2\sin(\nu L)\cos(\nu L) \tag{13.8a}$$

$$= \alpha\cos^2(\nu L) \tag{13.8b}$$

So $\|\phi_\nu(x)\|^2 = [(1+L/2)\alpha^2 + \nu^2 L/2]^{1/2}$, and the orthonormal basis formed for $L^2(0,L)$ satisfies the problem

$$\partial_t u(x,t) = k\partial_x^2 u(x,t) \tag{13.9}$$

where

$$u_n(x,t) = c_n e^{-\nu_n^2 kt}\phi_n(x) \tag{13.10}$$

We need to now find what all the c_n's are. (Yes, our work never ends; we're like Pinkerton's, we never rest.)

To do so, we let

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} c_n e^{-\nu_n^2 kt}\phi_n(x) \tag{13.11a}$$

$$u(x,0) = f(x) = \sum_{n=1}^{\infty}\langle f,\phi_n\rangle\phi_n \tag{13.11b}$$

$$= \sum_{n=1}^{\infty} u_n(x,0) = \sum_{n=1}^{\infty} c_n\phi_n \tag{13.11c}$$

Take the difference, we find

$$\sum_{n=1}^{\infty}(c_n - \langle f, \phi_n \rangle)\phi_n = 0 \tag{13.12}$$

because ϕ_n is a basis and this linear combination is zero, we conclude that

$$c_n = \langle f, \phi_n \rangle. \tag{13.13}$$

Change boundary conditions

Lets change the boundary conditions a bit to be

$$X''(x) + \nu^2 X(x) = 0, \qquad X(0) = X(L) = 0 \tag{13.14}$$

This corresponds to the heat equation

$$\partial_t u(x,t) = k\partial_x^2 u(x,t), \qquad u(0,t) = u(L,t) = 0 \tag{13.15}$$

The general solution of $X''(x) + \nu^2 X(x) = 0$ is

$$X(x) = c_1 \cos(\nu x) + c_2 \sin(\nu x) \tag{13.16}$$

Plug in the first boundary condition

$$X(0) = c_1 = 0 \Rightarrow c_1 = 0 \tag{13.17}$$

Thus

$$X(x) = c_2 \sin(\nu x) \tag{13.18}$$

which is normalized to be the normalized eigenfunction

$$X(x) = \sin(\nu x). \tag{13.19}$$

The eigenvalues are found by simpy plugging in the second boundary condition

$$X(L) = 0 \Rightarrow \sin(\nu L) = 0 \Rightarrow \nu = n\pi/L \tag{13.20}$$

The eigenvalues are thus

$$\nu_n = \frac{n}{L}\pi \tag{13.21}$$

So the eigenfunctions form the set $\{\sin(n\pi/L)\}_1^\infty$ which is the basis for $L^2(0,L)$. We choose $n \in \mathbb{N}$ because $X(x)$ is an odd function $X(-x) = -X(x)$.

Example 13.1. Consider an external source added to the heat equation

$$\partial_t u(x,t) = k\partial_x^2 u(x,t) + \underbrace{F(x,t)}_{\text{external source}} \tag{13.22}$$

with the boundary conditions $u(0,t) = u(L,t) = 0$ and $u(x,0) = 0$. For each t, the source $F(x,\cdot)$ is just a function of x and we know $F(x,\cdot) \in L^2(0,L)$. If this is the case, we can expand F in terms of the eigenbasis we just obtained. So *for each* t we expand

$$F(x,t) = \sum_{n=1}^{\infty} \beta_n(t) \sin\left(\frac{nx}{L}\pi\right) \tag{13.23}$$

If we know the form of F, we can compute the β_n coefficients directly.

We want to find $u(x,t)$, so we assume that $u(\cdot,t) \in C^1(0,\infty)$ and $u(x,\cdot) \in C^2(0,L)$. We can expand

$$u(x,t) = \sum_{n=1}^{\infty} \underbrace{b_n(t)}_{\text{unknown}} \sin\left(\frac{nx}{L}\pi\right) \tag{13.24}$$

where $b_n(t)$ is unknown. We need to find the coefficients $b_n(t)$ then we've found $u(x,t)$.

We can see that the boundary conditions are satisfied since

$$u(0,t) = \sum_{n=1}^{\infty} b_n(t)(0) = 0 \tag{13.25a}$$

$$u(L,t) = \sum_{n=1}^{\infty} b_n(t)(0) = 0 \tag{13.25b}$$

We also want the expansion to satisfy the initial condition

$$u(x,0) = 0 \Rightarrow b_n(0) = 0 \quad \forall n \in \mathbb{Z} \tag{13.26}$$

We have one fact about $b_n(t)$.

We can plug our series into our differential equation to find

$$\partial_t u(x,t) = k\partial_x^2 u(x,t) + F(x,t) \tag{13.27a}$$

$$\sum_{n=1}^{\infty} b_n'(t) \sin\left(\frac{nx}{L}\pi\right) = k \sum_{n=1}^{\infty} b_n(t) \left(\frac{-n^2\pi^2}{L^2}\right) \sin\left(\frac{nx}{L}\pi\right) + \sum_{n=1}^{\infty} \beta_n(t) \sin\left(\frac{nx}{L}\pi\right) \tag{13.27b}$$

$$\sum_{n=1}^{\infty} \left(b_n'(t) + \frac{n^2\pi^2}{L^2} k b_n(t)\right) \sin\left(\frac{nx}{L}\pi\right) = \sum_{n=1}^{\infty} \beta_n \sin\left(\frac{nx}{L}\pi\right) \tag{13.27c}$$

Now we are kind of happy, but it would be great if we could just associate the terms together? Well, we can, because the $\sin(\cdots)$ functions form an orthonormal basis, so we can then write

$$\sum_{n=1}^{\infty} \left(b_n'(t) + \frac{n^2\pi^2}{L^2} k b_n(t) - \beta_n(t)\right) \sin\left(\frac{nx}{L}\pi\right) = 0 \tag{13.28a}$$

$$\Rightarrow b_n'(t) + \frac{n^2\pi^2}{L^2} k b_n(t) - \beta_n(t) = 0 \tag{13.28b}$$

Thus we have

$$b_n'(t) + \frac{n^2\pi^2}{L^2}kb_n(t) = \beta_n(t). \tag{13.29}$$

Let $\lambda_n = kn^2\pi^2/L^2$, then we can solve

$$b_n'(t) + \lambda_n b_n(t) = \beta_n(t) \tag{13.30}$$

via the method of integrating factor. We have the integrating factor be

$$\exp(\int \lambda_n dt) = \exp(\lambda_n t) \tag{13.31}$$

so we multiply both sides by this quantity to find

$$e^{\lambda_n t}b_n'(t) + \lambda_n e^{\lambda_n t}b_n(t) = \beta_n(t)e^{\lambda_n t}. \tag{13.32}$$

We see that we can simplify this to be

$$\frac{d}{dt}(e^{\lambda_n t}b_n(t)) = \beta_n(t)e^{\lambda_n t} \tag{13.33}$$

and we can intgrate both sides with respect to t to find

$$\int_0^t \frac{d}{ds}(e^{\lambda_n s}b_n(s))ds = \int_0^t e^{\lambda_n s}\beta_n(s)ds \tag{13.34a}$$

$$e^{\lambda_n s}b_n(s)|_0^t = \int_0^t e^{\lambda_n s}\beta_n(s)ds \tag{13.34b}$$

$$e^{\lambda_n t}b_n(t) = \int_0^t e^{\lambda_n s}\beta_n(s)ds \tag{13.34c}$$

$$\Rightarrow b_n(t) = e^{-\lambda_n t}\int_0^t e^{\lambda_n s}\beta_n(s)ds \tag{13.34d}$$

SO we have just found the solution for $u(x,t)$ given some $F(x,t)$ external source.

Observe that we can do the same trick with the wave equation

$$\partial_t^2 u(x,t) = c^2\partial_x^2 u(x,t) \tag{13.35}$$

with boundary conditions

$$u(0,t) = u(L,t) = 0. \tag{13.36}$$

We use the same Sturm-Liouville trick. We need more initial conditions however

$$u(x,0) = f(x) \text{ and } \partial_t u(x,0) = g(x) \tag{13.37}$$

We can expand both f and g in terms of the basis. When we plug in

$$u(x,t) = \sum_n b_n(t)\phi_n(x) \tag{13.38}$$

we find

$$b_n''(t) + \frac{c^2 n^2 \pi^2}{L^2} b_n(t) = 0 \tag{13.39}$$

So the general solution is

$$b_n(t) = c_1 \cos\left(\frac{cn\pi}{L}t\right) + c_2 \sin\left(\frac{cn\pi}{L}t\right) \tag{13.40}$$

and we have just found the general solution in the form of a series.

Part II

Fourier Transforms

14 A Review of L^2 and an Introduction to the Fourier Transform

Note that for short hand usage we will write $L^2 = L^2(-\infty, \infty)$. More generally, we have

$$L^p = \{f \text{ defined on } \mathbb{R} : \int_{-\infty}^{\infty} |f(x)|^p dx\} \tag{14.1}$$

where $p \in \mathbb{N}$. In a very special case, when $p = 2$, we have a Hilbert space with an inner product defined by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)} dx \tag{14.2}$$

and an induced norm

$$\|f\| = \sqrt{\langle f, f \rangle} \tag{14.3}$$

All of the fundamental properties of the inner product and the norm still hold for arbitrary end points, so they still hold here.

The Cauchy-Schwarz inequality still holds

$$|\langle f, g \rangle| \leq \|f\| \|g\| \tag{14.4}$$

or equivalently

$$|\int f(x)\overline{g(x)} dx| \leq \left(\int |f(x)|^2 dx\right)^{1/2} \left(\int |g(x)|^2 dx\right)^{1/2} \tag{14.5}$$

This is not as useful as

$$\langle |f|, |g| \rangle = \int |f(x)g(x)| dx \leq \|f\| \|g\| \tag{14.6}$$

which some (e.g [2]) refer to as the Cauchy-Schwarz.

Now, if we examine L^1, it doesn't really have an inner product, but it has a norm defined on it

$$\|f\|_1 = \int |f(x)|dx. \tag{14.7}$$

Remark 14.1. Observe that $L^2(a,b) \subset L^1(a,b)$ for intervals, but this is not true for L^2 and L^1. There are $f \in L^2$ but $f \notin L^1$, and similarly there are $g \in L^1$ but $g \notin L^2$.

Example 14.1. Let

$$f(x) = \begin{cases} x^{-2/3} & \text{when } x > 1 \\ 0 & \text{otherwise} \end{cases} \tag{14.8}$$

Now

$$\int_1^\infty |f(x)|dx = \int_1^\infty x^{-2/3}dx \tag{14.9a}$$

$$= \lim_{b\to\infty} 3x^{1/3}|_1^b \tag{14.9b}$$

$$= 3(\lim_{b\to\infty} b^{1/3} - 1) = \infty \tag{14.9c}$$

which means that $f \notin L^1$. But

$$\int_1^\infty |f(x)|^2dx = \int_1^\infty x^{-4/3}dx \tag{14.10a}$$

$$= \lim_{b\to\infty} -3x^{-1/3}|_1^b \tag{14.10b}$$

$$= -3(\lim_{b\to\infty} b^{-1/3} - 1) \tag{14.10c}$$

$$= -3(-1) = 3 \tag{14.10d}$$

Thus $f \in L^2$. So L^1 functions go to zero but not as fast as L^2 functions.

Similarly we may show that

$$g(x) = \begin{cases} x^{-2/3} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{14.11}$$

Then we may show that $g \in L^1$ but $g \notin L^2$.

We will now give a list of useful facts without proofs.

1. If $f \in L^1$ and $|f(x)| \leq M$ for all $x \in \mathbb{R}$, then $f \in L^2$. So

$$\int |f(x)|^2dx \leq M \int |f(x)|dx < \infty \tag{14.12}$$

thus $f \in L^2$. In other words, $L^1 \cap L^2$ (the set of functions in both L^1 and L^2) contains bounded functions in both spaces.

2. If $f \in L^2$, and $f(x) = 0$ for x outside the interval $[a, b]$, then $f \in L^1$. Consider

$$\int |f(x)|dx = \int_a^b |f(x)|dx = \int_a^b 1 \cdot |f(x)|dx \tag{14.13}$$

then by the Cauchy-Schwarz inequality

$$\int_a^b 1 \cdot |f(x)|dx \leq \left(\int_a^b |f(x)|^2 dx\right) \left(\int_a^b 1^2 dx\right)^{1/2} \tag{14.14a}$$

$$\leq \sqrt{b-a}\|f(x)\|_2^{1/2} < \infty \tag{14.14b}$$

Definition 14.1. Let f be a function defined on the whole real line. Then the **Fourier transform** is

$$\widehat{f}(\xi) = \int_\infty^\infty f(x)e^{-ix\xi}dx \tag{14.15a}$$

$$= \langle f, e^{i\xi x}\rangle \tag{14.15b}$$

for all $\xi \in \mathbb{R}$. This is the continuous analog for the Fourier series coefficients.

The inversion formula is, if $f \in L^2$,

$$f(x) = \int \widehat{f}(\xi)e^{ix\xi}d\xi \tag{14.16}$$

for $x \in \mathbb{R}$.

14.1 Derivation of Fourier Transform

On $L^2(-\pi, \pi)$, we have the orthogonal basis $\{\exp(inx)\}_{-\infty}^{\infty}$. If we have $f \in L^2(-\pi, \pi)$, we may write it in the basis

$$f(x) = \sum c_n e^{inx} \tag{14.17}$$

where

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx \tag{14.18}$$

If f is 2π-periodic then the expansion is defined almost everwhere (at least it's defined on the interval $[-\pi, \pi]$).

We may scale this to the interval $f \in L^2(-\ell, \ell)$ where $\ell > 0$, then the basis becomes

$$\exp(inx) \to \exp\left(\frac{in\pi}{\ell}\right). \tag{14.19}$$

The Fourier expansion then becomes

$$f(x) = \sum \widetilde{c}_n e^{ix(n\pi/\ell)}, \quad \text{where } \widetilde{c}_n = \frac{1}{2\ell}\int_{-\ell}^{\ell} f(y)e^{-iny\pi/\ell}dy \tag{14.20}$$

For f defined on $\mathbb{R}$, we want to expand f as a superposition of $\exp(i\xi x)$...**HOW TO DO IT?!?**

For $[-\ell,\ell]$ when $\ell > 0$. The Fourier expansion on $[-\ell,\ell]$ and take the limit as $\ell \to \infty$:

$$f(x) = \sum_{-\infty}^{\infty} \frac{c_n}{2\ell} e^{in\pi x/\ell} \tag{14.21}$$

where

$$c_n = \langle f, e^{inx\pi/\ell}\rangle = \int_{-\ell}^{\ell} f(x) e^{-inx\pi/\ell} dx \tag{14.22}$$

We'll now turn this into a Riemann sum, let

$$\Delta\xi = \pi/\ell \tag{14.23}$$

Note interpretation of ξ_n as boundaries of subintervals

then subdividing the interval from $-\pi$ to π into 2ℓ intervals. Let

$$\xi_n = n\Delta\xi = \frac{n\pi}{\ell} = \text{ endpoints of subintervals of } [-\pi,\pi] \tag{14.24}$$

Then

$$\begin{aligned}
f(x) &= \sum \frac{c_n}{2\ell} e^{in\pi x/\ell} && (14.25a)\\
&= \sum c_n \frac{1}{2\ell} \cdot 1 \cdot e^{in\pi x/\ell} && (14.25b)\\
&= \sum c_n \left(\frac{1}{2\ell}\right)\left(\frac{\Delta\xi}{\pi/\ell}\right) e^{i\xi_n x} && (14.25c)\\
&= \sum c_n \left(\frac{1}{2\ell}\frac{\Delta\xi}{\pi/\ell}\right) e^{i\xi_n x} && (14.25d)\\
&= \sum \frac{c_n}{2\pi} e^{i\xi_n x} \Delta\xi && (14.25e)
\end{aligned}$$

where

$$c_n = \int_{-\ell}^{\ell} f(y) \exp(-i\xi_n y) dy \tag{14.26}$$

which when we take $\ell \to \infty$ is approximately

$$c_n \approx \int_{-\infty}^{\infty} f(y) e^{-i\xi_n y} dy \tag{14.27}$$

provided f vanishes rapidly as $y \to \pm\infty$. But this is the definition of $\widehat{f}(\xi_n)$. We then make this change to find $\Delta\xi \to d\xi$:

$$f(x) \approx \int_{-\infty}^{\infty} \frac{\widehat{f}(\xi_n)}{2\pi} e^{i\xi_n x} d\xi_n. \tag{14.28}$$

Thus we finish our derivation of the Fourier transform.

15 Convolution

Definition 15.1. The **convolution** of two functions is a function defined by

$$(f * g)(x) = \int_{\mathbb{R}} f(x-y)g(y)dy \tag{15.1}$$

provided that both integrals converge.

Example 15.1. Let $f(x) = x$, $g(x) = x^2$, $h(x) = \exp(-x^2)$, then

$$(f * g)(x) = \int_{-\infty}^{\infty} (x-y)y^2 dy \tag{15.2}$$

does not exist for arbitrary x. Thus the convolution is not well defined. However,

$$(f * h)(x) = \int_{-\infty}^{\infty} (x-y)e^{-y^2} dy = \sqrt{\pi}x \tag{15.3}$$

for all $x \in \mathbb{R}$.

Proof. (Proof of outrageous claim) Compute

$$\int_{-\infty}^{\infty} e^{-y^2} dy = \sqrt{\pi}. \tag{15.4}$$

We can take it on faith (which I'll do out of laziness), or we can bust open a can of complex analytical whoop ass on the problem and use residues. We have

$$(f * h)(x) = x\int_{\mathbb{R}} e^{-y^2} dy - \int_{\mathbb{R}} ye^{-y^2} dy \tag{15.5}$$

we see that $y\exp(-y^2)$ is odd, so its integral from $-\infty$ to $+\infty$ vanishes. We are left with

$$(f * h)(x) = x\int_{\mathbb{R}} e^{-y^2} dy \tag{15.6}$$

and this is necessarily

$$(f * h)(x) = x\sqrt{\pi} \tag{15.7}$$

which concludes our proof. □

Note

$$|(f * g)(x)| \leq \int |f(x-y)g(y)|dy \tag{15.8}$$

so if the integral converges absolutely, then the convolution exists. Some cases when the convolution is guarenteed:

convolution gauranteed when...

1. $f \in L^1(\mathbb{R})$, $|g(x)| \leq M$ for all $x \in \mathbb{R}$, the integral converges absolutely

2. $g \in L^1(\mathbb{R})$, $|f(x)| \leq M$ for all $x \in \mathbb{R}$, then the convlution exists

$$\int |f(x-y)g(y)|dy \leq M \int |f(x-y)|dy < +\infty \tag{15.9}$$

3. $f, g \in L^2(\mathbb{R})$, we have

$$\int |f(x-y)g(y)|dy \leq \left(\int |f(x-y)|^2 dy \right)^{1/2} \left(\int |g(y)|^2 dy \right)^{1/2} < +\infty \tag{15.10}$$

where we justify making it less than infinity by the properties of L^2.

There are many more times whene convolution is gauranteed.

But **WTF does the convolution mean anyways?**

Recall that the average of a function on $[a, b]$ is

$$avg(f) = \frac{1}{b-a} \int_a^b f(x)dx \tag{15.11}$$

We can generalize this to be a weighted average on $[a, b]$:

$$wt(f) = \frac{\int_a^b f(x)\omega(x)dx}{\int_a^b \omega(x)dx} \tag{15.12}$$

For our typical average, we usually just set $\omega(x) = 1$. So for

$$(f * g)(x) = \begin{pmatrix} \text{weighted average of } g \text{ around} \\ \text{the point } x \text{ with the weight} \\ \text{determined by the function } f, \\ \text{provided } f \text{ is normalized} \\ \int f(x)dx = 1 \end{pmatrix} \tag{15.13}$$

Alternatively, we can observe

$$(f * g)(x) = \int f(x-y)g(y)dy \tag{15.14}$$

so by plugging in the Riemann sum for the integral, we have

$$\int f(x-y)g(y)dy \approx \sum f(x-y_j)g(y_j)\Delta y_j. \tag{15.15}$$

The function $f(x - y_j)$ is just the function f translated along the x axis by the amount y_j, so the sum on the right is a linear combination of translates of f with coefficients $g(y_j)\Delta y_j$.

Example 15.2. Let

$$f = \begin{cases} 1 & \text{if } |x| \le a \\ 0 & \text{otherwise} \end{cases} \tag{15.16}$$

We normalize this to be

$$f = \begin{cases} \frac{1}{2a} & \text{if } |x| \le a \\ 0 & \text{otherwise} \end{cases} \tag{15.17}$$

So $\int_{-\infty}^{\infty} f(x)dx = 1$. Then for any g such that the convolution with f exists

$$(f * x)(x) = \int f(x-y)g(y)dy \tag{15.18a}$$

$$= \int_{|x-y|\le a} \frac{g(y)}{2a} dy \tag{15.18b}$$

$$= \int_{x-a} \frac{g(y)}{2a} dy \tag{15.18c}$$

$$= \text{average of } g \text{ on interval } [x-a, x+a] \tag{15.18d}$$

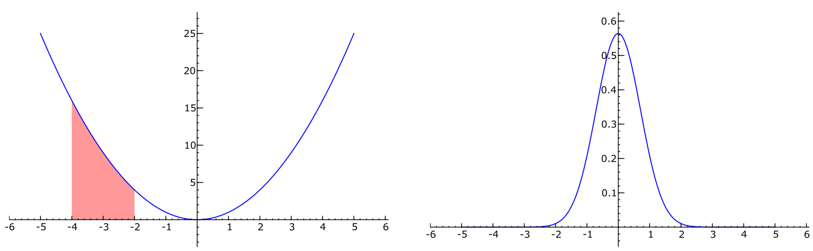

Figure 15.1: On the left, convolution with x^2; on the right, the Gaussian Kernel.

Example 15.3. Let $g(x) = x^2$, consider the convolution $(f*g)(x)$. Convolution is just sliding f along the x axis. You can see the integral as the red shaded region of figure (15.1).

Example 15.4. Let $f(x) = \exp(-x^2)$ and $g(x) = x^2$. Then we first normalize $f(x)$ by introducing $\widetilde{f}(x) = f(x)/\sqrt{\pi}$ so we see

$$\int \widetilde{f}(x)dx = 1. \tag{15.19}$$

This is a special function called the Gaussian Kernel *Gaussian Kernel*

$$\boxed{\widetilde{f}(x) = \frac{e^{-x^2}}{\sqrt{\pi}}} \tag{15.20}$$

For $|x| \geq 4$, $\widetilde{f}(x) \leq 6.4 \times 10^{-8}$. The values of $\widetilde{f}(x)$ that really matter are between $[x-4, x+4]$. One can see this reflected in the plot of the Gaussian in figure (15.1).

Scaling or Dilations

Suppose we shrink this down to $[x-\varepsilon, x+\varepsilon]$ for small ε by scaling or "dilating" f. We do this, given $\int f(x)dx = 1$, we define the **scaling** or **dilation** of f is given as

$$f_\varepsilon(x) = \frac{1}{\varepsilon} f\left(\frac{x}{\varepsilon}\right) \tag{15.21}$$

For $f(x) = \exp(-x^2)/\sqrt{\pi}$, we have

$$f_\varepsilon(x) = \frac{1}{\varepsilon\sqrt{\pi}} \exp(-(x/\varepsilon)^2). \tag{15.22}$$

As ε gets smaller, the peak gets higher, so $f_\varepsilon(x) \leq (6.4 \times 10^{-8})/\varepsilon$ for $|x| \geq 4\varepsilon$. What then is

$$(f_\varepsilon * g)(x) = \frac{1}{\varepsilon\sqrt{\pi}} \int \exp\left[-\left(\frac{x-y}{\varepsilon}\right)^2\right] g(y)dy \tag{15.23a}$$

$$\approx \frac{1}{\varepsilon\sqrt{\pi}} \int_{x-4\varepsilon}^{x+4\varepsilon} \exp\left[-\left(\frac{x-y}{\varepsilon}\right)^2\right] g(y)dy \tag{15.23b}$$

Theorem 15.1. Convolution has the same algebraic properties as multiplication

1. $f * (\alpha g + \beta h) = \alpha(f * g) + \beta(f * h)$
2. $f * g = g * f$ to prove this, do a change of variables

$$\begin{aligned}(f * g)(x) &= \int f(x-y)g(y)dy, \quad \text{let } z = x - y \\ &= \int f(z)g(x-z)dz \\ &= (g * f)(x)\end{aligned}$$

3. $f * (g * h) = (f * g) * h$

Theorem 15.2. (Convolution increases Smoothness) Suppose f is differentiable, $f * g$ and $f' * g$ are well defined (i.e. exist), then $f * g$ is differentiable and

$$(f * g)' = f' * g \tag{15.24}$$

Proof. Observe

$$(f * g)'(x) = \frac{d}{dx} \int f(x-y)g(y)dy \tag{15.25}$$

The integral converges for all x and x is not the variable of integration, so

$$\begin{aligned}\frac{d}{dx}\int f(x-y)g(y)dy &= \int \frac{d}{dx}\Big(f(x-y)g(y)\Big)dy \\ &= \int f'(x-y)g(y)dy \\ &= (f' * g)(x)\end{aligned}$$

which concludes our proof. $\square$

Implication: g may not have any derivatives, but if $g \in C^\infty(\mathbb{R})$, then we may form a new function $(f * g)$ which inherits all the smoothness of g, i.e. $(f * g) \in C^\infty$.

Three C^∞ functions

1. The Gaussian Kernel, which we have already seen,

$$G(x) = \frac{1}{\sqrt{\pi}} e^{-x^2} \tag{15.26}$$

with the conditions that $\int G(x)dx = 1$, it's even, and $G(x) \in C^\infty(\mathbb{R})$. Additionally, it's bounded!

2. The standard Cauchy distribution is another good choice

$$H(x) = \frac{1}{\pi}\frac{1}{1+x^2} \tag{15.27}$$

$\int H(x)dx = 1$, $H(x) \in C^\infty(\mathbb{R})$, $H(x)$ is even, it's bounded too.

3. The last is

$$K(x) = \begin{cases} \frac{1}{c} e^{-1/(1-x^2)} & \text{when } |x| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{15.28}$$

where $C = \int_{-1}^{1} \exp(-1/(1-x^2))dx$, $K(x) \in C^\infty(\mathbb{R})$, the only time we worry is at $x = \pm 1$. But it's clear that $\exp(-1/(1-x^2))$ vanishes for all derivatives of it. It's bounded, vanishes outside the interval from $[-1, 1]$.

Suppose $g \in L^1$, $\int g dx = 1$, dilates of g are — for any $\varepsilon > 0$ —

$$g_\varepsilon(x) = \frac{1}{\varepsilon} g\left(\frac{x}{\varepsilon}\right) \tag{15.29}$$

If $g'(x)$ exists, then

$$g'_\varepsilon(x) = \frac{1}{\varepsilon^2} g'\left(\frac{x}{\varepsilon}\right). \tag{15.30}$$

For the next theorem, let $\int g(x)dx = 1$,

$$\alpha = \int_{-\infty}^{0} g(x)dx$$

and

$$\beta = \int_0^\infty g(x)dx$$

so $\alpha + \beta = 1$. If g is even, then

$$\alpha = \beta = \frac{1}{2}.$$

Theorem 15.3. Suppose $g \in L^1$, and $\int g(x)dx = 1$. Suppose $f \in PC(\mathbb{R})$, and aso suppose either $|f(x)| \leq M$ for all $x \in \mathbb{R}$ or $g(x) = 0$ for x outside some finite interval, so $f * g$ is well defined. Then

$$\lim_{\varepsilon \to 0} (f * g)_\varepsilon(x) = \alpha f(x^+) + \beta f(x^-) \tag{15.31}$$

for all $x \in \mathbb{R}$. SO if x is continuous at x then $\lim_{\varepsilon \to 0}(f * g)_\varepsilon(x) = f(x)$. (This means that $|(f * g)_\varepsilon(x) - \alpha f(x^+) - \beta f(x^-)| < \delta$ for some $\delta > 0$ when ε is small enough; i.e. it's pointwise convergence.)

Alternatively (if additionally supposing that $|g(x)| \leq M$ for all x, and $f \in L^2$), then $(f * g)_\varepsilon \overset{(\varepsilon \to 0)}{\longrightarrow} f$ in norm.

16 Examples and Basic Properties of Fourier Transform

Let us consider a few examples computing the Fourier transform of a function. Remember the Fourier transform is

$$\widehat{f}(\xi) = \int f(x) e^{-i\xi x} dx. \tag{16.1}$$

Example 16.1. Let

$$\chi_a(x) = \begin{cases} 1 & \text{if } |x| < a \\ 0 & \text{otherwise} \end{cases} \tag{16.2}$$

be a characteristic function. Then

$$\widehat{\chi}_a(\xi) = \int_{-a}^{a} \xi_a(x) e^{-i\xi x} dx \tag{16.3a}$$

$$= \int_{-a}^{a} e^{-i\xi x} dx \tag{16.3b}$$

$$= \frac{i}{\xi}(e^{ia\xi} - e^{-ia\xi}) \tag{16.3c}$$

$$= -2\frac{\sin(a\xi)}{\xi}. \tag{16.3d}$$

Example 16.2. Consider

$$\mathcal{F}\left[e^{-ax^2/2}\right] = \sqrt{\frac{2\pi}{a}}e^{-\xi^2/2a} \tag{16.4}$$

where $\mathcal{F}[\cdot]$ is the fourier transform. We can see that

$$\begin{aligned}
\mathcal{F}\left[e^{-ax^2/2}\right] &= \int e^{-ax^2/2-i\xi x}dx && (16.5a)\\
&= \int e^{-y^2+(\xi^2/a)}\left(\frac{2}{\sqrt{a}}\right)dy \quad \text{where } y = x\sqrt{a/2}+i\xi/\sqrt{2a} && (16.5b)\\
&= \sqrt{\frac{2}{a}}e^{-\xi^2/2a}\int e^{-y^2}dy && (16.5c)\\
&= \sqrt{\frac{2\pi}{a}}e^{-\xi^2/2a}. && (16.5d)
\end{aligned}$$

Since we know that

$$\int e^{-x^2}dx = \sqrt{\pi}. \tag{16.6}$$

Now let us consider some of the basic properties of the Fourier transform.

1. (Shifting the Fourier Transform) For $a \in \mathbb{R}$,

$$\begin{aligned}
\mathcal{F}\left[f(x-a)\right] &= e^{-ia\xi}\mathcal{F}[f(x)]\\
&= e^{-a\xi}\widehat{f}(\xi)
\end{aligned}$$

2. (Dilation) For $\delta > 0$, $f_\delta(x) = f(x/\delta)/\delta$, then $\mathcal{F}(f_\delta) = \widehat{f}(\delta\xi)$. If $\delta > 1$, we're shrinking the width of f, but expanding the width of $\widehat{f}$. We also have

$$\mathcal{F}\left[f(\delta x)\right] = \widehat{f}_\delta(\xi) \tag{16.7}$$

3. If f is continuous, its first derivative $f'(x) \in PC(\mathbb{R})$ is piecewise continuous, and $f \in L^1$, then

Most important property of Fourier Transform!

$$\mathcal{F}\left[\frac{d}{dx}f(x)\right](\xi) = i\xi\widehat{f}(\xi). \tag{16.8}$$

This should be familiar, recall that fo the Fourier series we have f has coefficients c_n and f' has coefficients inc_n. On the other hand, if $xf(x) \in L^1$, then

$$\mathcal{F}\left[xf(x)\right](\xi) = i\frac{d}{d\xi}\widehat{f}(\xi). \tag{16.9}$$

4. If both $f, g \in L^1$, then

$$\mathcal{F}[f * g](\xi) = \widehat{f}(\xi)\widehat{g}(\xi) \tag{16.10}$$

The last property is the most important as it lets us change a given differential equation into an algebraic equation.

Proof. 1. (Shift) We see that

$$\begin{aligned}\mathcal{F}[f(x-a)] &= \int f(x-a)e^{-ix\xi}dx \\ &= \int f(y)e^{-i\xi(y+a)}dy \quad \text{where } y = x-a \\ &= e^{-ia\xi}\int f(y)e^{-i\xi y}dy \\ &= e^{-a\xi}\widehat{f}(\xi).\end{aligned}$$

2. (Dilation) Again by direct computation we see that

$$\begin{aligned}\mathcal{F}[f_\delta](\xi) &= \int \frac{1}{\delta}f\left(\frac{x}{\delta}\right)e^{-ix\xi}dx \\ &= \int f(y)e^{-i\delta y\xi}dy, \quad \text{with } y = x/\delta \\ &= \int f(y)e^{-iy(\delta\xi)}dy \\ &= \mathcal{F}[f(x)](\delta\xi) = \widehat{f}(\delta\xi)\end{aligned}$$

3. (Differentiation) Observe, once more by direct computation

$$\begin{aligned}\mathcal{F}\left[\frac{d}{dx}f(x)\right] &= \int f'(x)e^{-ix\xi}dx \\ &= f(x)e^{-ix\xi}|_{-\infty}^{\infty} - \int f(x)\frac{d}{dx}e^{-i\xi x}dx\end{aligned}$$

where we have just done integration by parts in the second step. We see that since $f \in L^1$ that $f(x) \to 0$ as $x \to \pm\infty$. So

$$\begin{aligned}f(x)e^{-ix\xi}|_{-\infty}^{\infty} - \int f(x)\frac{d}{dx}e^{-i\xi x}dx &= -\int f(x)\frac{d}{dx}e^{-i\xi x}dx \\ &= \int f(x)i\xi e^{-ix\xi}dx \\ &= i\xi\int f(x)e^{-ix\xi}dx \\ &= i\xi\widehat{f}(\xi)\end{aligned}$$

Similarly, we see that

$$xe^{-ix\xi} = i\frac{d}{d\xi}e^{-ix\xi} \tag{16.11}$$

so

$$\begin{aligned}
\mathcal{F}[xf(x)] &= \int xf(x)e^{-ix\xi}dx \\
&= \int \left(i\frac{d}{d\xi}e^{-i\xi x}\right) f(x)dx \\
&= i\frac{d}{d\xi}\int f(x)e^{-ix\xi}dx \\
&= i\frac{d}{d\xi}\widehat{f}(\xi)
\end{aligned}$$

4. If

$$\begin{aligned}
\mathcal{F}[f*g](\xi) &= \int \left(\int f(x-y)g(y)dy\right) e^{-ix\xi}dx \\
&= \int\int f(x-y)g(y)e^{-ix\xi}dxdy
\end{aligned}$$

Let $z = x - y$, then

$$\begin{aligned}
\int\int f(x-y)g(y)e^{-ix\xi}dxdy &= \int\int f(z)e^{-i\xi(z+y)}dzg(y)dy \\
&= \int \left(\int f(z)e^{-i\xi(z+y)}dz\right) g(y)dy \\
&= \int \left(\int f(z)e^{-i\xi z}e^{-iy\xi}dz\right) g(y)dy \\
&= \int \left(\int f(z)e^{-iz\xi}dz\right) g(y)e^{-iy\xi}dy \\
&= \int \widehat{f}(\xi)e^{-iy\xi}g(y)dy \\
&= \widehat{f}(\xi)\int g(y)e^{-iy\xi}dy \\
&= \widehat{f}(\xi)\widehat{g}(\xi)
\end{aligned}$$

which concludes our proof.

$\square$

Riemann-Lebesgue Theorem. If $f \in L^1$, then $\mathcal{F}[f](\xi) \to 0$ as $\xi \to \pm\infty$.

Sketch Of Proof. Intuitively we see

$$\begin{aligned}
\widehat{f}(\xi) &= \int e^{-i\xi x} f(x)dx \\
&= \int e^{-iz} f(z) \frac{dz}{\xi} \\
|\widehat{f}(\xi)| &\leq \int \left| f\left(\frac{z}{\xi}\right) \right| \frac{dz}{|\xi|} = \frac{1}{|\xi|} \int \left| f\left(\frac{z}{\xi}\right) \right| dz \\
&\leq \frac{1}{|\xi|} k
\end{aligned}$$

where k is a constant, since $f \in L^1$. So

$$\lim_{\xi \to \pm\infty} f(\xi) \leq \lim_{\xi \to \pm\infty} \frac{1}{|\xi|} k = 0. \tag{16.12}$$

This concludes our sketch of the proof.

What's the implication of the Riemann-Lebesgue lemma? Well, in addition to the property (3) of the Fourier transform, it implies the following:

If f is smooth, $\widehat{f}$ decays quickly. If $\widehat{f}$ decays fast, then f is smooth.

Example 16.3. If $f \in C^{(k-1)}$, $f^{(k)} \in PC(\mathbb{R})$ and $f^{(k)} \in L^1$, then we have

$$\mathcal{F}[f^{(k)}] = (i\xi)^k \widehat{f}(\xi). \tag{16.13}$$

The Riemann-Lebesgue lemma says $(i\xi)^k \widehat{f}(\xi) \to -$ as $|\xi| \to \infty$.

17 Inverse Fourier Transform, Fourier Transform on L^2

Lemma 17.1. If $f \in L^1$, then $\widehat{f}$ is continuous.

Proof. Observe by direct computation

$$\begin{aligned}
|\widehat{f}(\xi) - \widehat{f}(\eta)| &= \left| \int \left(e^{-i\xi x} - e^{-i\eta x}\right) f(x)dx \right| \\
&\leq 2 \int |f(x)|dx \quad \forall \eta, \xi \\
&\leq \int \left|e^{-i\xi x} - e^{-i\eta x}\right| |f(x)|dx
\end{aligned}$$

But as $\xi \to \eta$, $e^{-i\xi x} - e^{-i\eta x} \to 1 - 1 = 0$, so the quantity

$$\int \left|e^{-i\xi x} - e^{-i\eta x}\right| |f(x)|dx \to 0 \tag{17.1}$$

which concludes our proof. $\square$

The Fourier transform so far has been shown to be one way, mapping functions of x to functions of ξ. Can we go the other way? That is to say, is the Fourier transform invertible?

Inverse Fourier Transform

Fourier Inversion Theorem. Suppose $f \in L^1$ and $f \in PC(\mathbb{R})$, then

$$\lim_{\varepsilon \to 0} \frac{1}{2\pi} \int e^{i\xi x} e^{-\varepsilon^2 \xi^2/2} \widehat{f}(\xi) d\xi = \frac{1}{2} \left[f(x^+) + f(x^-) \right] \tag{17.2}$$

for all $x \in \mathbb{R}$. Moreover if both $f, \widehat{f} \in L^1$, then f is continuous, and

$$f(x) = \frac{1}{2\pi} \int e^{i\xi x} \widehat{f}(\xi) d\xi \tag{17.3}$$

for all $x \in \mathbb{R}$.

Proof. We have

$$\frac{1}{2\pi} \int e^{i\xi x} e^{-\varepsilon^2 \xi^2/2} \widehat{f}(\xi) d\xi = \frac{1}{2\pi} \int e^{i\xi x} e^{-\varepsilon^2 \xi^2/2} \int e^{-i\xi y} f(y) dy d\xi \tag{17.4}$$

We can interchange the variables of integration, thus

$$\begin{aligned}
\frac{1}{2\pi} \int e^{i\xi x} e^{-\varepsilon^2 \xi^2/2} \int e^{-i\xi y} f(y) dy d\xi &= \frac{1}{2\pi} \int \int e^{-i\xi(x-y)} e^{-\varepsilon^2 \xi^2/2} f(y) d\xi dy \\
&= \frac{1}{2\pi} \int f(y) \left[\int e^{-i\xi(y-x)} e^{-\varepsilon^2 \xi^2/2} \right] dy \\
&= \frac{1}{2\pi} \int f(y) \left[\sqrt{\frac{2\pi}{\varepsilon^2}} e^{-(y-x)^2/2\varepsilon^2} \right] dy \\
&= \frac{1}{\varepsilon \sqrt{2\pi}} \int f(y) e^{-(y-x)^2/2\varepsilon^2} dy \\
&= (f * \phi_\varepsilon)(x)
\end{aligned}$$

It turns out that $\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$ is the Gaussian. This function is even, and integral is evaluated to one. So it's C^∞. By theorem (15.3), $f * \phi_\varepsilon(x) \to \frac{1}{2}[f(x^-) + f(x^+)]$.

Suppose $f \in L^1$, then we need to show f is continuous. Since

$$|e^{i\xi x} e^{-\varepsilon^2 \xi^2/2} \widehat{f}(\xi)| \leq |\widehat{f}(\xi)| \tag{17.5}$$

for all ξ. We see that $|\exp(-\varepsilon^2 \xi^2/2)| < 1$ and $|\exp(i\xi x)| = 1$. So

$$\int e^{i\xi x} e^{-\varepsilon^2 \xi^2/2} \widehat{f}(\xi) d\xi \leq \int |\widehat{f}(\xi)| d\xi = \|\widehat{f}\|_{L^1} \tag{17.6}$$

for all $\varepsilon > 0$. So

$$\begin{aligned}
\mathcal{F}\{\widehat{f}\}(-x) &= \frac{1}{2\pi}\int e^{i\xi x}\widehat{f}(\xi)d\xi \\
&= \frac{1}{2\pi}\int e^{i\xi x}\left(\lim_{\varepsilon\to 0} e^{-\varepsilon^2\xi^2/2}\right)\widehat{f}(\xi)d\xi \\
&= \lim_{\varepsilon\to 0}\frac{1}{2\pi}\int e^{i\xi x}e^{-\varepsilon^2\xi^2/2}\widehat{f}(\xi)d\xi \\
&= \frac{1}{2}[f(x^+) + f(x^+)] \\
&= f(x)
\end{aligned}$$

where we justify the last step since f is continuous. So $\mathcal{F}\{\widehat{f}\}(-x)$ is continuous since $\widehat{f} \in L^1$, therefore f is continuous so we're done. $\square$

Remark 17.1. If $f \in L^1$ and $f \in PC$, then $\widehat{f}$ may not be in L^1. So we only have an approximation, which is the first part of the theorem, then

$$\underbrace{e^{-\varepsilon^2\xi^2/2}\widehat{f}(\xi)}_{\text{approx of } \widehat{f}} = \underbrace{\mathcal{F}\left\{f * \left(\frac{1}{\sqrt{2\pi}\varepsilon}e^{-(x/\varepsilon)^2/2}\right)\right\}}_{\text{approaches } f \text{ as } \varepsilon\to 0}$$

17.1 Consequences of the Inversion Theorem

1. If $\widehat{f} = \widehat{g}$, then $\widehat{F}\{f-g\} = \widehat{f} - \widehat{g} = 0$ which implies $f = g$. Thus the Fourier transform is unique. Further $\mathcal{F}^{-1}$ is well defined exactly by the inversion formula.

2. We have $\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$ in the inversion formula can be replaced by any (normalized) $g \in L^1$ with $\widehat{g} \in L^1$.

3. If $f \in L^1$, $\widehat{f} \in L^1$, then f is continuous and $f \in L^2$. So

$$\int |f(x)|^2 dx = \int |f(x)|\left|\int e^{i\xi x}\widehat{f}(\xi)d\xi\right| dx \tag{17.7}$$

but

$$\begin{aligned}
\left|\int e^{i\xi x}\widehat{f}(\xi)d\xi\right| &\le \int \left|e^{i\xi x}\widehat{f}(\xi)\right| d\xi \\
&\le \int |\widehat{f}(\xi)|d\xi
\end{aligned}$$

thus

$$\begin{aligned}
\int |f(x)|^2 dx &\le \int |f(x)||\widehat{f}(\xi)|d\xi dx \\
&\le \int |\widehat{f}(\xi)|d\xi \int |f(x)|dx.
\end{aligned}$$

17.2 Fourier Transform on L^2

Let $f \in L^2$, then $\widehat{f}(\xi) = \int \exp(-i\xi x) f(x) dx$ converges in L^2 norm. This exists for almost every ξ. Similarly

$$f(x) = \frac{1}{2\pi} \int e^{i\xi x} \widehat{f}(\xi) d\xi \tag{17.8}$$

for almost every x.

Plancherel Theorem. The Fourier Transform is an operator

$$\mathcal{F} : L^2 \to L^2 \tag{17.9}$$

and $\langle \widehat{f}, \widehat{g} \rangle = 2\pi \langle f, g \rangle$. If this is the case, this implies

$$\|\widehat{f}\|^2 = 2\pi \|f\|^2. \tag{17.10}$$

This is our Parseval identity in the continuous case.

Proof. We see by direct computation

$$\begin{aligned}
2\pi \langle f, g \rangle &= 2\pi \int f(x) \overline{g(x)} dx \\
&= \frac{2\pi}{2\pi} \int f(x) \overline{\int \widehat{g}(\xi) e^{i\xi x} d\xi} dx \\
&= \int f(x) \int \overline{\widehat{g}(\xi)} e^{-i\xi x} d\xi dx \\
&= \int \widehat{f}(\xi) \overline{\widehat{g}(\xi)} d\xi \\
&= \langle \widehat{f}, \widehat{g} \rangle
\end{aligned}$$

□

Example 17.1. Let

$$f(x) = \begin{cases} 1 & \text{when } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$g(x) = \begin{cases} |x| & \text{when } |x| \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Let us compute $\mathcal{F}[f * g]$.

Direct Computation: We see that we only really need to compute the convolution for $x \geq 0$ because both f and g are even functions, so $(f * g)$ is also

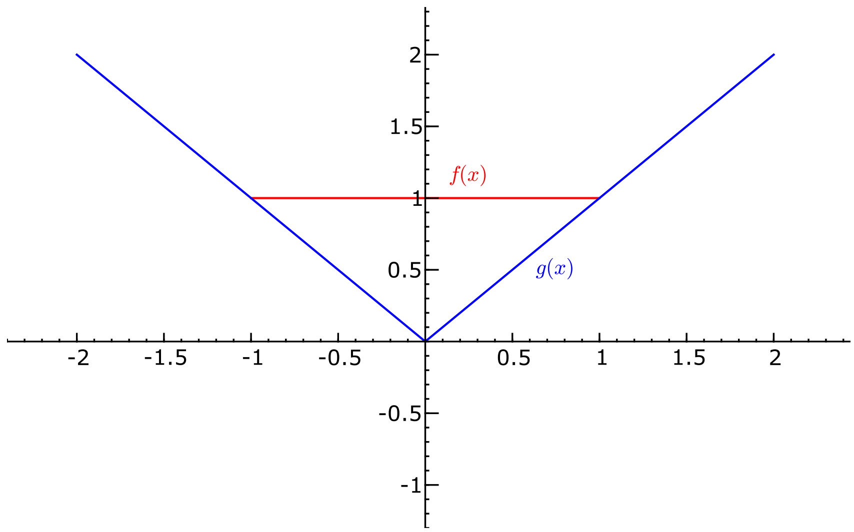

Figure 17.1: Plot of f and g

even. We can see this assertion by

$$\begin{aligned}(f*g)(-x) &= \int f(y)g(-x-y)dy, \quad \text{let } z=-y \\ &= \int f(-z)g(-x+z)dz \\ &= \int f(z)g(x-z)dz \\ &= (f*g)(x)\end{aligned}$$

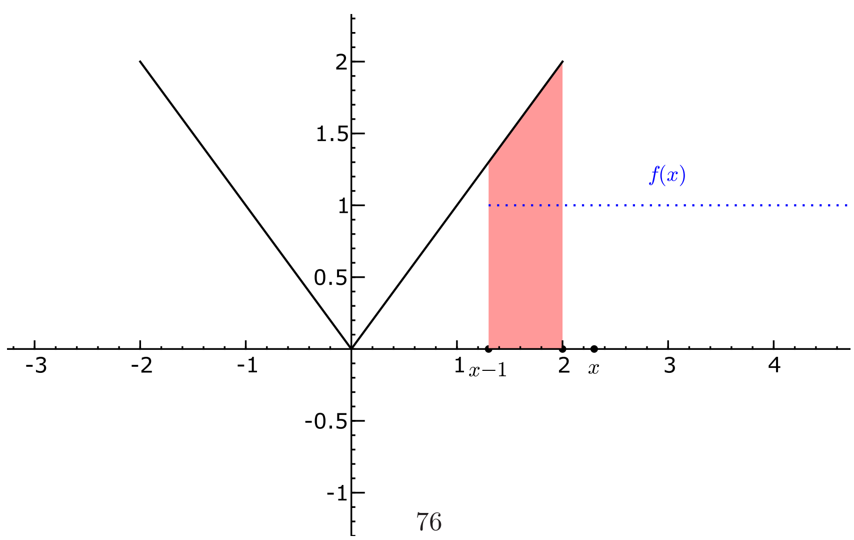

Figure 17.2: When $1 \le x \le 3$

which is justified by the even-ness of f and g. We also see that when $x > 3$, $(f * g)(x) = 0$. When $1 \leq x \leq 3$, we see that

$$(f * g)(x) = \int_{x-1}^{2} y dy = \frac{1}{2}(4 - (x-1)^2) \tag{17.11}$$

by direct computation. The integral is shown in figure (17.2).

For $0 \leq x \leq 1$, we have

$$\begin{aligned}
(f * g)(x) &= \int_{x-1}^{0} y dy + \int_{0}^{x+1} y dy && (17.12a)\\
&= \frac{-1}{2} y^2 \Big|_{x-1}^{0} + \frac{1}{2} y^2 \Big|_{0}^{x+1} && (17.12b)\\
&= x^2 + 1. && (17.12c)
\end{aligned}$$

The integral is doodled in figure (17.3).

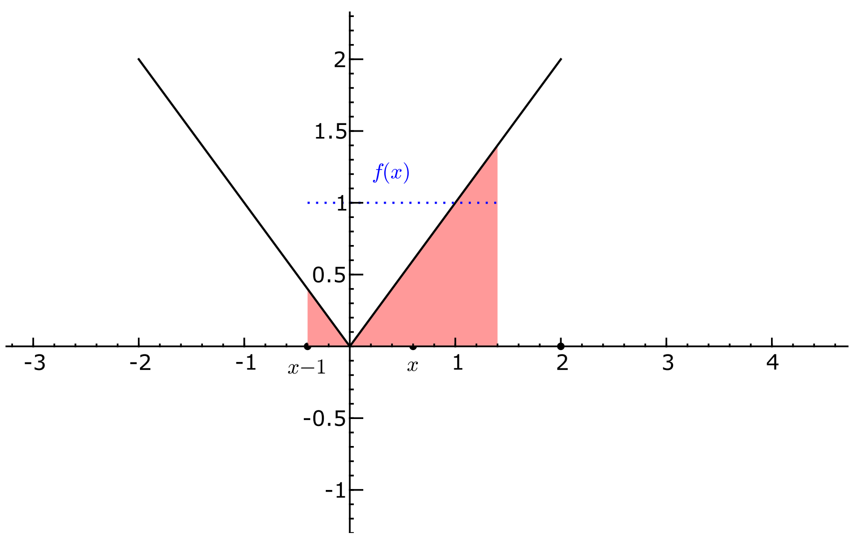

Figure 17.3: For $0 \leq x \leq 1$

We then find

$$(f * g)(x) = \begin{cases} x^2 + 1, & |x| \leq 1 \\ 2 - \frac{1}{2}(|x| - 1)^2, & 1 \leq |x| \leq 3 \\ 0 & \text{otherwise.} \end{cases} \tag{17.13}$$

The Fourier transform is then

$$\mathcal{F}[(f * g)] = \int_{-1}^{1} (x^2 + 1) e^{-i\xi x} dx + \int_{1 \leq |x| \leq 3} (2 - \frac{1}{2}(|x| - 1)^2) e^{-i\xi x} dx \tag{17.14}$$

This is a bit of a difficult thing to evaluate, so perhaps we shouldn't even try.

On the other hand, we can find

$$\mathcal{F}[f] = \widehat{f}(\xi) = 2\frac{\sin(\xi)}{\xi} \tag{17.15}$$

and similarly

$$\begin{aligned}
\int_0^A xe^{-i\xi x}dx &= \frac{ix}{\xi}e^{-i\xi x}\Big|_0^A - \frac{i}{\xi}\int_0^A e^{-i\xi x}dx && (17.16a)\\
&= \frac{iA}{\xi}e^{-iA\xi} - \frac{1}{\xi^2} + \frac{1}{\xi^2}e^{-iA\xi} && (17.16b)\\
&= \left(\frac{1}{\xi^2} + \frac{iA}{\xi}\right)e^{-iA\xi} - \frac{1}{\xi^2} && (17.16c)\\
\mathcal{F}[g] &= \int_0^2 xe^{-i\xi x}dx - \int_{-2}^0 xe^{-i\xi x}dx && (17.16d)\\
&= \int_0^2 xe^{-i\xi x}dx + \int_0^{-2} xe^{-i\xi x}dx && (17.16e)\\
&= \left(\frac{1}{\xi^2} + \frac{i2}{\xi}\right)e^{-i2\xi} + \left(\frac{1}{\xi^2} - \frac{i2}{\xi}\right)e^{i2\xi} - \frac{2}{\xi^2} && (17.16f)\\
&= \frac{2\cos(2\xi)}{\xi^2} + \frac{4\sin(2\xi)}{\xi} - \frac{2}{\xi^2} && (17.16g)
\end{aligned}$$

Thus

$$\mathcal{F}[f*g] = \widehat{f}(\xi)\widehat{g}(\xi) = \left(2\frac{\sin(\xi)}{\xi}\right)\left[\frac{2\cos(2\xi)}{\xi^2} + \frac{4\sin(2\xi)}{\xi} - \frac{2}{\xi^2}\right] \tag{17.17}$$

which would have been an impossible integral to perform if we did it the direct, naive way.

17.3 Applications of the Fourier Transform

Applications: Differential Equations

One of the most useful applications of the Fourier transform is to solve differential equations. Usually, it is on unbounded domains. What does this mean? Well, if the function has boundaries of $(-\infty,\infty)$, or $(-\infty,0)$, or $(0,\infty)$, then it's domain is unbounded.

Example 17.2. (Heat Equation) The heat equation is

$$\partial_t u(x,t) = k\partial_x^2 u(x,t), \quad -\infty \le x \le \infty. \tag{17.18}$$

There is no boundary condition for u but we have the initial condition $u(x,0) = f(x)$. We can take the Fourier transform in the x variable to find

$$\partial_t \widehat{u}(\xi,t) = k(i\xi)^2\widehat{u}(\xi,t) = -k\xi^2 u(\xi,t) \tag{17.19}$$

This is a first order differential equation in time, which has the solution

$$\widehat{u}(\xi,t) = \widehat{f}(\xi)e^{-k\xi^2 t} \tag{17.20}$$

If we define

$$K_t(x) = \mathcal{F}^{-1}[e^{-k\xi^2 t}] = \frac{1}{\sqrt{4\pi kt}}e^{-x^2/(4kt)} \tag{17.21}$$

(where we just used our relation of the Fourier transform of the Gaussian to justify this), then

$$\widehat{u}(\xi,t) = \mathcal{F}[f * K_t] \tag{17.22}$$

thus

$$u(x,t) = (f * K_t)(x) = \frac{1}{\sqrt{4\pi kt}}\int f(y)e^{-(x-y)^2/4kt}dy. \tag{17.23}$$

Example 17.3. (Wave Equation) The wave equation for an unbounded wave is

$$\partial_x^2 u(x,y) + \partial_y^2 u(x,y) = 0, \quad -\infty < x < \infty, \quad y > 0 \tag{17.24}$$

We would like to look for bounded solutions. So we take the Fourier transform in x and find

$$-\xi^2\widehat{u} + \partial_y^2\widehat{u} = 0 \tag{17.25}$$

The initial condition becomes $\widehat{u}(\xi,0) = \widehat{f}(\xi)$. So we've changed the wave equation into a second order ODE, the characteristic equation is

$$-\xi^2 + r^2 = 0 \Rightarrow r = \pm|\xi|. \tag{17.26}$$

So, the general solution is

$$\widehat{u}(\xi,y) = c_1(\xi)e^{|\xi|y} + c_2(\xi)e^{-|\xi|y} \tag{17.27}$$

Observe that as $\xi \to \pm\infty$, the first term goes to infinity. So, due to our demand of making u bounded, we need to set $c_1 = 0$. Thus our solution becomes (by imposing our initial condition)

$$\widehat{u}(\xi,y) = c_2(\xi)e^{-|\xi|y} \tag{17.28a}$$

$$= \widehat{f}(\xi)e^{-|\xi|y}. \tag{17.28b}$$

Once again, this is a Fourier transform of f convoluted with some function. We then write

$$\widehat{u}(\xi,y) = \widehat{f}(\xi)e^{-|\xi|y} = \mathcal{F}[f * P_y] \tag{17.29}$$

so we define

$$P_y(x) = \mathcal{F}^{-1}[e^{-|\xi|y}] \tag{17.30}$$

which is a special function. We call it the **Poisson Kernel** and we can write it out explicitly as

$$\boxed{P_y(x) = \frac{y}{\pi(x^2+y^2)}.} \tag{17.31}$$

We can now write the general solution of the wave equation as

$$u(x,y) = (f * P_y)(x) = \int \frac{yf(x-y)}{\pi(x^2+y^2)}dy. \tag{17.32}$$

This concludes our example.

Applications: Signal Analysis

The other application is signal analysis. We represent a signal as a function of time, $f(t)$, representing the amplitude of a signal at time t (e.g. a sound signal, $f(t)$ would be the volume of the sound, etc.).

By the Fourier inversion formula

$$f(t) = \frac{1}{2\pi}\int e^{i\omega t}\widehat{f}(\omega)d\omega \tag{17.33}$$

where $\widehat{f}(\omega)$ is the Fourier transform

$$\widehat{f}(\omega) = \int f(t)e^{-i\omega t}dt \tag{17.34}$$

where t is time, and ω is frequency; we have t in units of (e.g) seconds, and ω in units of Herz ("cycles per second").

The Fourier inversion represents $f(t)$ as a (continuous) superposition of periodic and simple waves. That is

$$e^{i\omega t} = \text{ periodic simple waves} \tag{17.35}$$

and

$$\mathcal{F}^{-1}[\widehat{f}(\omega)] = \begin{pmatrix} \text{representation of } f(t) \text{ as} \\ \text{continuous superposition of} \\ \text{simple periodic waves.} \end{pmatrix}. \tag{17.36}$$

We also model systems (e.g. an electrical system, or a telephone system) with an operator L

$$\begin{aligned} L :& f \to L[f] \\ & \text{input} \to \text{output} \end{aligned}$$

When L is linear, we have a linear system. If L is shift-invariant, then L commutes with translations. What do we mean by this? Well, if

$$L[f(x)] = g(x) \Rightarrow L[f(x+k)] = g(x+k) \tag{17.37}$$

for some arbitrary constant k, then an operator that shifts f by 1

$$E[f(x)] = f(x+1) \tag{17.38}$$

can be interchanged with L:

$$L[E[f(x)]] = E[L[f(x)]]. \tag{17.39}$$

For instance, the AM radio operator is a linear, shift invariant system.

Part III

Discrete Fourier Transforms

18 Sampling Theorem

Oftentimes in signal processing, we have to work with discrete packets instead of continuous signals. Because of this, we often use a discretized Fourier transform. We will introduce a few notions first before getting to the Discrete Fourier transform.

Definition 18.1. A signal $f(t)$ is **bandlimited** if $\widehat{f}(\omega)$ vanishes for all $|\omega| > \Omega$ where Ω is a constant called the **bandwidth**.

Sampling Theorem. Suppose $f \in L^2$ and $\widehat{f}(\omega) = 0$ for $|\omega| > \Omega$. Then

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{\Omega}\right) \frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}. \tag{18.1}$$

So we can have a function defined on the real line, and can be reconstructed from countably many values we know about it, despite *it having uncountably many values.*

We may interpret the sine function as the orthogonal basis in L^2 where $\{\widehat{f}(\omega) = 0 \text{ for } |\omega| > \Omega\}$ is a subspace of L^2 (i.e. it is closed under function addition and scalar multiplication).

Proof. Since $f \in L^2$, $\widehat{f} \in L^2(-\Omega, \Omega)$ so we may expand $\widehat{f}$ in a Fourier series! We have to extend $\widehat{f}$ to be periodic, and then look at the particular interval of interest $(-\Omega, \Omega)$. Now the Fourier series of the function is

$$\widehat{f}(\omega) = \sum_{n=-\infty}^{\infty} c_{-n} e^{-in\omega\pi/\Omega} \tag{18.2}$$

and further

$$\begin{aligned}
c_{-n} &= \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \widehat{f}(\omega) e^{in\omega\pi/\Omega} d\omega && (18.3a)\\
&= \frac{1}{2\Omega} \int_{-\infty}^{\infty} \widehat{f}(\omega) e^{in\omega\pi/\Omega} d\omega \text{ (extending from } [-\Omega, \Omega] \text{ to } \mathbb{R}) && (18.3b)\\
&= \frac{\pi}{\Omega} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\omega) e^{in\omega\pi/\Omega} d\omega \right) && (18.3c)\\
&= \frac{\pi}{\Omega} \left[f\left(\frac{n\pi}{\Omega}\right) \right] = \frac{\pi}{\Omega} f\left(\frac{n\pi}{\Omega}\right) && (18.3d)
\end{aligned}$$

which is really just the inverse Fourier Transform. Thus we have

$$\widehat{f}(\omega) = \sum_{n=-\infty}^{\infty} \frac{\pi}{\Omega} f\left(\frac{n\pi}{\Omega}\right) e^{-in\omega\pi/\Omega} \tag{18.4}$$

We then perform the inverse Fourier transform to get

$$\begin{aligned}
f(t) &= \frac{1}{2\pi}\int_{-\Omega}^{\Omega} \widehat{f}(\omega) e^{i\omega t} d\omega && \text{(18.5a)}\\
&= \frac{1}{2\Omega}\int_{-\Omega}^{\Omega}\left(\sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{\Omega}\right) e^{-in\omega\pi/\Omega}\right) e^{i\omega t} d\omega && \text{(18.5b)}\\
&= \frac{1}{2\Omega}\sum f\left(\frac{n\pi}{\Omega}\right)\int_{-\Omega}^{\Omega} e^{i\omega(t-n\pi/\Omega)} d\omega && \text{(18.5c)}\\
&= \frac{1}{2\Omega}\sum f\left(\frac{n\pi}{\Omega}\right)\Omega\left[\frac{e^{i(\Omega t-n\pi)} - e^{-i(\Omega t-n\pi)}}{i(\Omega t-n\pi)}\right] && \text{(18.5d)}\\
&= \sum f\left(\frac{n\pi}{\Omega}\right)\frac{\sin(\Omega t-n\pi)}{(\Omega t-n\pi)} && \text{(18.5e)}
\end{aligned}$$

This concludes our proof. □

NOTE! ★★★★★

Here we must emphasize

$$\boxed{\widehat{f}(\omega) = \sum \frac{\pi}{\Omega} f\left(\frac{n\pi}{\Omega}\right) e^{-in\omega\pi/\Omega}} \tag{18.6}$$

If we sample at N points equally distanced from each other, letting $t_n = n(\pi/\Omega)$, we have

$$\widehat{f}(\omega) = \frac{\pi}{\Omega}\sum f(t_n)e^{-it_n\omega} \tag{18.7}$$

We assume the sample points are periodic.

Dual Formulation For Sampling Frequency

Suppose we have a signal $f \in L^2$ such that $f(t) = 0$ for $|t| > L > 0$. Then we say the signal is **time limited**. Then

$$\widehat{f}(\omega) = \sum_{-\infty}^{\infty} \widehat{f}\left(\frac{n\pi}{L}\right)\frac{\sin(L\omega - n\pi)}{(L\omega - n\pi)}. \tag{18.8}$$

We then have a small modification to the Sampling Theorem.

Theorem 18.1. (Modified Sampling Theorem) Suppose we have $f \in L^2$ and $\widehat{f}(\omega) = 0$ for ω outside of $[a, b]$. Then

$$f(t) = \sum f\left(\frac{2\pi n}{b-a}\right) e^{-i\left(\frac{a+b}{b-a}\right)n\pi t}\left[\frac{\sin\left(\left(\frac{b-a}{2}\right)t - n\pi\right)}{\left(\frac{b-a}{2}\right)t - n\pi}\right] \tag{18.9}$$

19 Uncertainty Principle

There is one last significant principle that we will cover in signal processing – Heisenberg's famous uncertainty principle.

Definition 19.1. For $f \in L^2$, the **dispersion of** f **about a point** a is defined as

$$\Delta_a f = \frac{\int (x-a)^2 |f(x)|^2 dx}{\int |f(x)|^2 dx}. \tag{19.1}$$

This tells us how concentrated the function f is near the point a. The smaller it is, the more concentrated f is; the larger it is, the less concentrated f is.

Example 19.1. Consider the rectangle function $\chi_{1/2}(t)$ which is 1 if $t \in [-1/2, 1/2]$ and 0 otherwise. We can explicitly compute

$$\Delta_a \chi_{1/2} = \frac{\int_{-1/2}^{1/2} (x-a)^2 |\chi_{1/2}(x)|^2 dx}{\int_{-1/2}^{1/2} |\chi_{1/2}(x)|^2 dx} \tag{19.2a}$$

$$= \int_{-1/2}^{1/2} (x-a)^2 dx \tag{19.2b}$$

$$= \frac{1}{3}(x-a)^3 |_{x=-1/2}^{x=1/2} \tag{19.2c}$$

$$= a^2 + \frac{1}{12}. \tag{19.2d}$$

Note its smallest at $a = 0$ and it increases as $|a| \to \infty$.

Theorem 19.1. (Heisenberg's Uncertainty Principle) Given some $f \in L^2$, then

$$(\Delta_a f)\left(\Delta_\alpha \widehat{f}\right) \geq \frac{1}{4} \tag{19.3}$$

for any $a, \alpha \in \mathbb{R}$.

This amounts to nothing more than saying "For any signal $f \in L^2$, f *cannot* be *both* time limited and band limited."

References

[1] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics Part II.* McGraw-Hill Science, 1953.

[2] G. B. Folland, *Fourier Analysis and its Applications.* Brooks/Cole Publishing, 1992.

[3] W. Rudin, *Principles of Mathematical Analysis.* McGraw-Hill Science/Engineering/Math, third edition ed., 1976.

[4] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics Part I.* McGraw-Hill Science, 1953.

Index