Harmonic Analysis

Jacob Denson

December 5, 2017

Table Of Contents

Part I

Classical Fourier Analysis

Deep mathematical knowledge often arises hand in hand with the recognition of symmetry. Nowhere is this more clear than in the foundations of harmonic analysis, where we attempt to understand 'oscillating' mathematical objects, in their various forms. In the mid 18th century, problems in mathematical physics led D. Bernoulli, D'Alembert, Lagrange, and Euler to consider functions representable as a trigonometric series

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(nt) + b_n \sin(kt)$$

In 1811, Joseph Fourier has the audacity to announce that *all* functions were representable in this form. His conviction is the reason the classical theory of harmonic analysis is often named Fourier analysis, where we analyze the degree to which Fourier's proclamation holds. In the 1820s, Poisson, Cauchy, and Dirichlet all attempted to form rigorous proofs that 'Fourier summation' holds for all functions. Their work is responsible for most of the modern subject of analysis we know today. The biggest question we will ask is how we interpret the convergence of Fourier series. Pointwise convergence is not enough to justify most analytic techniques, and what's more, under pointwise convergence, the representation of a function by Fourier series need not be unique. Uniform convergence is useful, but too hopeful to obtain for all functions, and need not even hold for continuous functions. This means we must introduce more subtle methods like Feyér and Abel convergence.

Chapter 1

Springs, Strings, and Symmetry

1.1 The Wave Equation

Let's begin by taking a look at the problem which inspired Fourier and the mathematicians of his time to consider Fourier summation. Consider the physical problem of determining the motion of a spring undergoing simple harmonic motion, whose acceleration can be described by the differential equation $\ddot{x} = -k^2 x$ for some $k > 0$. It is well known that solutions of this equation are oscillating vibrations of the form

$$x = A\cos(kt) + B\sin(kt) = C\cos(kt + \phi)$$

where the two representations are connected by the trigonometric equality

$$\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$$

ϕ is known as the *phase* of the oscillation, C is the *amplitude*, and $k/2\pi$ the *frequency*. No serious effort was required on our part of produce these equations – they were known to Newton, and to Hooke before him, and require only the basic methods of the calculus. But this equation is very important; it forebodes that trigonometric functions will occur over and over again in the study of oscillatory behaviour. Instead, consider a tethered string vibrating under the influence of tension, whose motion is describable by the **wave equation**

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

where $u(t,x)$ models the motion of the spring over time and space. We will assume $x \in [0,\pi]$, and $t \in \mathbf{R}$. Why we normalize space to $[0,\pi]$ will become clear later on. There is an obvious connection between the dynamics of the spring and string; both describe motion under the effects of tension. What makes the string's motion tricky to analyze is that the motion is infinite dimensional; the physical state of the string at any particular time is described by a function $u(t,\cdot)$ on the real line, which consists of a specifying the position of infinitely many points. This makes the analysis of the wave equation much harder than the harmonic oscillator.

If you've seen a string vibrate, you'll notice that it follows a motion with an initial pattern which is perturbed back and forth vertically. These are standing waves, described by a motion of the form

$$u(t,x) = \psi(x)\nu(t)$$

where ν is periodic. Because of 'separating the variables' of the equation, determining all possible standing waves is much easier than determining all solutions to the wave equation. Any such equation which satisfies the wave equation must satisfy

$$\psi''(x)\nu(t) = \psi(x)\nu''(t)$$

or

$$\frac{\psi''(x)}{\psi(x)} = \frac{\nu''(t)}{\nu(t)}$$

Since the left side is independent of t, and the right side independent of x, the value the equations describe must be independent of both t and x, hence constant over the entire region to a value λ, where we obtain the equations

$$\psi''(x) = \lambda\psi(x) \qquad \nu''(t) = \lambda\nu(t)$$

We can assume $\lambda < 0$, for otherwise our standing wave solution will not oscillate. Thus we return to the solution of the spring equation and find

$$\psi(x) = a\cos(mx) + b\sin(mx) \qquad \nu(t) = a'\cos(mt) + b'\sin(mt)$$

where $m^2 = -\lambda$. Since $\psi(0) = \psi(\pi) = 0$, we have $a = 0$, , and $m \in \mathbf{Z}$. Our final expression can then be rewritten as

$$u(t,x) = \sin(mx)(A\cos(mt) + B\sin(mt)) = A\sin(mx)\cos(mt - \phi)$$

These are the harmonics. If you've ever learned to play music, these are the 'pure tones', which overlap to form an interesting and pleasant harmony.

It was Fourier who had the audacity to suggest that one could produce *all* solutions to the wave equation from these base tones. Since the wave equation is a *linear* partial differential equation, the set of solutions forms a linear class of functions, and we can therefore obtain a more complicated family of solutions to the wave equation of the form

$$u(t,x) = \sum_{m=1}^{n} \sin(mx)(A_m \cos(mt) + B_m \sin(mt))$$

Fourier said that these were *all* such solutions, provided we take $n \to \infty$, and consider an infinite series. Now given the initial conditions $u(0,x) = f(x)$, we find

$$f(x) = \sum_{m=0}^{\infty} A_m \sin(mx)$$

and given an initial velocity function $\partial_t u(0,x) = g(x)$, by performing a formal differentiation, we should have

$$g(x) = \sum_{m=0}^{\infty} mB_m \sin(mx)$$

Thus in order to find the constants A_m, B_m which give the motion of a string in terms of harmonic frequencies, it suffices to decompose an arbitrary one dimensional function on $[0,\pi]$ into the sum of sinusoidal functions of differing frequency. The first problem of Fourier analysis is the investigation of the limits of this method; How do we obtain the coefficients of the sum from the function itself, and how can we ensure these coefficients reflect the original function?

The question of obtaining the coefficients can be approached by a formal calculation. Suppose that a function f has an expansion

$$f(x) = \sum_{n=0}^{\infty} A_n \sin(nx)$$

Using the fact that the sin functions are orthogonal, in the sense that

$$\int_0^{\pi} \sin(mx)\sin(nx) = \begin{cases} 0 & m \neq n \\ \frac{\pi}{2} & m = n \end{cases}$$

We find that

$$\int_0^\pi f(x)\sin(mx)dx = \int_0^\pi \sum_{n=0}^\infty A_n \sin(nx)\sin(mx)$$
$$= \sum_{n=0}^\infty \int_0^\pi A_n \sin(nx)\sin(mx) = \frac{\pi}{2}A_m$$

Given any function $f : [0,\pi] \to \mathbf{R}$, a reasonable candidate for the coefficients is

$$A_n = \frac{2}{\pi}\int_0^\pi f(x)\sin(mx)$$

These values will be known as the **Fourier coefficients** of the function f.

1.2 The Heat Equation

Now we come to a quite different physical situation. Suppose we have a two-dimensional region D, with a heat distribution fixed on the buondary, upon which temperature fluctuates in the interior. The equation modelling the evolution of the heat distribution over time is the **heat equation**

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \Delta u$$

where Δu is the Laplacian operator. Here we are describing evolution over functions in two dimensions, which is also an infinite dimensional configuration.

To simplify again, we start by looking at only the steady state heat equations, those functions u satisfying $\Delta u = 0$, known as **harmonic functions**. Normally, we fix the boundary of a set C, and attempt to find a solution on the interior satisfying the boundary condition - physically, we fix a temperature on the boundary, wait for a long time, and see how the heat disperses on the interior. For now, let's consider functions on the unit disk $\mathbf{D}$, with fixed values on the boundary S^1. In this domain, we can switch to polar coordinates, in which the Laplacian operator takes the form

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}$$

We then apply the method of separation of coordinates. If $\Delta u = 0$, then

$$r^2 \frac{\partial^2 u}{\partial r^2} + r \frac{\partial u}{\partial r} = -\frac{\partial^2 u}{\partial \theta^2}$$

Writing $u(r,\theta) = f(r)g(\theta)$, the equation above reads

$$r^2 f''(r)g(\theta) + rf'(r)g(\theta) = -f(r)g''(\theta)$$

$$\frac{r^2 f''(r) + rf'(r)}{f(r)} = -\frac{g''(\theta)}{g(\theta)}$$

This means that both sides are equal to a constant λ^2 for some λ (If the constant value was negative, g wouldn't be periodic). Solving these equations tells us

$$g''(\theta) = -\lambda g(\theta) \qquad r^2 f''(r) + rf'(r) - \lambda f(r) = 0$$

Then we have

$$g(\theta) = A\cos(\lambda\theta) + B\sin(\lambda\theta)$$

Since g is 2π periodic, we require λ to be an integer m. The equation for f can be solved when $m \neq 0$ to be

$$f(r) = Ar^m + Br^{-m}$$

and for physical reasons, we force $f(r)$ to be bounded at zero, so $B = 0$, and we find the only solutions with separable variables are

$$u(r,\theta) = [A\cos(m\theta) + B\sin(m\theta)]r^m = A\cos(m\theta - \phi)r^m$$

When $m = 0$, the solution is just constant, because the solutions to $rf''(r) + f'(r) = 0$ are described by the equation $f(r) = A\log(r) + B$, which is unbounded near the origin unless $A = 0$. After our previous work, we would hope that all solutions are of the form

$$u(r,\theta) = C + \sum_{m=1}^{\infty} [A_m \cos(m\theta) + B_m \sin(m\theta)]r^m$$

If we know the values of u at $r = 1$, then we may apply the same expansion technique of the wave equation, except now we are trying to expand a functions on $[-\pi, \pi]$ in sines *and* cosines. Noticing that sines and cosines remain orthogonal under integration on $[-\pi, \pi]$, and that

$$\int_{-\pi}^{\pi} \sin^2(mt)\, dt = \int_{-\pi}^{\pi} \cos^2(mt)\, dt = \pi$$

the coefficents of the expansion should be

$$A_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(t)\, dt \qquad B_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(t)\, dt$$

If we begin with a function on $[0, \pi]$, and enlarge the domain to $[-\pi, \pi]$ by making the function odd, then the B_m all vanish, and we obtain the same expansion as on $[0, \pi]$. Since an arbitrary function f on $[-\pi, \pi]$ can be written as the sum of odd and even functions, expansion on $[-\pi, \pi]$ in terms of sin and cos is no more general than an expansion on $[0, \pi]$, and for our analysis, choosing either method is up to style. However, in the next section, we will introduce an even more elegant notation, applying complex exponentials, which will make the problem the simplest possible.

1.3 The Fundamental Oscillator

We are working with 2π-periodic functions $f : \mathbf{R} \to \mathbf{R}$, and attempting to decompose them into summations of sines and cosines. We now introduce a third object which encompasses sines and cosines together. First, define the circle group

$$\mathbf{T} = \{z \in \mathbf{C} : |z| = 1\}$$

Functions from $\mathbf{T}$ to $\mathbf{R}$ naturally correspond to 2π-periodic functions; given $g : \mathbf{T} \to \mathbf{R}$, the correspondence is given by the equation

$$f(t) = g(e^{it})$$

Thus, when defining 2π periodic functions, we shall make no distinction between a function 'defined in terms of t' and a function 'defined in terms of z', after making the explicit identification $z = e^{it}$. Then an expansion of the form

$$f(t) = \sum_{k=0}^{\infty} A_k \cos(kt) + \sum B_k \sin(kt)$$

using Euler's identity $e^{it} = \cos t + i \sin t$, leads to an expansion

$$\begin{aligned} f(z) &= \sum_{k=0}^{\infty} A_k \mathfrak{Re}[z^k] + B_k \mathfrak{Im}[z^k] \\ &= \sum_{k=0}^{\infty} A_k \left(\frac{z^k + z^{-k}}{2} \right) - i B_k \left(\frac{z^k - z^{-k}}{2} \right) = \sum_{k=-\infty}^{\infty} C_k z^k \end{aligned}$$

so a Fourier expansion on $[0, 2\pi]$ is really just a power series expansion on the circle in disguise. Thus expanding a real-valued function in e^{kit} is the same as expanding the function in terms of sines and cosines. The complex exponentials e^{kit} have the same orthogonality properties as sin and cos, so given a function f, the coefficients C_k can be found by the expansion

$$C_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-kit} dt$$

Thus a periodic function f gives rise to a function

$$\widehat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-kit} dt = ⨍_{-\pi}^{\pi} f(t) e^{-kit} \, dt$$

defined on **Z**, called the **Fourier series** of f, which measures the average value of f when it is 'twisted' by an oscillation with frequency k. If f is a complex-valued function on the circle group, we define the integral

$$\int_{\mathbf{T}} f(z) \, dz = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) \, dt$$

The notation for the Fourier transform then becomes

$$\widehat{f}(n) = \int_{S^1} f(z) z^{-n}$$

the most austere and elegant way to write the transform. It is the Fourier series representation in terms of complex exponentials which we will use for the rest of the book. We note, however, that no deep knowledge of the complex numbers was used here, or is used in any basic Fourier analysis – for most purposes, the exponential is just a simple way to represent sums of sines and cosines.

1.4 Basic Properties of Fourier Series

One of the most important properties of the Fourier series is that it asks nicely in terms of certain linear transformations of the functions it acts upon. This is summarized in this table of properties.

Theorem 1.1. *The Fourier series transforms in the following ways*

- *If f is real-valued, then $\widehat{f}(-n) = \overline{\widehat{f}(n)}$.*
- *If $(L_s f)(t) = f(t+s)$, then $\widehat{L_s f}(n) = e^{ins}\widehat{f}(n)$, so the frequencies of f are rotated by a factor which increases as the frequencies increase.*
- *We can shift frequencies by linear rotation. If we define $(M_m f)(t) = e^{mit} f$, then $\widehat{M_m f}(n) = \widehat{f}(n-m)$.*
- *If f is odd, then $\widehat{f}(-n) = -\widehat{f}(n)$. In particular, if f is real-valued, then the Fourier coefficients of f are purely imaginary, and*

  $$\widehat{f}(n) = -⨍_0^\pi f(t)\ \sin(nt)$$

- *If f is even, then $\widehat{f}(-n) = \widehat{f}(n)$, and if f is real-valued, then the Fourier ceofficients of f are real, and*

  $$\widehat{f}(n) = ⨍_0^\pi f(t)\ \cos(nt)$$

- *If we define $\overline{f}(x) = \overline{f(x)}$, then*

  $$\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}$$

- *If f is continuously differentiable, then $\widehat{f'}(n) = in\widehat{f}(n)$.*

These relations are all easy exercises in transforming integrals over $\mathbf{T}$, and are left to the reader to verify.

1.5 Examples of Expansions

Before we get to the real work, let's start by computing some Fourier series, to use as examples. We also illustrate the convergence properties of the series, which we shall look at in more detail later. The brunt of the calculation is left as an exercise.

Example. *Consider 'plucking' a string, by pinching a point p on a string of length π and moving it up. An equation modelling this type of configuration is*

$$f(x) = \begin{cases} \frac{x}{p} & : 0 \leqslant x \leqslant p \\ \frac{\pi - x}{\pi - p} & : p \leqslant x \leqslant \pi \end{cases}$$

To calculate Fourier coefficients, we extend f to be a 2π periodic odd function. Then, the Fourier coefficients are purely imaginary, with

$$\hat{f}(n) = \frac{-i}{\pi} \int_0^\pi f(x) \sin(nx) = \frac{-i}{\pi} \left[\frac{1}{p} \int_0^p x \sin(nx) + \frac{1}{\pi - p} \int_p^\pi (\pi - x) \sin(nx) \right]$$

Integration by parts tells us that

$$\int_0^p f(x) \sin(nx) = \frac{\sin(np)}{n^2} - \frac{\cos(np)p}{n}$$

$$\int_p^\pi (\pi - x) \sin(nx) = \frac{\cos(np)(\pi - p)}{n} + \frac{\sin(np)}{n^2}$$

Putting these together, we find

$$\widehat{f}(n) = -i \frac{\sin(np)}{n^2 p(\pi - p)}$$

and so, in some sense, we can obtain the identity

$$f(t) = \sum_{k=-\infty}^{\infty} -i \frac{\sin(np)}{n^2 p(\pi - p)} e^{kit} = \sum_{k=1}^{\infty} \frac{2\sin(np)}{p(\pi - p)} \sin(nt)$$

Referring back to our discussion of the wave equation, this means that if we let the string go, and let it perturb back and forth, the motion will be described by the infinite series

$$\sum_{n=1}^{\infty} \frac{2\sin(np)}{n^2 p(\pi - p)} \sin(nx) \cos(nt)$$

This converges absolutely and uniformly across time and space.

Example. *Consider the function f, defined on $[0,\pi]$ by $f(x) = x(\pi - x)$, made odd so that the function is defined on $[-\pi,\pi]$. The Fourier series is then purely imaginary, and in fact,*

$$-i\hat{f}(n) = \begin{cases} 4(\pi n^3)^{-1} & n \text{ odd} \\ 0 & n \text{ even} \end{cases}$$

and we may write

$$f(x) \sim \sum_{n \text{ odd}} \frac{4i}{\pi n^3}[e^{nix} - e^{-nix}] = \sum_{n \text{ odd}} \frac{8}{\pi n^3} \sin(nx)$$

This sum converges absolutely and uniformly on the entire real line.

Example. *The tent function*

$$f(x) = \begin{cases} 1 - \frac{|x|}{\delta} & : |x| < \delta \\ 0 & : |x| \geqslant \delta \end{cases}$$

is even, and therefore has a purely real Fourier expansion

$$\hat{f}(0) = \frac{\delta}{2\pi} \qquad \hat{f}(n) = \frac{1 - \cos(n\delta)}{\delta \pi n^2}$$

so we obtain an expansion

$$f(x) \sim \frac{\delta}{2\pi} + \sum_{n \neq 0} \frac{1 - \cos(n\delta)}{\delta \pi n^2} e^{inx} = \frac{\delta}{2\pi} + 2\sum_{n=1}^{\infty} \frac{1 - \cos(n\delta)}{\delta \pi n^2} \cos(nx)$$

This sum also converges absolutely and uniformly.

Example. *Consider the characteristic function*

$$\chi_{(a,b)}(x) = \begin{cases} 1 & : x \in (a,b) \\ 0 & : x \notin (a,b) \end{cases}$$

Then

$$\widehat{\chi}_{(a,b)}(n) = \frac{1}{2\pi} \int_a^b e^{-inx} = \frac{e^{-ina} - e^{-inb}}{2\pi i n}$$

Hence we may write

$$\chi_{(a,b)}(x) = \frac{b-a}{2\pi} + \sum_{n\neq 0} \frac{e^{-ina}-e^{-inb}}{2\pi in} e^{inx}$$
$$= \frac{b-a}{2\pi} + \sum_{n=1}^{\infty} \frac{\sin(nb)-\sin(na)}{\pi n}\cos(nx) + \frac{\cos(na)-\cos(nb)}{\pi n}\sin(nx)$$

This sum does not converge absolutely for any value of x (except when a and b are chosen trivially). To see this, note that

$$\left|\frac{e^{-inb}-e^{-ina}}{2\pi n}\right| = \left|\frac{1-e^{in(b-a)}}{2\pi n}\right| \geqslant \left|\frac{\sin(n(b-a))}{2\pi n}\right|$$

so that it suffices to show $\sum |\sin(nx)| n^{-1} = \infty$ for every $x \notin \pi\mathbf{Z}$. This follows because enough of the values of $|\sin(nx)|$ are large, so that we may apply the divergence of $\sum n^{-1}$ become applicable. First, assume $x \in (0, \pi/2)$. If

$$m\pi - x/2 < nx < m\pi + x/2$$

for some $m \in \mathbf{Z}$, then

$$m\pi + x/2 < (n+1)x < m\pi + 3x/2 < (m+1)\pi - x/2$$

so that if $nx \in (-x/2, x/2) + \pi\mathbf{Z}$, $(n+1)x \notin (-x/2, x/2) + \pi\mathbf{Z}$. For y outside of $(-x/2, x/2) + \pi\mathbf{Z}$, we have $|\sin(y)| > |\sin(x/2)|$, and therefore for any n,

$$\frac{|sin(nx)|}{n} + \frac{|\sin((n+1)x)|}{n+1} > \frac{|\sin(x/2)|}{n+1}$$

and thus

$$\sum_{n=1}^{\infty} \frac{|\sin(nx)|}{n} = \sum_{n=1}^{\infty} \frac{|\sin(2nx)|}{2n} + \frac{|\sin((2n+1)x)|}{2n+1}$$
$$> |\sin(x/2)| \sum_{n=1}^{\infty} \frac{1}{2n+1} = \infty$$

In general, we may replace x with $x - k\pi$, with no effect to the values of the sum, so we may assume $0 < x < \pi$. If $\pi/2 < x < \pi$, then

$$\sin(nx) = \sin(n(\pi - x))$$

and $0 < \pi - x < \pi/2$, completing the proof, except when $x = \pi$, in which case

$$\sum_{n=1}^{\infty} \left| \frac{1 - e^{in\pi}}{2\pi n} \right| = \sum_{n \text{ even}} \left| \frac{1}{\pi n} \right| = \infty$$

Thus the convergence of a Fourier series need not be absolute.

Example. *We can often find formulas for certain fourier summations from taking the corresponding power series. For instance, the power series expansion*

$$\log\left(\frac{1}{1-x}\right) = \sum_{k=1}^{\infty} \frac{z^k}{k}$$

which converges pointwise for every $z \in \mathbf{D}$ but $z = 1$, implies that for $x \notin 2\pi\mathbf{Z}$,

$$\sum_{k=1}^{\infty} \frac{\cos(kx)}{k} = \mathfrak{Re}\left(\log\left(\frac{1}{1-e^{ix}}\right)\right) = -\frac{1}{2}\log(2 - 2\cos(x))$$
$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \mathfrak{Im}\left(\log\left(\frac{1}{1-e^{ix}}\right)\right) = \arctan\left(\frac{\sin(x)}{1-\cos(x)}\right)$$

where we agree that $\arctan(\pm\infty) = \pm\pi/2$. *If a power series' radius of convergence exceeds 1, then it is likely that the corresponding Fourier series taken on the circle will be pleasant, whereas if the power series' radius is equal to 1, we can expect nasty behaviour on the boundary. In Complex analysis, one avoids talking about the boundary of the holomorphic function's definition, whereas in Fourier analysis we have to embrace the boundary points, which makes the theory a little more pathological.*

Example. *If f is a* **trigonometric polynomial***, meaning there are coefficients a_n such that*

$$f(x) = \sum_{n=-N}^{N} a_n e^{nit}$$

then it is easy to see that $\widehat{f}(n) = a_n$. In particular, we will be interested in the analysis of the **Dirichlet kernel**

$$D_N(x) = \sum_{n=-N}^{N} e^{nit}$$

which assigns a unit mass to each integer frequency between $-N$ and N. By the geometric series summation formula, we may write the Dirichlet kernel as

$$\begin{aligned}
1+\sum_{n=1}^{N} e^{nit}+e^{-nit} &= 1+e^{it}\frac{e^{Nit}-1}{e^{it}-1}+e^{-it}\frac{e^{-Nit}-1}{e^{-it}-1}\\
&= 1+e^{it}\frac{e^{Nit}-1}{e^{it}-1}+\frac{e^{-Nit}-1}{1-e^{it}} = \frac{e^{(N+1)it}-e^{-Nit}}{e^{it}-1}\\
&= \frac{e^{(N+1/2)it - e^{-(N+1/2)it}}}{e^{it/2}-e^{-it/2}} = \frac{\sin((N+1/2)t)}{\sin(x/2)}
\end{aligned}$$

D_N has average value $1/2\pi$ on $[-\pi,\pi]$, but the average value of $|D_N|$ becomes very large as N tends to ∞. We thus see that D_N has large oscillation, which causes D_N to be small but $|D_N|$ to be very large.

Example. *The* **Poisson kernel** *P_r is defined on $[-\pi,\pi]$, and for $0 \leqslant r < 1$, by the power series*

$$P_r(t) = \sum_{n=-\infty}^{\infty} r^{|n|}e^{int}$$

P_r arises in the discussion of heat distributions on the unit disk, because

$$P_r(t) = 1 + 2\sum_{n=1}^{\infty} r^n \cos(nt)$$

so P_r 'fits the mold' we gave of solutions to the heat equation, and in some sense, it is the fundamental solution, because all solutions can be rewritten in the form

$$\sum a_n r^{|n|}e^{int}$$

for some coefficients a_n, which is P_r with certain frequencies amplified. For $r < 1$, the sum defining P_r converges uniformly on the disk, and as such we obtain that $\widehat{P_r}(n) = r^{|n|}$, so we see that the high frequency parts of P_r decrease geometrically. We can also apply an (infinite) geometric series summation to obtain that

$$\begin{aligned}
\sum r^{|n|}e^{int} &= 1+\frac{re^{it}}{1-re^{it}}+\frac{re^{-it}}{1-re^{-it}} = 1+\frac{2r\cos t-2r^2}{(1-re^{it})(1-re^{-it})}\\
&= 1+\frac{2r\cos t-2r^2}{1-2r\cos t+r^2} = \frac{1-r^2}{1-2r\cos t+r^2}
\end{aligned}$$

As $r \to 1$, the function concentrates at the origin, and there does not appear to be too much oscillation in the function, reflecting the fact that Poisson kernel is much better behaved and the Dirichlet kernel.

Chapter 2

Fourier Series Convergence

Let's focus in on the problem we introduced in the last chapter. For each function $f : \mathbf{T} \to \mathbf{C}$ (which we assume from now on to be Riemann integrable), we have an associated **formal trigonometric series**

$$\sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{inx}$$

In some sense, f should be able to be approximated by the finite degree trigonometric polynomial

$$\sum_{n=-N}^{N} \widehat{f}(n) e^{inx}$$

At this point, we haven't deduced any reason for these sums to converge to f analytically. To understand the convergence, we define the m'th partial sum

$$S_m(f)(x) = \sum_{n=-m}^{m} \hat{f}(n) e^{inx}$$

The reason we have symmetry on both hand sides is so that, when f is real valued, $S_m(f)$ will just be a sum of cosines and sines. The first relation we can expect is pointwise convergence; is it true that for every x,

$$\lim_{m \to \infty} S_m(f)(x) = f(x)$$

Perhaps if we're lucky, we'll get uniform convergence as well. Unfortunately, we will show that there are even examples of continuous periodic functions whose partial sums diverge somewhere, so we must search for more exotic methods of convergence.

2.1 Do Functions Have Unique Series?

If the Fourier series of every function converged pointwise, we could conclude that if f and g have the same fourier coefficients, they must necessarily be equal. This is clearly not true, for if we alter a function at a point, the fourier series, defined by averaging over the entire region, remains the same. Nonetheless, if a function is continuous editing the function at a point will break continuity, so we may have some hope of uniqueness of the expansion.

Theorem 2.1. *If $\hat{f}(n) = 0$ for all n, then f vanishes wherever it is continuous.*

Proof. We shall prove this for real-valued functions. For every trigonometric polynomial $P(x) = \sum a_n e^{-nix}$, we have

$$\int_{-\pi}^{\pi} f(x)P(x)dx = 2\pi \sum a_n \widehat{h}(n) = 0$$

Suppose that f is continuous at zero, and assume without loss of generality that $f(0) > 0$. Pick δ such that if $|x| < \delta$, $|f(x)| > f(0)/2$. Consider the trigonometric polynomial

$$P(x) = \varepsilon + \cos x = \varepsilon + \frac{e^{ix} + e^{-ix}}{2}$$

where ε is small enough that $P(x) > A > 1$ for $|x| < \delta/2$, $P(x) > 0$ for $\delta/2 \leqslant |x| < \delta$, and $P(x) < B < 1$ for $|x| \geqslant \delta$. Consider the series of trigonometric polynomials

$$P_n(x) = (\varepsilon + \cos x)^n$$

For which we have

$$\left| \int_{-\pi}^{\pi} P_n(x) f(x) dx \right| \geqslant \int_{|x|<\delta} P_n(x) f(x) dx - \left| \int_{\delta \leqslant |x|} P_n(x) f(x) dx \right|$$

Now

$$\left| \int_{\delta \leqslant |x|} P_n(x) f(x) dx \right| \leqslant \int_{\delta \leqslant |x|} B^n |f(x)| \, dx \leqslant B^n \|f\|_1$$

whereas

$$\int_{|x|<\delta} P_n(x)f(x)dx = \int_{|x|<\delta/2} P_n(x)f(x) + \int_{\delta/2\leqslant|x|<\delta} P_n(x)f(x)$$
$$\geqslant \int_{|x|<\delta/2} P_n(x)f(x) \geqslant \delta A^n f(0)$$

and so we conclude

$$0 = \left| \int_{-\pi}^{\pi} P_n(x)f(x)dx \right| \geqslant \delta A^n f(0) - B^n \|f\|_1$$

Regardless of the values of $\|f\|_1, f(0)$, and δ, eventually, we find that the right hand side of the equation is positive, which is impossible. In general, if f is complex valued, then we may write $f = u + iv$, where

$$u(x) = \frac{f(x) + \overline{f(x)}}{2} \quad v(x) = \frac{f(x) - \overline{f(x)}}{2i}$$

The Fourier coefficients of $\overline{f}$ all vanish, because the coefficients of f vanish, and so we conclude the coefficients of u and v vanish. f is continuous at x if and only if it is continuous at u and v, and we know from the previous case this means that both u and v vanish at that point. □

Corollary 2.2. *If $\widehat{f} = \widehat{g}$, where f and g are continuous, then $f = g$.*

Proof. Because $f - g$ is continuous with vanishing Fourier coefficients. □

Later, we will see that this theorem can be generalized to not-necessarily continuous functions. Then the two Fourier series agree if and only if the two functions agree except on a 'small' set, that is, a set of measure zero.

Corollary 2.3. *If a continuous function f has absolutely convergent Fourier coefficients, then it's Fourier series converges uniformly to f.*

Proof. If $\sum |\widehat{f}(n)| < \infty$, then the functions $S_m(f)$ converge uniformly to a function g, which necessarily must be continuous. We may apply uniform convergence to conclude

$$\hat{g}(n) = \lim_{m\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} S_m(f)(t)e^{-int} = \hat{f}(n)$$

Hence $\hat{f} = \hat{g}$, so $f = g$. □

2.2 Convergence In $C^2(\mathbf{T})$

The easiest place to verify convergence is in the space of continuously differentiable functions. It turns out that differentiability gives the Fourier series a sort of 'decay' property, both over the torus and in more general contexts, which make the convergence properties easy to verify.

Theorem 2.4. *If $f \in C^2(\mathbf{T})$, then $\widehat{f}(n) = O(n^{-2})$, and so f's Fourier series converges uniformly to f.*

Proof. We know that $\widehat{f''}(n) = in\widehat{f}'(n) = -n^2\widehat{f}(n)$. Since we have the elementary estimate

$$|\widehat{f}''(n)| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f''(t)e^{-int} \right| \leqslant \frac{\|f''\|_1}{2\pi}$$

This means $\widehat{f}(n) = O(n^{-2})$. Every series that is $O(n^{-2})$ converges absolutely by the comparison test, and so the corresponding Fourier series converges uniformly. But this means that it converges to a continuous function, and so it must converge to f. $\square$

More generally, if $f \in C^k(\mathbf{T})$, then $\widehat{f}(n) = O(n^{-k})$. Later on, we will show that we obtain uniform convergence even in $C^1(\mathbf{T})$, and even if f is just Hölder continuous for $\alpha > 1/2$, in the sense that there is a constant C such that for any t, s,

$$|f(t) - f(s)| \leqslant C|t - s|^\alpha$$

which is slightly stronger than mere continuity.

2.3 Convolution and Kernel Methods

The notion of the convolution of two functions f and g is a key tool in Fourier analysis, both as a way to regularize functions, and as an operator that transforms nicely when we take Fourier series. Given two functions f and g, we define

$$(f * g)(t) = \fint_{-\pi}^{\pi} f(s)g(t - s)$$

which we can think of smoothing g by taking it's averages in a neighbourhood of a point, according to a distribution with density f, and this is indeed the case if f is positive and integrates to 2π. The first of the following properties for convolution is an easy exercise in integral transformations.

Theorem 2.5. *Convolution has the following properties:*

- *Convolution is a commutative, associative, bilinear operation.*
- $f * g$ *is continuous*
- $\widehat{f * g} = \widehat{f}\widehat{g}$.
- *If* f *is in* $C^1(\mathbf{T})$, *then* $f * g$ *is* C^1, *and* $(f * g)' = f' * g$. *In particular, if* f *is* $C^n(\mathbf{T})$, *and* g *is in* $C^m(\mathbf{T})$, *then* $f * g$ *is in* $C^{n+m}(\mathbf{T})$, *so convolution is 'additively smoothing'.*

Proof. Assume that f and g are continuous functions on $\mathbf{T}$. Then

$$[(f * g)(t) - (f * g)(s)] = \frac{1}{2\pi} \int f(u)[g(t-u) - g(s-u)]$$

Since $\mathbf{T}$ is compact, g is uniformly continuous, and so we may find δ such that if $|t-s| < \delta$, then $|g(t-u)-g(s-u)| < \varepsilon$. But then the integral above is bounded by $(2\pi)^{-1}\varepsilon\|f\|_1$, and this gives the continuity of convolution. If f and g are general Riemann integrable functions, then there exists continuous functions f_k, and g_k with $\|f_k\|_\infty \leqslant \|f\|_\infty$, $\|g_k\|_\infty \leqslant \|g\|_\infty$, which converge to f and g in L^1. But then

$$(f * g) - (f_k * g_k) = (f - f_k) * g + f_k * (g - g_k)$$

Now

$$|((f - f_k) * g)(t)| = \frac{1}{2\pi} \left| \int (f - f_k)(s) g(t-s) \right| \leqslant \frac{1}{2\pi} \|g\|_\infty \|f - f_k\|_1$$

$$|f_k * (g - g_k)| = \frac{1}{2\pi} \left| \int f_k(t-s)(g - g_k)(s) \right| \leqslant \frac{1}{2\pi} \|f_k\|_\infty \|g - g_k\|_1 \leqslant \frac{1}{2\pi} \|f\|_\infty \|g - g_k\|_1$$

so we see $f_k * g_k$ converges uniformly to $f * g$. But since each $f_k * g_k$ is continuous, this means $f * g$ must also be continuous.

To obtain the product identity for the Fourier series, we can apply Fubini's theorem to write

$$\begin{aligned}
\widehat{f * g}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f * g)(t) e^{-nit}\, dt \\
&= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(s) g(t-s) e^{-nit}\, ds\, dt \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \int_{-\pi}^{\pi} (L_{-s} g)(t) e^{-nit}\, dt\, ds \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) e^{-ins} \widehat{g}(n)\, ds \\
&= \widehat{f}(n) \widehat{g}(n)
\end{aligned}$$

and this is exactly the identity required.

The fact that $f * g$ is differentiable, and that the derivative is $f' * g$ when f is differentiable, follows because, using the mean value theorem, that the partial sums

$$\frac{f(x+h) - f(x)}{h}$$

converge uniformly in x to $f'(x)$, hence

$$(f * g)'(x) = \lim_{h \to 0} \int \frac{f(x+h-y) - f(x-y)}{h} g(x) = \int f'(x-y) g(x) = (f' * g)$$

and this is the required property. $\square$

We know that suitably smooth functions have convergent Fourier series. If we want to establish the converence of the series corresponding to a function f, we might want to 'smooth' f by convolving it with a function g. Provided that $\widehat{g}$ is 'close' to 1, the above theorem says that $\widehat{f * g}$ will be close to $\widehat{g}$. If we can establish the convergence properties on the convolution $f * g$, then we can probably obtain results about f. The family of functions g we will consider for which $\widehat{g}$ approximates 1 are called **good kernels**. In particular, a good kernel is a sequence of functions K_n on $\mathbf{T}$, for which

$$⨍_{-\pi}^{\pi} K_n(t) = 1$$

the functions K_n are bounded in L^1, and for any $\delta > 0$,

$$\int_{\delta < |x|} |K_n(x)| \to 0$$

so the functions K_n become concentrated at the origin. The L^1 boundedness follows from the first condition if we assume the K_n are positive. As $n \to \infty$, we find that for any m, the fact that K_n vanishes outside of small neighbourhoods implies that

$$\widehat{K_n}(m) = \frac{1}{2\pi} \int K_n(x) e^{-mit} \to 1$$

so if f is any function, then the Fourier series of $f * K_n$ converges to the Fourier series of f pointwise. Even better, $f * K_n$ approximates f, because the 'averages' of f with respect to K_n become concentrated around f.

Theorem 2.6. *If f is any function, then $(f * K_n)(x) \to f(x)$ whenever f is continuous at x.*

Proof. We have

$$[(f * K_n)(x) - f(x)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x-y) - f(x)] K_n(y)\, dy$$

Now fix δ such that $|f(x-y) - f(x)| < \varepsilon$ for $|y| < \varepsilon$. Then

$$\begin{aligned}
|(f * K_n)(x) - f(x)| &= \frac{1}{2\pi} \int_{|y| < \varepsilon} |f(x-y) - f(x)||K_n(y)|\, dy \\
&\quad + \frac{1}{2\pi} \int_{|y| \geqslant \varepsilon} |f(x-y) - f(x)||K_n(y)|\, dy \\
&\leqslant \frac{\varepsilon}{2\pi} \|K_n\|_1 + \frac{2\|f\|_\infty}{2\pi} \int_{|y| \geqslant \varepsilon} |K_n(y)|\, dy \\
&\leqslant \frac{\varepsilon}{2\pi} \|K_n\|_1 + o(1)
\end{aligned}$$

Taking δ smaller and smaller, we can make the two values as close as we desire, so we obtain pointwise convergence. Note that if f is continuous everywhere, then it is uniformly continuous, hence this bound is uniform, and $f * K_n$ therefore converges to f uniformly. $\square$

Recall the definition of the Dirichlet kernel

$$D_N(t) = \sum_{n=-N}^{N} e^{nit}$$

Because of the properties of convolution, we can represent the partial Fourier approximations of any function f by convolution with the Dirichlet kernel, in the sense that $S_n(f) = f * D_n$. If D_n was a good kernel, then we would obtain that the partial sums of S_n converge uniformly. This initially seems a good strategy, because $⨍ D_N(t) = 1$. However, we find

$$\begin{aligned}
\int_{-\pi}^{\pi} |D_n(t)| = \int_{-\pi}^{\pi} \left| \frac{\sin((N+1/2)t)}{\sin(t/2)} \right| &= 2\int_0^{\pi} \frac{|\sin((N+1/2)t)|}{\sin(t/2)} \\
\geqslant 4\int_0^{\pi} \frac{|\sin((N+1/2)t)|}{t} &= 4\int_0^{(2N+1)\pi} \frac{|\sin(u/2)|}{u} \\
\geqslant \sum_{n=0}^{N-1} \frac{2}{n\pi} \int_0^{2\pi} \sin(u/2) &= \frac{8}{\pi} \sum_{n=0}^{N-1} \frac{1}{n} = \Omega(\log N)
\end{aligned}$$

So the L^1 norm of D_n grows, albeit slowly, to ∞. This reflects the fact that D_n oscillates very frequently. Because of this, pointwise convergence of the Fourier series is much more subtle than that provided by good kernels. However, in the next section we show that, if we interpret the convergence in a different manner, we get a family of good kernels, and therefore we obtain pointwise convergence for suitable reinterpretations of partial sums.

2.4 Countercultural Methods of Summation

The standard method of summation suffices for much of analysis. Given a sequence $a_0, a_1, \dots$, we define the infinite sum as the limit of partial sums.

$$\sum_{k=0}^{\infty} a_k = \lim_{n \to \infty} \sum_{k=0}^{n} a_k$$

Similarily,

$$\sum_{k=-\infty}^{\infty} a_k = \sum_{k=0}^{\infty} a_k + a_{-k} = \lim_{n \to \infty} \sum_{k=-n}^{n} a_k$$

Some sums, like $\sum_{k=1}^{\infty} k$, obviously diverge, whereas other sums, like $\sum 1/n$, 'just' fail to converge because they grow suitably slowly towards infinity over time. Since the time of Euler, a new method of summation developed by Cesaro was introduced which 'regularized' certain terms by considering averaging the sums over time. Rather than considering limits of partial sums, we consider limits of averages of sums, known as Cesaro means. Letting $s_n = \sum_{k=0}^{n} a_k$, we define the Cesaro means

$$\sigma_n = \frac{s_0 + \dots + s_n}{n+1}$$

and then consider a sequence as Cesaro summable if the σ_n converge. If the normal summation exists, then the Cesaro limit exists, and is equal to the original sum. However, the Cesaro summation is stronger than normal convergence.

Example. *In the sense of Cesaro, we have*

$$1 - 1 + 1 - + \dots = 1/2$$

which reflects the fact that the partials sums do 'converge', but to two different numbers 0 *and* 1*, which the series oscillates between, and the Cesaro means average these two points of convergence out.*

Another notion of regularization sums emerged from Complex analysis, called Abel summation. Given a sequence $\{a_i\}$, we can consider the power series $\sum a_k r^k$. If this is well defined for $|r| < 1$, we can consider the Abel means $A_r = \sum a_k r^k$, and ask if $\lim_{r \to 1} A_r$ exists, which should be 'almost like' $\sum a_k$. If this limit exists, we call it the Abel sum of the sequence.

Example. *In the Abel sense, we have*

$$1 - 2 + 3 - 4 + 5 - \dots = \frac{1}{4}$$

because

$$\sum_{k=0}^{\infty} (-1)^k (k+1) z^k = \frac{1}{(1+z)^2}$$

the coefficients here are $\Omega(N)$*, so they can't be Cesaro summable.*

Abel summation is even more general than Cesaro summation.

Theorem 2.7. *If a sequence is Cesaro summable, it is Abel summable, and to the same value.*

Proof. Let $\{a_i\}$ be a Cesaro summable sequence, which we may without loss of generality assume converges to 0. Now $(n+1)\sigma_n - n\sigma_{n-1} = s_n$, so

$$(1-r)^2 \sum_{k=0}^{n} (k+1)\sigma_k r^k = (1-r) \sum_{k=0}^{n} s_k r^k = \sum_{k=0}^{n} a_k r^k$$

As $n \to \infty$, the left side tends to a well defined value for $r < 1$, hence the same is true for $\sum_{k=0}^{n} a_k r^k$. Given $\varepsilon > 0$, let N be large enough that $|\sigma_n| < \varepsilon$ for $n > N$, and let M be a bound for all $|\sigma_n|$. Then

$$\begin{aligned}
\left|(1-r)^2 \sum_{k=0}^{\infty} (k+1)\sigma_k r^k\right| &\leqslant (1-r)^2 \left(\sum_{k=0}^{N} (k+1)|\sigma_k| r^k + \varepsilon \sum_{k=N+1}^{\infty} (k+1) r^k \right) \\
&= (1-r)^2 \left(\sum_{k=0}^{N} (k+1)(|\sigma_k| - \varepsilon) r^k + \varepsilon \left[\frac{r^{n+1}}{1-r} + \frac{1}{(1-r)^2} \right] \right) \\
&\leqslant (1-r)^2 M \sum_{k=0}^{N} (k+1) r^k + \varepsilon r^{n+1}(1-r) + \varepsilon \\
&\leqslant (1-r)^2 M \frac{(N+1)(N+2)}{2} + \varepsilon r^{n+1}(1-r) + \varepsilon
\end{aligned}$$

Fixing N, and letting $r \to 1$, we may make the complicated sum on the end as small as possible, so the absolute value of the infinite sum is less than ε. Thus the Abel limit converges to zero. $\square$

Note that the Cesaro means of the Fourier series of f are given by

$$\sigma_N(f) = \frac{S_0(f) + \cdots + S_{N-1}(f)}{N} = f * \left(\frac{D_0 + \cdots + D_{N-1}}{N} \right) = f * F_N$$

The convergence properties of the Cesaro means therefore relate to the properties of the **Fejér kernel** F_N. We find that

$$F_N(x) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) e^{nit} = \frac{1}{N} \frac{\sin^2(Nx/2)}{\sin^2(x/2)}$$

so the oscillations of the Dirichlet kernel are slightly dampened, and as a result, F_N is a good kernel.

Theorem 2.8 (Fejér's Theorem)**.** *The Fourier series of f is Cesaro summable to f at every point of continuity, and is uniformly Cesaro summable if f is continuous everywhere.*

Relating Abel summations to Fourier series requires a little bit more careful work, since we do not consider limits of finite sums. Note that the Abel sum is

$$A_r(f) = \sum_{n=-\infty}^{\infty} \widehat{f}(n) r^n e^{nit}$$

Using the fact that the partial sums of the *Poisson kernel*

$$P_r(t) = \sum r^{|n|} e^{nit}$$

converge uniformly, we can calculate that $A_r(f) = P_r * f$. Thankfully, we find P_r is a good kernel, and the Abel means work nicely with Fourier series.

2.5 Returning to Heat Distributions on the Disk

Using the Abel summability properties of functions, we can solve Laplace's equation on a disk uniquely given any boundary conditions.

Theorem 2.9. *If f is integrable on $\mathbf{T}$, then the function*

$$u(re^{it}) = (f * P_r)(t)$$

is a $C^2(\mathbf{D}^\circ)$ harmonic function such that at any point of continuity t of f,

$$\lim_{r \uparrow 1} u(re^{it}) = f(t)$$

If f is continuous everywhere, then this limit is uniform, and u is the unique $C^2(\mathbf{D}^\circ)$ function with these properties.

Proof. Note that because P_r is C^∞ for all $r < 1$ (the power series and its derivatives converge uniformly), $f * P_r$ is C^∞, and so

$$u(re^{it}) = \sum \widehat{f}(n) r^{|n|} e^{nit}$$

which converges uniformly in every bounded disk around the origin of radius less than one. This justifies the calculation

$$\frac{\partial^2 u}{\partial t^2} = -\sum \widehat{f}(n) n^2 r^{|n|} e^{nit}$$

$$\frac{\partial (f * P_r)}{\partial r} = \sum \widehat{f}(n)|n| r^{|n|-1} e^{nit} \qquad \frac{\partial (f * P_r)}{\partial r^2} = \sum \widehat{f}(n)|n|(|n|-1) r^{|n|-2} e^{nit}$$

and now we find that on the interior of the unit disk,

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial t^2} = \sum [|n|(|n|-1) + |n| - n^2] \widehat{f}(n) r^{|n|-2} e^{nit} = 0$$

Abel summability results imply that $u(re^{it}) \to f(t)$ at every point of continuity of f, and the convergence is uniform if f is continuous.

Now instead, suppose that u solves $\Delta u = 0$ in the interior of the unit disk, and converges uniformly to f at the boundary. Because u is in $C^2(\mathbf{D}^\circ)$, we have a uniform sum

$$u_r(t) = \sum a_n(r) e^{nit}$$

where

$$a_n(r) = ⨍_{-\pi}^{\pi} u(re^{it}) e^{-int}$$

But then the $a_n(r)$ are differentiable, and

$$a_n'(r) = ⨍_{-\pi}^{\pi} \frac{\partial u}{\partial r}(re^{it}) e^{-int} \qquad a_n''(r) = ⨍_{-\pi}^{\pi} \frac{\partial^2 u}{\partial r^2}(re^{it}) e^{-int}$$

$$⨍_{-\pi}^{\pi} \frac{\partial^2 u}{\partial t^2}(re^{it}) e^{int} = -n^2 a_n(r)$$

and so $a_n'' + a_n'/r - n^2 a_n/r^2 = 0$, which is an ordinary differential equation whose only bounded solution is $a_n(r) = A_n r^n$. But now, if we let $r \to \infty$, we conclude that

$$A_n = \lim_{r \to 1} ⨍_{-\pi}^{\pi} u(re^{it}) e^{-int} = ⨍_{-\pi}^{\pi} u(e^{it}) e^{-int} = \widehat{f}(n)$$

so u is just $f * P_r$. $\square$

A final remark is that, if u is only required to converge to f *pointwise* on the boundary, then the conclusion may fail. We provide an example where u tends to zero pointwise on the boundary, but does not vanish on the interior of the unit disk.

Example. *If P_r is the Poisson kernel, define a function*

$$u(r,t) = \frac{\partial P_r}{\partial t}$$

Then u is harmonic in the unit disk, because by commutativity of derivatives,

$$\Delta u = \frac{\partial \Delta P_r}{\partial t} = \frac{\partial 0}{\partial t} = 0$$

and so

$$\begin{aligned}
u(r,t) &= \sum_{n=1}^{\infty} inr^n[e^{int} - e^{-int}] \\
&= i\left[\frac{re^{it}}{(re^{it}-1)^2} - \frac{re^{-it}}{(re^{-it}-1)^2}\right] \\
&= i\left[\frac{re^{-it} + r^{-1}e^{it} - re^{it} - r^{-1}e^{-it}}{(re^{it}-1)^2(re^{-it}-1)^2}\right] \\
&= \frac{(r-r^{-1})\sin(t)}{(re^{it}-1)^2(re^{-it}-1)^2} \\
&= \frac{(r^2-1)\sin(t)}{r(re^{it}-1)^2(re^{-it}-1)^2}
\end{aligned}$$

In this form, it is easy to see that for a fixed t, as $r \to 1$, $u(r,t) \to 0$. However, the denominator tells us this convergence isn't uniform.

2.6 L^2 convergence of Fourier Series

2.7 A Continuous Function with Divergent Fourier Series

Analysis was built to analyze continuous functions, so we would hope the method of fourier expansion would work for all continuous functions. Unfortunately, this is not so. The behaviour of the Dirichlet kernel away from the origin already tells us that the convergence of Fourier series is subtle. We shall take advantage of this to construct a continuous function with divergent fourier series at a point.

To start with, we shall consider the series

$$f(t) \sim \sum_{n \neq 0} \frac{e^{int}}{n}$$

where f is an odd function equaling $i(\pi - t)$ for $t \in (0, \pi]$. Such a function is nice to use, because its Fourier representation is simple, yet very close to diverging. Indeed, if we break the series into the pair

$$\sum_{n=1}^{\infty} \frac{e^{int}}{n} \qquad \sum_{n=-\infty}^{-1} \frac{e^{int}}{n}$$

Then these series no longer are the Fourier representations of a Riemann integrable function. For instance, if $g(t) \sim \sum_{n=1}^{\infty} \frac{e^{int}}{n}$, then the Abel means

$A_r(f)(t) =$

Chapter 3

The Fourier Transform

For the last 4 chapters, we have been discussing the role of Fourier analysis on $[-\pi, \pi]$. Is there any way to extend this to functions on $(-\infty, \infty)$? If f is such a function, we can certainly compute the fourier expansion by restricting f to $[-\pi, \pi]$, though there is no guarantee that the fourier expansion will converge outside of $[-\pi, \pi]$, to a function that looks anything like f. In general, we can also expand the function on $[-x, x]$, obtaining expansions of the form

$$f(t) \sim \sum_{n=-\infty}^{\infty} \frac{a_n}{2x} e^{nit/x}$$

where

$$a_n = \int_{-x}^{x} f(t) e^{-yit} dt$$

as we take x to ∞, we might expect the limit of the expansions on $[-x, x]$ to converge on all of $\mathbf{R}$, provided they converge to anything meaningful. The trick to guessing the convergence is to view these expansions as Riemann sums, sampling the **Fourier transform**

$$\widehat{f}(x) = \int_{-\infty}^{\infty} f(t) e^{-tix} dt$$

which results in the relation

$$f \sim \int_{-\infty}^{\infty} \widehat{f}(y) e^{yit} dy$$

This is the inversion formula, which essentially says that $\hat{f}$ is another representation of f (for we may obtain f uniquely, given that we know $\hat{f}$). The duality of a function and its Fourier transform shall be the main focus in this chapter.

For a general f, we may not even be able to define $\hat{f}$ for all real values, so it is hopeless to pursue the inversion formula for all functions. Thus, to understand the Fourier transform, we restrict ourselves to certain subclasses of all functions. This also gives us insight into the transform, for it tells us upon which subspaces the transformation performs well. First, we recall the definition of an integral over $\mathbf{R}$.

$$\int_{-\infty}^{\infty} f(x)dx = \lim_{y\to\infty} \int_{-y}^{y} f(x)dx$$

As should be expected this far into analysis, these types of limits do not play nicely with certain manipulations which will soon become essential. In the theory of series, we restrict our understanding to absolutely converging sequences; in integration, the corresponding objects are functions f such that

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

such a function shall be called absolutely integrable. It is then clear that $\int f(x)dx$ exists, because if

$$\int_{-\infty}^{\infty} |f(x)| - \int_{-a}^{a} |f(x)| < \varepsilon$$

Then for $b > a$,

$$\left| \int_{-b}^{b} f(x)dx - \int_{-a}^{a} f(x)dx \right| \leqslant \int_{-b}^{-a} |f(x)|dx + \int_{a}^{b} |f(x)|dx < \varepsilon$$

This is essentially the same proof as that for the convergence of absolutely convergent series.

If $f(x)$ is an absolutely integrable function, then $f(x)e^{-nix}$ is absolutely integrable, since $|e^{-nix}| = 1$ for all x. Thus we see that the Fourier transform is well defined for all real values. However, we still may not be able to interpret the inversion formula in this setting, because $\hat{f}$ may not be absolutely integrable. In the theory of Fourier series, we found that the smoothness of f had a direct relationship with the decay of $\widehat{f}$. We find respite in the refinement of our space of functions, considered by Schwartz and very useful in the analysis of the Fourier transform.

The **Schwartz space** consists of all smooth functions f (continous derivatives of all orders) which rapidly decrease at infinity. That is, for any $k > 0$, $l \in \mathbf{Z}$,

$$\sup |x|^k |f^{(l)}(x)| < \infty$$

We denote this space by $\mathcal{S}$. The Schwartz space is closed under addition, scalar multiplication, differentiation, and multiplication by polynomials.

It is not even obvious that $\mathcal{S}$ contains functions other than those which are constant, but there is a central example. Consider the Gaussian, defined by

$$f(x) = e^{-x^2}$$

s

Lemma 3.1. *If f is an increasing function which tends to ∞, and g is a decreasing function which tends to $-\infty$, then for any absolutely integrable h,*

$$\lim_{x \to \infty} \int_{g(x)}^{f(x)} h = \int_{-\infty}^{\infty} h$$

Proof. We shall prove the theorem assuming $h \geqslant 0$. In this case the limit above is increasing in x, and since

$$\int_{g(x)}^{f(x)} h \leqslant \int_{-M}^{M} h \leqslant \int_{-\infty}^{\infty} h$$

taking limits of both sides, we find

$$\lim_{x \to \infty} \int_{g(x)}^{f(x)} h \leqslant \int_{-\infty}^{\infty} h$$

Similarily, if we take x big enough that $N \leqslant f(x)$, $g(x) \leqslant -N$, then

$$\int_{g(x)}^{f(x)} h \geqslant \int_{-N}^{N} h$$

As we take x to ∞, we may also increase N to ∞, and we find

$$\lim_{x\to\infty}\int_{g(x)}^{f(x)} h \geqslant \int_{-\infty}^{\infty} h$$

and now we've squeezed the limit between the same value. The general case where h is not necessarily positive results by comparing the growth of h with the growth of $|h|$, as in the last theorem. $\square$

Corollary 3.2 (Translation Invariance)**.** *If f is absolutely integrable, and $h \in$ $\mathbf{R}$, then*

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} f(x+h)dx$$

Lemma 3.3. *If $\delta > 0$, and f is integrable on $\mathbf{R}$, then*

$$\int_{-\infty}^{\infty} f(\delta x)dx = \frac{1}{\delta}\int_{-\infty}^{\infty} f(x)$$

Proof. By the change of variables formula,

$$\int_{-N}^{N} f(\delta x)dx = \frac{1}{\delta}\int_{-\delta N}^{\delta N} f(y)dy$$

We then take limits of both sides of the equation. $\square$

We say a continuous function f is of **moderate decrease** if $|f| = O\left(\frac{1}{1+|x|^2}\right)$. Certainly then f is absolutely integrable.

Theorem 3.4. *If f is of moderate decrease, then*

$$\lim_{h\to 0}\int_{-\infty}^{\infty} |f(x-h) - f(x)|dx = 0$$

On the Schwarz space of infinitely differentiable rapidly vanishing functions, the Fourier transform is well defined

$$\widehat{f}(y) = \int e^{-2\pi i x y} f(x)dx$$

and is an automorphism of the vector space. We have the basic properties

$$\mathcal{F}(f(x+a)) = e^{2\pi i a}\mathcal{F}(f(x))$$

$$\mathcal{F}(e^{2\pi iax}f(x)) = f(x-a)$$

$$\mathcal{F}(f(bx)) = \frac{1}{b}f(x/b)$$

$$\mathcal{F}\left(\frac{df}{dx}\right) = 2\pi ix\frac{d\widehat{f}(x)}{dx}$$

$$\mathcal{F}(-2\pi ixf(x)) = \frac{d\widehat{f}(x)}{dx}$$

Theorem 3.5. *The Gaussian distribution function $f(x) = e^{-\pi x^2}$ satisfies $\widehat{f} = f$ (that is, the Gaussian functions are eigenvectors of the Fourier transform).*

Proof. We have

$$\frac{d\widehat{f}}{dx} = 2\pi i \int xe^{-2\pi ixy - \pi y^2}dy$$

□

Chapter 4

Applications

4.1 The Wirtinger Inequality on an Interval

Theorem 4.1. *Given $f \in C^1[-\pi,\pi]$ with $\int_{-\pi}^{\pi} f(t)dt = 0$,*

$$\int_{-\pi}^{\pi} |f(t)|^2 \leqslant \int_{-\pi}^{\pi} |f'(t)|^2$$

Proof. Consider the fourier series

$$f(t) \sim \sum a_n e^{nit} \qquad f'(t) \sim \sum ina_n e^{nit}$$

Then $a_0 = 0$, and so

$$\int_{-\pi}^{\pi} |f(t)|^2 \, dt = 2\pi \sum |a_n|^2 \leqslant 2\pi \sum n^2 |a_n|^2 = \int_{-\pi}^{\pi} |f'(t)|^2 \, dt$$

equality holds here if and only if $a_i = 0$ for $i > 1$, in which case we find

$$f(t) = Ae^{nit} + \overline{A}e^{-nit} = B\cos(t) + C\sin(t)$$

for some constants $A \in \mathbf{C}$, $B, C \in \mathbf{R}$. □

Corollary 4.2. *Given $f \in C^1[a,b]$ with $\int_a^b f(t)\, dt = 0$,*

$$\int_a^b |f(t)|^2 dt \leqslant \left(\frac{b-a}{\pi}\right)^2 \int_a^b |f'(t)|^2 \, dt$$

4.2 Energy Preservation in the String equation

Solutions to the string equation are

If $u(t, x)$

Part II

A More Sophisticated Viewpoint

Part III

Abstract Harmonic Analysis

The main property of spaces where Fourier analysis applies is symmetry – for a function $\mathbf{R}$, we can translate and invert about the axis. On $\mathbf{R}^n$ we have rotational symmetry as well. It turns out that we can apply Fourier analysis to any 'space with symmetry', that is, functions on an Abelian group. We shall begin with the study of finite abelian groups, where convergence questions disappear, and with it much of the analytical questions. We then proceed to generalize to a study of infinite abelian groups with topological structure.

Chapter 5

Finite Character Theory

Let us review our achievements so far. We have found several important families of functions on the spaces we have studied, and shown they can be used to approximate arbitrary functions. On the circle group $\mathbf{T}$, the functions take the form of the power maps

$$\phi_n : z \mapsto z^n$$

for $n \in \mathbf{Z}$. The important properties of these functions is that

- The functions are orthogonal to one another.
- A large family of functions can be approximated by linear combinations of the power maps.
- The power maps are multiplicative: $\phi_n(zw) = \phi_n(z)\phi_n(w)$.

The existence of a family with these properties is not dependant on much more than the symmetry properties of $\mathbf{T}$, and we can therefore generalize the properties of the fourier series to a large number of groups. In this chapter, we consider the generalization to finite abelian groups.

The last property of these exponentials should be immediately recognizable to any student of group theory. It implies the exponentials are homomorphisms from the circle group to itself. This is the easiest of the three properties to generalize to arbitrary groups; we shall call a homomorphism from a finite abelian group to $\mathbf{T}$ a **character**. For any abelian group G, we can put all characters together to form the character group $\Gamma(G)$, which forms an abelian group under pointwise multiplication.

$$(fg)(z) = f(z)g(z)$$

It is these functions which are 'primitive' in synthesizing functions defined on the group.

Example. *If μ_n is the set of nth roots of unity, then $\Gamma(\mu_n)$ consists of the power maps $\phi_m : z \mapsto z^m$, for $m \in \mathbf{Z}$. Since $\phi_i \phi_j = \phi_{i+j}$, and $\phi_i = \phi_j$ if and only if $i - j$ is divisible by n, this also shows that $\Gamma(\mu_n) \cong \mu_n$. Because*

$$\phi(\omega)^n = \phi(\omega^n) = \phi(1) = 1$$

we see that any character on μ_n is really a homomorphism from μ_n to μ_n. Since the homomorphisms on μ_n are determined by their action on this primitive root, there can only be at most n characters on μ_n, since there are only n elements in μ_n. Our derivation then shows us that the ϕ_n enumerate all such characters, which completes our proof.

Example. *The group $\mathbf{Z}_n$ is isomorphic to μ_n under the identification $n \mapsto \omega^n$, where ω is a primitive root of unity. This means that we do not need to distinguish functions 'defined in terms of n' and 'defined in terms of ω', assuming the correspondance $n = \omega^n$, as with the correspondence with functions on $\mathbf{T}$ and periodic functions on $\mathbf{R}$. The characters of $\mathbf{Z}_n$ are then exactly the maps $n \mapsto \omega^{kn}$. This follows from the general fact that if $f : G \to H$ is an isomorphism of abelian groups, the map $F : \phi \mapsto \phi \circ f$ is an isomorphism from $\Gamma(H)$ to $\Gamma(G)$.*

Example. *If K is a finite field, then the set K^* of non-zero elements is a group under multiplication. In fact, a rather sneaky algebraic proof shows the existence of elements of K, known as primitive elements, which generate the multiplicative group of all numbers. Thus K is cyclic, and therefore isomorphic to μ_n, where $n = |K| - 1$. The characters of K are then easily found under the correspondence. This is a fact that is incredibly important in analytic number theory, especially in the theory of Dirichlet characters.*

Example. *For a fixed n, the set of invertible elements of $\mathbf{Z}_n$ form a group under multiplication, denoted $\mathbf{Z}_n^*$. Any character from $\mathbf{Z}_n^*$ maps into $\mu_{\varphi(n)}$, because the order of each element in $\mathbf{Z}_n^*$ divides $\varphi(n)$. The groups are in general non-cyclic. For instance ,$\mathbf{Z}_8^* \cong \mathbf{Z}_2^3$. However, we can always break down a finite abelian group into cyclic subgroups to calculate the character group; a simple argument shows that $\Gamma(G \times H) \cong \Gamma(G) \times \Gamma(H)$, where we identify (f,g) with the map $(x,y) \mapsto f(x)g(y)$.*

5.1 Fourier analysis on Cyclic Groups

We shall start our study of abstract Fourier analysis by looking at Fourier analysis on μ_n. Geometrically, these points uniformly distribute themselves over $\mathbf{T}$, and therefore μ_n provides a good finite approximation to $\mathbf{T}$. Functions from μ_n to $\mathbf{C}$ are really just functions from $[n] = \{1, \dots, n\}$ to $\mathbf{C}$, so we're really computing the Fourier analysis of finite domain functions, in a way which encodes the translational symmetry of the function relative to translational shifts on $\mathbf{Z}_n$.

There is a trick which we can use to obtain quick results in μ_n. Given a function $f : [n] \to \mathbf{C}$, consider the n-periodic function defined by

$$g(t) = \sum_{k=1}^{n} f(k)\mathbf{I}\left(t \in (k - 1/2, k + 1/2)\right)$$

Classical Fourier analysis of this function (and the fact that g is differentiable at each k) give us an expansion of $f(k)$ as an infinite series in the $e^{2\pi i k l/n}$ for $l \in \mathbf{Z}$, which may be summed up over equivalence classes of kl modulo n to give us a finite expansion. This method certainly 'works', but does not generalize to finite abelian groups which cannot be expressed as subgroups of $\mathbf{T}$. It turns out that the theory of characters in the correct generalization to arbitrary finite abelian groups. Here, we notice that the maps $\phi_k : m \mapsto e^{2\pi i m k/n}$ are exactly the set of characters from $\mathbf{Z}_n$ to $\mathbf{C}$. Dual to this trick, we can recover much of the theory of Fourier series on $\mathbf{T}$ from looking at approximations of functions on μ_n.

The correct generalization of Fourier analysis is to analyze the set of 'square integrable functions' on μ_n, which is really just the set of all functions from μ_n to $\mathbf{C}$ since the domain is finite. We make V into an inner product space by defining

$$\langle f, g \rangle = \sum_{k=1}^{n} f(k)\overline{g(k)}$$

We claim that the characters $\phi_i : z \mapsto z^i$ are orthonormal in this space, since

$$\langle \phi_i, \phi_j \rangle = \sum_{k=1}^{n} \omega^{k(i-j)}$$

If $i = j$, we may sum up to find $\langle e_i, e_j \rangle = n$. Otherwise we use a standard summation formula to find

$$\sum_{k=1}^{n} \omega^{k(i-j)} = \omega^{(i-j)} \frac{\omega^{n(i-j)} - 1}{\omega^{(i-j)} - 1}$$

But $\omega^{n(i-j)} = 1$, since it is an n'th root of unity, so the sum is zero.

Since the family of characters is orthonormal, they are linearly independant. Because V is only n dimensional, the family of characters must span the space. Thus, for any function f, we have

$$f(k) = \sum_{m=1}^{n} \frac{\langle f, \phi_m \rangle}{\langle \phi_m, \phi_m \rangle} \omega^{mk} = \sum_{m=1}^{n} \left(\frac{1}{n} \sum_{l=1}^{n} f(l) \omega^{-ml} \right) \omega^{mk}$$

We can perform this calculation, with slightly more care, on an arbitrary abelian group to obtain the same results.

5.2 An Arbitrary Finite Abelian Group

It should be easy to guess how we proceed for a general finite abelian group. Given some group G, we study the character group $\hat{G}$, and how $\hat{G}$ represents general functions from G to $\mathbf{C}$. We shall let V be the space of all such functions, and on it we define the inner product

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{a \in G} f(a) \overline{g(a)}$$

If there's any justice in the world, the characters of G should be independant.

Theorem 5.1. *The set $\Gamma(G)$ of characters is an orthonormal set.*

Proof. If e is a character of G, then $|e(a)| = 1$ for each a, and so

$$\langle e, e \rangle = \frac{1}{|G|} \sum_{a \in G} |e(a)| = 1$$

Now if $e \neq 1$ is a non-trivial character, then

$$\sum_{a \in G} e(a) = 0$$

Now for any $b \in G$, the map $a \mapsto ba$ is a bijection of G, and so

$$e(b) \sum_{a \in G} e(a) = \sum_{a \in G} e(ba) = \sum_{a \in G} e(a)$$

Implying either $e(b) = 1$, or $\sum_{a \in G} e(a) = 0$. Finally, if $e \neq e'$ are characters,

$$\langle e, e' \rangle = \frac{1}{|G|} \sum_{a \in G} e(a) e'(a)^{-1} = 0$$

since $e(e')^{-1} \neq 1$ is a character. $\square$

Because elements of $\Gamma(G)$ are orthonormal, they are linearly independent over the space of functions on G, and we obtain a bound $|\Gamma(G)| \leqslant |G|$. All that remains is to show equality. This can be shown very simply by applying the structure theorem for finite abelian groups. First, note it is true for all cyclic groups. Second, note that if it is true for two groups G and H, it is true for $G \times H$, because

$$\Gamma(G \times H) \cong \Gamma(G) \times \Gamma(H)$$

since a finite abelian group is a finite product of cyclic groups, this proves the theorem. There are more basic ways to prove this statement however, without using much more than elementary linear algebra. What's more, the linear algebraic techniques generalize to the theory of unitary representations in harmonic analysis over infinite groups.

Theorem 5.2. *Let $\{T_1, \dots, T_n\}$ be a family of commuting unitary matrices. Then there is a basis $v_1, \dots, v_m \in \mathbf{C}^m$ which are eigenvectors for each T_i.*

Proof. For $n = 1$, the theorem is the standard spectral theorem. For induction, suppose that the $T_1, \dots, T_{k-1}$ are simultaneously diagonalizable. Write

$$\mathbf{C}^m = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_l}$$

where λ_i are the eigenvalues of T_k, and V_{λ_i} are the corresponding eigenspaces. Then if $v \in V_{\lambda_i}$, and $j < k$,

$$T_k T_j v = T_j T_k v = \lambda_i T_j v$$

so $T_j(V_{\lambda_i}) = V_{\lambda_i}$. Now on each V_{λ_i}, we may apply the induction hypotheis to diagonalize the $T_1, \dots, T_{k-1}$. Putting this together, we simultaneously diagonalize $T_1, \dots, T_k$. $\square$

This theorem enables us to prove the character theory in a much more simple manner. Let V be the space of complex valued functions on G, and define, for $a \in G$, the map $(T_a f)(b) = f(ab)$. An orthonormal basis of V is the set $\mathcal{B}$ of χ_a, for $a \in G$, which maps a to 1, and maps any other element to zero. In this basis, we find $T_a \chi_b = \chi_{ba^{-1}}$, so the matrix of T_a with respect to this basis is a permutation matrix, hence T_a is unitary. The operators commute, since $T_a T_b = T_{ab} = T_{ba} = T_b T_a$. Hence these operators can be simultaneously diagonalized. That is, there is a family $f_1, \dots, f_n$ such that for each $a \in G$, $T_a f_i = \lambda_{ia} f_i$. We may assume $f_i(1) = 1$ for each i. Then, for any $a \in G$, we have $f_i(a) = f_i(a \cdot 1) = \lambda_{ia} f_i(1) = \lambda_{ia}$, so for any $b \in G$, $f_i(ab) = \lambda_{ia} f_i(b) = f_i(a) f_i(b)$. This shows each f_i is a character, completing the proof. We summarize our discussion in the following theorem.

Theorem 5.3. *Let G be a finite abelian group. Then $\Gamma(G)$ has the same cardinality as G, and forms an orthogonal basis for the space of complex valued functions on G. For any function $f : G \to \mathbf{C}$,*

$$f(a) = \sum_{e \in \hat{G}} \langle f, e \rangle \, e(a) = \sum_{e \in \hat{G}} \hat{f}(e) e(a) \qquad \langle f, g \rangle = \frac{1}{|G|} \sum_{a \in G} f(a) \overline{g(a)}$$

In this context, we also have Parseval's theorem

$$\|f(a)\|^2 = \sum_{e \in \hat{G}} |\hat{f}(e)|^2 \qquad \langle f, g \rangle = \sum_{e \in \hat{G}} \hat{f}(e) \overline{\hat{g}(e)}$$

5.3 Convolutions

There is a version of convolutions for finite functions, which is analogous to the convolutions on $\mathbf{R}$. Given two functions f, g, we define

$$(f * g)(a) = \frac{1}{|G|} \sum_{b \in G} f(b) g(b^{-1} a)$$

The mapping $b \mapsto ab^{-1}$ is a bijection of G, and so we also have

$$(f * g)(a) = \frac{1}{|G|} \sum_{b \in G} f(ab^{-1}) g(b) = (g * f)(a)$$

Now for $e \in \hat{G}$,

$$\begin{aligned}\widehat{(f * g)}(e) &= \frac{1}{|G|}\sum_{a \in G}(f * g)(a)\overline{e(a)} \\ &= \frac{1}{|G|^2}\sum_{a,b \in G} f(ab)g(b^{-1})\overline{e(a)}\end{aligned}$$

The bijection $a \mapsto ab^{-1}$ shows that

$$\begin{aligned}\widehat{(f * g)}(e) &= \frac{1}{|G|^2}\sum_{a,b} f(a)g(b^{-1})\overline{e(a)}\overline{e(b^{-1})} \\ &= \frac{1}{|G|}\left(\sum_a f(a)\overline{e(a)}\right)\frac{1}{|G|}\left(\sum_b g(b)\overline{e(b)}\right) \\ &= \hat{f}(e)\hat{g}(e)\end{aligned}$$

In the finite case we do not need approximations to the identity, for we have an identity for convolution. Define $D : G \to \mathbf{C}$ by

$$D(a) = \sum_{e \in \hat{G}} e(a)$$

We claim that $D(a) = |G|$ if $a = 1$, and $D(a) = 0$ otherwise. Note that since $|G| = |\hat{G}|$, the character space of $\hat{G}$ is isomorphic to G. Indeed, for each $a \in G$, we have the maps $\hat{a} : e \mapsto e(a)$, which is a character of $\hat{G}$. Suppose $e(a) = 1$ for all characters e. Then $e(a) = e(1)$ for all characters e, and for any function $f : G \to \mathbf{C}$, we have $f(a) = f(1)$, implying $a = 1$. Thus we obtain $|G|$ distinct maps $\hat{a}$, which therefore form the space of all characters. It therefore follows from a previous argument that if $a \neq 1$, then

$$\sum_{e \in \hat{G}} e(a) = 0$$

Now $f * D = f$, because

$$\widehat{D}(e) = \frac{1}{|G|}\sum_{a \in G} D(a)\overline{e(a)} = \overline{e}(1) = 1$$

D is essentially the finite dimensional version of the 'Dirac delta function', since it has unit mass, and acts as the identity in convolution.

5.4 The Fast Fourier Transform

The main use of the fourier series on μ_n is to approximate the Fourier transform on $\mathbf{T}$, where we need to compute integrals explicitly. If we have a function $f \in L^1(\mathbf{T})$, then f may be approximated in $L^1(\mathbf{T})$ by step functions of the form

$$f_n(t) = \sum_{k=1}^{n} a_k \mathbf{I}(x \in (2\pi(k-1)/n, 2\pi k/n))$$

And then $\hat{f}_n \to \hat{f}$ uniformly. The Fourier transform of f_n is the same as the Fourier transform of the corresponding function $k \mapsto a_k$ on $\mathbf{Z}_n$, and thus we can approximate the Fourier transform on $\mathbf{T}$ by a discrete computation on $\mathbf{Z}_n$. Looking at the formula in the definition of the discrete transform, we find that we can compute the Fourier coefficients of a function $f : \mathbf{Z}_n \to \mathbf{C}$ in $O(n^2)$ addition and multiplication operations. It turns out that there is a much better method of computation which employs a divide and conquer approach, which works when n is a power of 2, reducing the calculation to $O(n \log n)$ multiplications.

To see this, consider a particular division in the group $\mathbf{Z}_{2n}$. Given $f : \mathbf{Z}_{2n} \to \mathbf{C}$, define two functions $g, h : \mathbf{Z}_n \to \mathbf{C}$, defined by $g(k) = f(2k)$, and $h(k) = f(2k+1)$. Then g and h encode all the information in f, and if $\nu = e^{\pi i/n}$ is the canonical generator of $\mathbf{Z}_{2n}$, we have

$$\hat{f}(m) = \frac{\hat{g}(m) + \hat{h}(m)\nu^m}{2}$$

Because

$$\begin{aligned}\frac{1}{2n}\sum_{k=1}^{n}\left(g(k)\omega^{-km} + h(m)\omega^{-km}\nu^m\right) &= \frac{1}{2n}\sum_{k=1}^{n} f(2k)\nu^{-2km} + f(2k+1)\nu^{-(2k+1)m} \\ &= \frac{1}{2n}\sum_{k=1}^{2n} f(k)\nu^{-km}\end{aligned}$$

This is essentially a discrete analogue of the Poission summation formula, which we will generalize later when we study the harmonic analysis of abelian groups. If $H(m)$ is the number of operations needed to calculate the Fourier transform of a function on μ_{2^n} using the above recursive formula, then the above relation tells us $H(2m) = 2H(m) + 3(2m)$.

If $G(n) = H(2^n)$, then $G(n) = 2G(n-1) + 32^n$, and $G(0) = 1$, and it follows that

$$G(n) = 2^n + 3\sum_{k=1}^{n} 2^k 2^{n-k} = 2^n(1+3n)$$

Hence for $m = 2^n$, we have $H(m) = m(1 + 3\log(m)) = O(m \log m)$. Similar techniques show that one can compute the inverse Fourier transform in $O(m \log m)$ operations (essentially by swapping the root ν with ν^{-1}).

5.5 Dirichlet's Theorem

We now apply the theory of Fourier series on finite abelian groups to prove Dirichlet's theorem.

Theorem 5.4. *If m and n are relatively prime, then the set*

$$\{m + kn : k \in \mathbf{N}\}$$

contains infinitely many prime numbers.

An exploration of this requries the Riemann-Zeta function, defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

The function is defined on $(1, \infty)$, since for $s > 1$ the map $t \mapsto 1/t^s$ is decreasing, and so

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \leqslant 1 + \int_1^{\infty} \frac{1}{t^s} = 1 + \lim_{n \to \infty} \frac{1}{s-1}\left[1 - 1/n^{s-1}\right] = 1 + \frac{1}{s-1}$$

The series converges uniformly on $[1+\varepsilon, N]$ for any $\varepsilon > 0$, so ζ is continuous on $(1, \infty)$. As $t \to 1$, $\zeta(t) \to \infty$, because $n^s \to n$ for each n, and if for a fixed M we make s close enough to 1 such that $|n/n^s - 1| < 1/2$ for $1 \leqslant n \leqslant M$, then

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \geqslant \sum_{n=1}^{M} \frac{1}{n^s} = \sum_{n=1}^{M} \frac{1}{n}\frac{n}{n^s} \geqslant \frac{1}{2}\sum_{n=1}^{M} \frac{1}{n}$$

Letting $M \to \infty$, we obtain that $\sum_{n=1}^{\infty} \frac{1}{n^s} \to \infty$ as $s \to 1$.

The Riemann-Zeta function is very good at giving us information about the prime integers, because it encodes much of the information about the prime numbers.

Theorem 5.5. *For any $s > 1$,*

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - p^s}$$

Proof. The general idea is this – we may write

$$\prod_{p \text{ prime}} \frac{1}{1-p^s} = \prod_{p \text{ prime}} (1 + 1/p^s + 1/p^{2s} + \dots)$$

If we expand this product out formally, enumating the primes to be $p_1, p_2, \dots$, we find

$$\prod_{p \leqslant n} (1 + 1/p^s + 1/p^{2s} + \dots) = \sum_{n_1, n_2, \dots = 0}^{\infty} \frac{1}{p_1^{n_1}}$$

□

Chapter 6

Topological Groups

In abstract harmonic analysis, the main subject matter is the **topological group**, a group G equipped with a topology which makes the operation of multiplication and inversion continuous. In the mid 20th century, it was realized that basic Fourier analysis could be generalized to a large class of groups. The nicest generalization occurs over the locally compact groups, which simplifies the theory considerably.

Example. *There are a few groups we should keep in mind for intuition in the general topological group.*

- *The classical groups $\mathbf{R}^n$ and $\mathbf{T}^n$, from which Fourier analysis originated.*
- *The group μ of roots of unity, rational numbers $\mathbf{Q}$, and cyclic groups $\mathbf{Z}_n$.*
- *The matrix subgroups of the general linear group $GL(n)$.*
- *The product $\mathbf{T}^\omega$ of Torii, occurring in the study of Dirichlet series.*
- *The product $\mathbf{Z}_2^\omega$, which occurs in probability theory, and other contexts.*
- *The field of p-adic numbers $\mathbf{Q}_p$, which are the completion of $\mathbf{Q}$ with respect to the absolute value $|p^{-m}q|_p = p^m$.*

6.1 Basic Results

The topological structure of a topological group naturally possesses large amounts of symmetry, simplifying the spatial structure. For any topological group, the maps

$$x \mapsto gx \qquad x \mapsto xg \qquad x \mapsto x^{-1}$$

are homeomorphisms. Thus if U is a neighbourhood of x, then gU is a neighbourhood of gx, Ug a neighbourhood of xg, and U^{-1} a neighbourhood of x^{-1}, and as we vary U through all neighbourhoods of x, we obtain all neighbourhoods of the other points. Understanding the topological structure at any point reduces to studying the neighbourhoods of the identity element of the group.

In topological group theory it is even more important than in basic group theory to discuss set multiplication. If U and V are subsets of a group, then we define

$$U^{-1} = \{x^{-1} : x \in U\} \qquad UV = \{xy : x \in U, y \in V\}$$

We let $V^2 = VV$, $V^3 = VVV$, and so on.

Theorem 6.1. *Let U and V be subsets of a topological group.*

(i) If U is open, then UV is open.

(ii) If U is compact, and V closed, then UV is closed.

(iii) If U and V are connected, UV is connected.

(iv) If U and V are compact, then UV is compact.

Proof. To see that (i) holds, we see that

$$UV = \bigcup_{x \in V} Ux$$

and each Ux is open. To see (ii), suppose $u_i v_i \to x$. Since U is compact, there is a subnet u_{i_k} converging to y. Then $y \in U$, and we find

$$v_{i_k} = u_{i_k}^{-1}(u_{i_k} v_{i_k}) \to y^{-1}x$$

Thus $y^{-1}x \in V$, and so $x = yy^{-1}x \in UV$. (iii) follows immediately from the continuity of multiplication, and the fact that $U \times V$ is connected, and (iv) follows from similar reasoning. $\square$

Example. *If U is merely closed, then (ii) need not hold. For instance, in* $\mathbf{R}$*, take* $U = \alpha\mathbf{Z}$*, and* $V = \mathbf{Z}$*, where α is an irrational number. Then* $U + V = \alpha\mathbf{Z} + \mathbf{Z}$ *is dense in* $\mathbf{R}$*, and is hense not closed.*

There are useful ways we can construct neighbourhoods under the group operations, which we list below.

Lemma 6.2. *Let U be a neighbourhood of the identity. Then*

(1) There is an open V such that $V^2 \subset U$.

(2) There is an open V such that $V^{-1} \subset U$.

(3) For any $x \in U$, there is an open V such that $xV \subset U$.

(4) For any x, there is an open V such that $xVx^{-1} \subset U$.

Proof. (1) follows simply from the continuity of multiplication, and (2) from the continuity of inversion. (3) is verified because $x^{-1}U$ is a neighbourhood of the origin, so if $V = x^{-1}U$, then $xV = U \subset U$. Finally (4) follows in a manner analogously to (3) because $x^{-1}Ux$ contains the origin. □

If $\mathcal{U}$ is an open basis at the origin, then it is only a slight generalization to show that for any of the above situations, we can always select $V \in \mathcal{U}$. Conversely, suppose that $\mathcal{V}$ is a family of subsets of a (not yet topological) group G containing e such that (1), (2), (3), and (4) hold. Then the family $\mathcal{V}' = \{xV : V \in \mathcal{V}, x \in G\}$ forms a subbasis for a topology on G which forms a topological group. If $\mathcal{V}$ also has the base property, then $\mathcal{V}'$ is a basis.

Theorem 6.3. *If K and C are disjoint, K is compact, and C is closed, then there is a neighbourhood V of the origin for which KV and CV is disjoint. If G is locally compact, then we can select V such that KV is precompact.*

Proof. For each $x \in K$, C^c is an open neighbourhood containing x, so by applying the last lemma recursively we find that there is a symmetric neighbourhood V_x such that $xV_x^4 \subset C^c$. Since K is compact, finitely many of the xV_x cover K. If we then let V be the open set obtained by intersecting the finite subfamily of the V_x, then KV is disjoint from CV. □

Taking K to be a point, we find that any open neighbourhood of a point contains a closed neighbourhood. Provided points are closed, we can set C to be a point as well.

Corollary 6.4. *Every Kolmogorov topological group is Hausdorff.*

Related to this theorem is the

Theorem 6.5. *For any set $A \subset G$,*

$$\overline{A} = \bigcap_V AV$$

Where V ranges over the set of neighbourhoods of the origin.

Proof. If $x \notin \overline{A}$, then the last theorem guarantees that there is V for which $\overline{A}V$ and Ax are disjoint. We conclude $\bigcap AV \subset \overline{A}$. Conversely, any neighbourhood contains a closed neighbourhood, so that $\overline{A} \subset AV$ for a fixed V, and hence $\overline{A} \subset \bigcap AV$. □

Theorem 6.6. *Every open subgroup of G is closed.*

Proof. Let H be an open subgroup of G. Then

$$\overline{H} = \bigcap_V HV$$

If W is a neighbourhood of the origin contained in H, then we find

$$\overline{H} \subset HW \subset H$$

so H is closed. □

We see that open subgroups of a group therefore correspond to connected components of the group, so that connected groups have no proper open subgroups. This also tells us that a locally compact group is σ-compact on each of its components, for if V is a pre-compact neighbourhood of the origin, then $V^2, V^3, \dots$ are all precompact, and $\bigcup_{k=1}^{\infty} V^k$ is an open subgroup of G, which therefore contains the component of e, and is σ-compact. Since the topology of a topological group is homogenous, we can conclude that all components of the group are σ compact.

6.2 Quotient Groups

If G is a topological group, and H is a subgroup, then G/H can be given a topological structure in the obvious way. The quotient map is open, because VH is open in G for any open set V, and if H is normal, G/H is also a topological group, because multiplication is just induced from the quotient map of $G \times G$ to $G/H \times G/H$, and inversion from G to G/H. We should think the quotient structure is pleasant, but if no conditions on H are given, then G/H can have pathological structure. One particular example is the quotient $\mathbf{T}/\mu_\infty$ of the torus modulo the roots of unity, where the quotient is lumpy.

Theorem 6.7. *If H is closed, G/H is Hausdorff.*

Proof. If $x \neq y \in G/H$, then xHy^{-1} is a closed set in G, not containing e, so we may conclude there is a neighbourhood V for which V and $VxHy^{-1}$ are disjoint, so VyH and VxH are disjoint. This implies that the open sets $V(xH)$ and $V(yH)$ are disjoint in G/H. □

Theorem 6.8. *If G is locally compact, G/H is also.*

Proof. If $\{U_i\}$ is a basis of precompact neighbourhoods at the origin, then U_iH is a family of precompact neighbourhoods of the origin in G/H, and is in fact a basis, for if V is any neighbourhood of the origin, there is $U_i \subset \pi^{-1}(V)$, and so $U_iH \subset V$. □

If G is a non-Hausdorff group, then $\overline{\{e\}} \neq \{e\}$, and $G/\overline{\{e\}}$ is Hausdorff. Thus we can get away with assuming all our topological groups are Hausdorff, because a slight modification in the algebraic structure of the topological group gives us this property.

6.3 Uniform Continuity

An advantage of the real line $\mathbf{R}$ is that continuity can be explained in a *uniform sense*, because we can transport any topological questions about a certain point x to questions about topological structure near the origin via the map $g \mapsto x^{-1}g$. We can then define a uniformly continuous function $f : \mathbf{R} \to \mathbf{R}$ to be a function possessing, for every $\varepsilon > 0$, a $\delta > 0$ such that if $|y| < \delta$, $|f(x+y) - f(x)| < \varepsilon$. Instead of having to specify a δ for every point on the domain, the δ works uniformly everywhere. The group structure is all we need to talk about these questions.

We say a function $f : G \to H$ between topological groups is (left) uniformly continuous if, for any open neighbourhood U of the origin in H, there is a neighbourhood V of the origin in G such that for each x, $f(xV) \subset f(x)U$. Right continuity requires $f(Vx) \subset Uf(x)$. The requirement of distinguishing between left and right uniformity is important when we study non-commutative groups, for there are certainly left uniform maps which are not right uniform in these groups. If $f : G \to \mathbf{C}$, then left uniform continuity is equivalent to the fact that $\|L_x f - f\|_\infty \to 0$ as $x \to 1$, where $(L_x f)(y) = f(xy)$. Right uniform continuity requires $\|R_x f - f\|_\infty \to 0$, where $(R_x f)(y) = f(yx)$. R_x is a homomorphism, but L_x is what is called an antihomomorphism.

Example. *Let G be any Hausdorff non-commutative topological group, with sequences x_i and y_i for which $x_i y_i \to e$, $y_i x_i \to z \neq e$. Then the uniform structures on G are not equivalent.*

It is hopeless to express uniform continuity in terms of a new topology on G, because the topology only gives a local description of continuity, which prevents us from describing things uniformly across the whole group. However, we can express uniform continuity in terms of a new topology on $G \times G$. If $U \subset G$ is an open neighbourhood of the origin, let

$$L_U = \{(x,y) : yx^{-1} \in U\} \qquad R_U = \{(x,y) : x^{-1}y \in U\}$$

The family of all L_U (resp. R_U) is known as the left (right) uniform structure on G, denoted $LU(G)$ and $RU(G)$. Fix a map $f : G \to H$, and consider the map

$$g(x,y) = (f(x), f(y))$$

from G^2 to H^2. Then f is left (right) uniformly continuous if and only if g is continuous with respect to $LU(G)$ and $LU(H)$ ($RU(G)$ and $RU(H)$). $LU(G)$ and $RU(G)$ are weaker than the product topologies on G and H, which reflects the fact that uniform continuity is a strong condition than normal continuity. We can also consider uniform maps with respect to $LU(G)$ and $RU(H)$, and so on and so forth. We can also consider uniform continuity on functions defined on an open subset of a group.

Example. *Here are a few examples of easily verified continuous maps.*

- *If the identity map on G is left-right uniformly continuous, then $LU(G) = RU(G)$, and so uniform continuity is invariant of the uniform structure chosen.*

- *Translation maps $x \mapsto axb$, for $a, b \in G$, are left and right uniform.*

- *Inversion is uniformly continuous.*

Theorem 6.9. *All continuous maps on compact subsets of topological groups are uniformly continuous.*

Proof. Let K be a compact subset of a group G, and let $f : K \to H$ be a continuous map into a topological group. We claim that f is then uniformly continuous. Fix an open neighbourhood V of the origin, and let V' be a symmetric neighbourhood such that $V'^2 \subset V$. For any x, there is U_x such that

$$f(x)^{-1} f(xU_x) \subset V'$$

Choose U'_x such that $U'^2_x \subset U_x$. The xU'_x cover K, so there is a finite subcover corresponding to sets $U'_{x_1}, \dots, U'_{x_n}$. Let $U = U'_{x_1} \cap \dots \cap U'_{x_n}$. Fix $y \in G$, and suppose $y \in x_k U'_{x_k}$. Then

$$\begin{aligned} f(y)^{-1} f(yU) &= f(y)^{-1} f(x_k) f(x_k)^{-1} f(yU) \\ &\subset f(y)^{-1} f(x_k) f(x_k)^{-1} f(x_k U x_k) \\ &\subset f(y)^{-1} f(x_k) V' \\ &\subset V'^2 \subset V \end{aligned}$$

So that f is left uniformly continuous. Right uniform continuity is proven in the exact same way. ☐

Corollary 6.10. *All maps with compact support are uniformly continuous.*

Corollary 6.11. *Uniform continuity on compact groups is invariant of the uniform structure chosen.*

6.4 Ordered Groups

In this section we describe a general class of groups which contain both interesting and pathological examples. Let G be a group with an ordering

$<$ preserved by the group operations, so that $a < b$ implies both $ag < bg$ and $ga < gb$. We now prove that the order topology gives G the structure of a normal topological group (the normality follows because of general properties of order topologies).

First note, that $a < b$ implies $a^{-1} < b^{-1}$. This results from a simple algebraic trick, because

$$a^{-1} = a^{-1}bb^{-1} > a^{-1}ab^{-1} = b^{-1}$$

This implies that the inverse image of an interval (a,b) under inversion is (b^{-1}, a^{-1}), hence inversion is continuous.

Now let $e < b < a$. We claim that there is then $e < c$ such that $c^2 < a$. This follows because if $b^2 \geqslant a$, then $b \geqslant ab^{-1}$ and so

$$(ab^{-1})^2 = ab^{-1}ab^{-1} \leqslant ab^{-1}b = a$$

Now suppose $a < e < b$. If $\inf\{y : y > e\} = x > e$, then $(x^{-1}, x) = \{e\}$, and the topology on G is discrete, hence the continuity of operations is obvious. Otherwise, we may always find c such that $c^2 < b$, $a < c^{-2}$, and then if $c^{-1} < g, h < c$, then

$$a < c^{-2} < gh < c^2 < b$$

so multiplication is continuous at every pair (x, x^{-1}). In the general case, if $a < gh < b$, then $g^{-1}ah^{-1} < e < g^{-1}bh^{-1}$, so there is c such that if $c^{-1} < g', h' < c$, then $g^{-1}ah^{-1} < g'h' < g^{-1}bh^{-1}$, so $a < gg'h'h < b$. The set of gg', where $c^{-1} < g' < c$, is really just the set of $gc^{-1} < x < gc$, and the set of $h'h$ is really just the set of $c^{-1}h < x < ch$. Thus multiplication is continuous everywhere.

Example (Dieudonne)**.** *For any well ordered set S, the dictionary ordering on $\mathbf{R}^S$ induces a linear ordering inducing a topological group structure on the set of maps from S to $\mathbf{R}$.*

Let us study Dieudonne's topological group in more detail. If S is a finite set, or more generally possesses a maximal element w, then the topology on $\mathbf{R}^S$ can be defined such that $f_i \to f$ if eventually $f_i(s) = f(s)$ for all $s < w$ simultaneously, and $f_i(w) \to f(w)$. Thus $\mathbf{R}^S$ is isomorphic (topologically) to a discrete union of a certain number of copies of $\mathbf{R}$, one for each tuple in $S - \{w\}$.

If S has a countable cofinal subset $\{s_i\}$, the topology is no longer so simple, but $\mathbf{R}^S$ is still first countable, because the sets

$$U_i = \{f : (\forall w < s_i : f(w) = 0)\}$$

provide a countable neighbourhood basis of the origin.

The strangest properties of $\mathbf{R}^S$ occur when S has no countable cofinal set. Suppose that $f_i \to f$. We claim that it follows that $f_i = f$ eventually. To prove by contradiction, we assume without loss of generality (by thinning the sequence) that no f_i is equal to f. For each f_i, find the largest $w_i \in S$ such that for $s < w_i$, $f_i(s) = f(s)$ (since S is well ordered, the set of elements for which $f_i(s) \neq f(s)$ has a minimal element). Then the w_i form a countable cofinal set, because if $v \in S$ is arbitrary, the f_i eventually satisfy $f_i(s) = f(s)$ for $s < v$, hence the corresponding w_i is greater than v_i. Hence, if $f_i \to f$ in $\mathbf{R}^S$, where S does not have a countable cofinal subset, then eventually $f_i = f$. We conclude all countable sets in $\mathbf{R}^S$ are closed, and this proof easily generalises to show that if S does not have a cofinal set of cardinality $\mathfrak{a}$, then every set of cardinality $\leqslant \mathfrak{a}$ is closed.

The simple corollary to this proof is that compact subsets are finite. Let $X = f_1, f_2, \dots$ be a denumerable, compact set. Since all subsets of X are compact, we may assume $f_1 < f_2 < \dots$ (or $f_1 > f_2 > \dots$, which does not change the proof in any interesting way). There is certainly $g \in \mathbf{R}^S$ such that $g < f_1$, and then the sets $(g, f_2), (f_1, f_3), (f_2, f_4), \dots$ form an open cover of X with no finite subcover, hence X cannot be compact. We conclude that the only compact subsets of $\mathbf{R}^S$ are finite.

Furthermore, the class of open sets is closed under countable intersections. Consider a series of functions

$$f_1 \leqslant f_2 \leqslant \dots < h < \dots \leqslant g_2 \leqslant g_1$$

Suppose that $f_i \leqslant k < h < k' \leqslant g_j$. Then the intersection of the (f_i, g_i) contains an interval (k, k') around h, so that the intersection is open near h. The only other possiblity is that $f_i \to h$ or $g_i \to h$, which can only occur if $f_i = h$ or $g_i = h$ eventually, in which case we cannot have $f_i < h$, $h < g_i$. We conclude the intersection of countably many intervals is open, because we can always adjust any intersection to an intersection of this form without changing the resulting intersecting set (except if the set is empty, in which case the claim is trivial). The general case results from noting that any open set in an ordered group is a union of intervals.

6.5 Topological Groups arising from Normal subgroups

Let G be a group, and $\mathcal{N}$ a family of normal subgroups closed under intersection. If we interpret $\mathcal{N}$ as a neighbourhood base at the origin, the resulting topology gives G the structure of a totally disconnected topological group, which is Hausdorff if and only if $\bigcap \mathcal{N} = \{e\}$. First note that $g_i \to g$ if g_i is eventually in gN, for every $N \in \mathcal{N}$, which implies $g_i^{-1} \in Ng^{-1} = g^{-1}N$, hence inversion is continuous. Furthermore, if h_i is eventually in hN, then $g_i h_i \in gNhN = ghN$, so multiplication is continuous. Finally note that $N^c = \bigcup_{g \neq e} gN$ is open, so that every open set is closed.

Example. *Consider $\mathcal{N} = \{\mathbf{Z}, 2\mathbf{Z}, 3\mathbf{Z}, \dots\}$. Then $\mathcal{N}$ induces a Hausdorff topology on $\mathbf{Z}$, such that $g_i \to g$, if and only if g_i is eventually in $g + n\mathbf{Z}$ for all n. In this topology, the series $1, 2, 3, \dots$ converges to zero!*

This example gives us a novel proof, due to Furstenburg, that there are infinitely many primes. Suppose that there were only finitely many, $\{p_1, p_2, \dots, p_n\}$. By the fundamental theorem of arithmetic,

$$\{-1, 1\} = (\mathbf{Z}p_1)^c \cap \dots \cap (\mathbf{Z}p_n)^c$$

and is therefore an open set. But this is clearly not the case as open sets must contain infinite sequences.

Chapter 7

The Haar Measure

One of the reasons that we isolate locally compact groups to study is that they possess an incredibly useful object allowing us to understand functions on the group, and thus the group itself. A **left (right) Haar measure** for a group G is a Radon measure μ for which $\mu(xE) = \mu(E)$ for any $x \in G$ and measurable E ($\mu(Ex) = \mu(E)$ for all x and E). For commutative groups, all left Haar measures are right Haar measures, but in non-commutative groups this need not hold. However, if μ is a right Haar measure, then $\nu(E) = \mu(E^{-1})$ is a left Haar measure, so there is no loss of generality in focusing our study on left Haar measures.

Example. *The example of a Haar measure that everyone knows is the Lebesgue measure on* $\mathbf{R}$ *(or* $\mathbf{R}^n$*). It commutes with translations because it is the measure induced by the linear functional corresponding to Riemann integration on* $C_c^+(\mathbf{R}^n)$*. A similar theory of Darboux integration can be applied to linearly ordered groups, leading to the construction of a Haar measure on such a group.*

Example. *If G is a Lie group, consider a* 2*-tensor* $g_e \in T_e^2(G)$ *inducing an inner product at the origin. Then the diffeomorphism* $f : a \mapsto b^{-1}a$ *allows us to consider* $g_b = f^*\lambda \in T_b^2(G)$*, and this is easily verified to be an inner product, hence we have a Riemannian metric. The associated Riemannian volume element can be integrated, producing a Haar measure on G.*

Example. *If G and H have Haar measures μ and ν, then $G \times H$ has a Haar measure $\mu \times \nu$, so that the class of topological groups with Haar measures is closed under the product operation. We can even allow infinite products, provided that the groups involved are compact, and the Haar measures are normalized to probability measures. This gives us probability measures on* F_2^ω *and*

$\mathbf{T}^\omega$, *which essentially measures the probability of an infinite sequence of coin flips.*

Example. *dx/x is a Haar measure for the multiplicative group of positive real numbers, since*

$$\int_a^b \frac{1}{x} = \log(b) - \log(a) = \log(cb) - \log(ca) = \int_{ca}^{cb} \frac{1}{x}$$

If we take the multiplicative group of all non-negative real numbers, the Haar measure becomes $dx/|x|$.

Example. *$dxdy/(x^2+y^2)$ is a Haar measure for the multiplicative group of complex numbers, since we have a basis of 'arcs' around the origin, and by a change of variables to polar coordinates, we verify the integral is changed by multiplication. Another way to obtain this measure is by noticing that $\mathbf{C}^\times$ is topologically isomorphic to the product of the circle group and the multiplicative group of real numbers, and hence the measure obtained should be the product of these measures. Since*

$$\frac{dxdy}{x^2+y^2} = \frac{drd\theta}{r}$$

We see that this is just the product of the Haar measure on $\mathbf{R}^+$, dr/r, and the Haar measure on $\mathbf{T}$, $d\theta$.

Example. *The space $M_n(\mathbf{R})$ of all n by n real matrices under addition has a Haar measure dM, which is essentially the Lebesgue measure on $\mathbf{R}^{n^2}$. If we consider the measure on $GL_n(\mathbf{R})$, defined by*

$$\frac{dM}{det(M)^n}$$

To see this, note the determinant of the map $M \mapsto NM$ on $M_n(\mathbf{R})$ is $det(N)^n$, because we can view $M_n(\mathbf{R})$ as the product of $\mathbf{R}^n$ n times, multiplication operates on the space componentwise, and the volume of the image of the unit paralelliped in each $\mathbf{R}^n$ is $det(N)$. Since the multiplicative group of complex numbers $z = x+iy$ can be identified with the group of matrices of the form

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

and the measure on $\mathbf{C} - \{0\}$ then takes the form $dM/det(M)$. More generally, if G is an open subset of $\mathbf{R}^n$, and left multiplication acts affinely, $xy = A(x)y + b(x)$, then $dx/|det(A(x))|$ is a left Haar measure on G, where dx is Lebesgue measure.

It turns out that there is a Haar measure on any locally compact group, and what's more, it is unique up to scaling. The construction of the measure involves constructing a positive linear functional $\phi : C_c(G) \to \mathbf{R}$ such that $\phi(L_x f) = \phi(f)$ for all x. The Riesz representation theorem then guarantees the existence of a Radon measure μ which represents this linear functional, and one then immediately verifies that this measure is a Haar measure.

Theorem 7.1. *Every locally compact group G has a Haar measure.*

Proof. The idea of the proof is fairly simple. If μ was a Haar measure, $f \in C_c^+(G)$ was fixed, and $\phi \in C_c^+(G)$ was a function supported on a small set, and behaving like a step function, then we could approximate f well by translates of ϕ,

$$f(x) \approx \sum c_i (L_{x_i} \phi)$$

Hence

$$\int f(x) d\mu \approx \sum c_i \int L_{x_i} \phi = \sum c_i \int \phi$$

If $\int \phi = 1$, then we could approximate $\int f(x) d\mu$ as literal sums of coefficients c_i. Since μ is outer regular, and ϕ is supported on neighbourhoods, one can show $\int f(x) d\mu$ is the infinum of $\sum c_i$, over all choices of $c_i > 0$ and $\int \phi \geqslant 1$, for which $f \leqslant \sum c_i L_{x_i} \phi$. Without the integral, we cannot measure the size of the functions ϕ, so we have to normalize by a different factor. We define $(f : \phi)$ to be the infinum of the sums $\sum c_i$, where $f \leqslant \sum c_i L_{x_i} \phi$ for some $x_i \in G$. We would then have

$$\int f d\mu \leqslant (f : \phi) \int \phi d\mu$$

If k is fixed with $\int k = 1$, then we would have

$$\int f d\mu \leqslant (f : \phi)(\phi : k)$$

We cannot change k if we wish to provide a limiting result in ϕ, so we notice that $(f:g)(g:h) \leqslant (f:h)$, which allows us to write

$$\int f\, d\mu \leqslant \frac{(f:\phi)}{(k:\phi)}$$

Taking the support of ϕ to be smaller and smaller, this value should approximate the integral perfectly accurately.

Define the linear functional

$$I_\phi(f) = \frac{(f:\phi)}{(k:\phi)}$$

Then I_ϕ is a sublinear, monotone, function with a functional bound

$$(k:f)^{-1} \leqslant I_\phi(f) \leqslant (f:k)$$

Which effectively says that, regardless of how badly we choose ϕ, the approximation factor $(f:\phi)$ is normalized by the approximation factor $(k:\phi)$ so that the integral is bounded. Now we need only prove that I_ϕ approximates a linear functional well enough that we can perform a limiting process to obtain a Haar integral. If $\varepsilon > 0$, and $g \in C_c^+(G)$ with $g = 1$ on $\text{supp}(f_1 + f_2)$, then the functions

$$h = f_1 + f_2 + \varepsilon g$$

$$h_1 = f_1/h \qquad h_2 = f_2/h$$

are in $C_0^+(G)$, if we define $h_i(x) = 0$ if $f_i(x) = 0$. This implies that there is a neighbourhood V of e such that if $x \in V$, and y is arbitrary, then

$$|h_1(xy) - h_1(y)| \leqslant \varepsilon \qquad |h_2(xy) - h_2(y)| < \varepsilon$$

If $\text{supp}(\phi) \subset V$, and $h \leqslant \sum c_i L_{x_i}\phi$, then

$$f_j(x) = h(x)h_j(x) \leqslant \sum c_i \phi(x_i x) h_j(x) \leqslant \sum c_i \phi(x_i x)\left[h_j(x_i^{-1}) + \varepsilon\right]$$

since we may assume that $x_i x \in \text{supp}(\phi) \subset V$. Then, because $h_1 + h_2 \leqslant 1$,

$$(f_1:\phi) + (f_2:\phi) \leqslant \sum c_j[h_1(x_j^{-1}) + \varepsilon] + \sum c_j[h_2(x_j^{-1}) + \varepsilon] \leqslant \sum c_j[1 + 2\varepsilon]$$

Now we find, by taking infinums, that

$$I_\phi(f_1) + I_\phi(f_2) \leqslant I_\phi(h)(1+2\varepsilon) \leqslant [I_\phi(f_1+f_2) + \varepsilon I_\phi(g)][1+2\varepsilon]$$

Since g is fixed, and we have a bound $I_\phi(g) \leqslant (g:k)$, we may always find a neighbourhood V (dependant on f_1, f_2) for any $\varepsilon > 0$ such that

$$I_\phi(f_1) + I_\phi(f_2) \leqslant I_\phi(f_1+f_2) + \varepsilon$$

if $\text{supp}(\phi) \subset V$.

Now we have estimates on how well I_ϕ approximates a linear function, so we can apply a limiting process. Consider the product

$$X = \prod_{f \in C_0^+(G)} [(k:f)^{-1}, (k:f_0)]$$

a compact space, by Tychonoff's theorem, consisting of $F : C_c^+(G) \to \mathbf{R}$ such that $(k:f)^{-1} \leqslant F(f) \leqslant (f:k)$. For each neighbourhood V of the identity, let $K(V)$ be the closure of the set of I_ϕ such that $\text{supp}(\phi) \subset V$. Then the set of all $K(V)$ has the finite intersection property, so we conclude there is some $I : C_c^+(G) \to \mathbf{R}$ contained in $\bigcap K(V)$. This means that every neighbourhood of I contains I_ϕ with $\text{supp}(\phi) \subset V$, for all ϕ. This means that if $f_1, f_2 \in C_c^+(G)$, $\varepsilon > 0$, and V is arbitrary, there is ϕ with $\text{supp}(\phi) \subset V$, and

$$|I(f_1) - I_\phi(f_1)| < \varepsilon \quad |I(f_2) - I_\phi(f_2)| < \varepsilon$$

$$|I(f_1+f_2) - I_\phi(f_1+f_2)| < \varepsilon$$

this implies that if V is chosen small enough, then

$$|I(f_1+f_2) - (I(f_1) - I(f_2))| \leqslant 2\varepsilon + |I_\phi(f_1+f_2) - (I_\phi(f_1) + I_\phi(f_2))| < 3\varepsilon$$

Taking $\varepsilon \to 0$, we conclude I is linear. Similar limiting arguments show that I is homogenous of degree 1, and commutes with all left translations. We conclude the extension of I to a linear functional on $C_0(G)$ is well defined, and the Radon measure obtained by the Riesz representation theorem is a Haar measure. □

We shall prove that the Haar measure is unique, but first we show an incredibly useful regularity property.

Proposition 7.2. *If U is open, and μ is a Haar measure, then $\mu(U) > 0$. It follows that if f is in $C_c^+(G)$, then $\int f\,d\mu > 0$.*

Proof. If $\mu(U) = 0$, then for any $x_1, \ldots, x_n \in G$,

$$\mu\left(\bigcup_{i=1}^{n} x_i U\right) \leqslant \sum_{i=1}^{n} \mu(x_i U) = 0$$

If K is compact, then K can be covered by finitely many translates of U, so $\mu(K) = 0$. But then $\mu = 0$ by regularity, a contradiction. □

Theorem 7.3. *Haar measures are unique up to a multiplicative constant.*

Proof. Let μ and ν be Haar measures. Fix a compact neighbourhood V of the identity. If $f, g \in C_c^+(G)$, consider the compact sets

$$A = \mathrm{supp}(f)V \cup V\mathrm{supp}(f) \qquad B = \mathrm{supp}(g)V \cup V\mathrm{supp}(g)$$

Then the functions $F_y(x) = f(xy) - f(yx)$ and $G_y(x) = g(xy) - g(yx)$ are supported on A and B. There is a neighbourhood $W \subset V$ of the identity such that $\|F_y\|_\infty, \|G_y\|_\infty < \varepsilon$ if $y \in W$. Now find $h \in C_c^+(G)$ with $h(x) = h(x^{-1})$ and $\mathrm{supp}(h) \subset W$ (take $h(x) = k(x)k(x^{-1})$ for some function $k \in C_c^+(G)$ with $\mathrm{supp}(k) \subset W$, and $k = 1$ on a symmetric neighbourhood of the origin). Then

$$\begin{aligned}\left(\int h\,d\mu\right)\left(\int f\,d\lambda\right) &= \int h(y)f(x)\,d\mu(y)\,d\lambda(x)\\ &= \int h(y)f(yx)\,d\mu(y)\,d\lambda(x)\end{aligned}$$

and

$$\begin{aligned}\left(\int h\,d\lambda\right)\left(\int f\,d\mu\right) &= \int h(x)f(y)\,d\mu(y)\,d\lambda(x)\\ &= \int h(y^{-1}x)f(y)\,d\mu(y)\,d\lambda(x)\\ &= \int h(x^{-1}y)f(y)\,d\mu(y)\,d\lambda(x)\\ &= \int h(y)f(xy)\,d\mu(y)\,d\lambda(x)\end{aligned}$$

Hence, applying Fubini's theorem,

$$\left|\int h d\mu \int f d\lambda - \int h d\lambda \int f d\mu\right| \leqslant \int h(y)|F_y(x)| d\mu(y) d\lambda(x)$$
$$\leqslant \varepsilon \lambda(A) \int h d\mu$$

In the same way, we find this is also true when f is swapped with g, and A with B. Dividing this inequalities by $\int h d\mu \int f d\mu$, we find

$$\left|\frac{\int f d\lambda}{\int f d\mu} - \frac{\int h d\lambda}{\int h d\mu}\right| \leqslant \frac{\varepsilon \lambda(A)}{\int f d\mu}$$

and this inequality holds with f swapped out with g, A with B. We then combine these inequalities to conclude

$$\left|\frac{\int f d\lambda}{\int f d\mu} - \frac{\int g d\lambda}{\int g d\mu}\right| \leqslant \varepsilon \left[\frac{\lambda(A)}{\int f d\mu} + \frac{\lambda(B)}{\int g d\mu}\right]$$

Taking ε to zero, we find $\lambda(A), \lambda(B)$ remain bounded, and hence

$$\frac{\int f d\lambda}{\int f d\mu} = \frac{\int g d\lambda}{\int g d\mu}$$

Thus there is a cosntant $c > 0$ such that $\int f d\lambda = c \int f d\mu$ for any function $f \in C_c^+(G)$, and we conclude that $\lambda = c\mu$. □

The theorem can also be proven by looking at the translation invariant properties of the derivative $f = d\mu/d\nu$, where $\nu = \mu + \lambda$ (We assume our group is σ compact for now). Consider the function $g(x) = f(yx)$. Then

$$\int_A g(x) d\nu = \int_{yA} f(x) d\nu = \mu(yA) = \mu(A)$$

so g is derivative, and thus $f = g$ almost everywhere. Our interpretation is that for a fixed y, $f(yx) = f(x)$ almost everywhere with respect to ν. Then (applying a discrete version of Fubini's theorem), we find that for almost all x with respect to ν, $f(yx) = f(x)$ holds for almost all y. But this implies that there exists an x for which $f(yx) = f(x)$ holds almost everywhere. Thus for any measurable A,

$$\mu(A) = \int_A f(y) d\nu(y) = f(x)\nu(A) = f(x)\mu(A) + f(x)\nu(A)$$

Now $(1 - f(x))\mu(A) = f(x)\nu(A)$ for all A, implying (since $\mu, \nu \neq 0$), that $f(x) \neq 0, 1$, and so

$$\frac{1 - f(x)}{f(x)}\mu(A) = \nu(A)$$

for all A. This shows the uniqueness property for all σ compact groups. If G is an arbitrary group with two measures μ and ν, then there is c such that $\mu = c\nu$ on every component of G, and thus on the union of countably many components. If A intersects uncountably many components, then either $\mu(A) = \nu(A) = \infty$, or the intersection of A on each set has positive measure on only countably many components, and in either case we have $\mu(A) = \nu(A)$.

7.1 Fubini, Radon Nikodym, and Duality

Before we continue, we briefly mention that integration theory is particularly nice over locally compact groups, even if we do not have σ finiteness. This essentially follows because the component of the identity in G is σ compact (take a compact neighbourhood and its iterated multiples), hence all components in G are σ compact. The three theorems that break down outside of the σ compact domain are Fubini's theorem, the Radon Nikodym theory, and the duality between $L^1(X)$ and $L^\infty(X)$. We show here that all three hold if X is a locally compact topological group.

First, suppose that $f \in L^1(G \times G)$. Then the essential support of f is contained within countably many components of $G \times G$ (which are simply products of components in G). Thus f is supported on a σ compact subset of $G \times G$ (as a locally compact topological group, each component of $G \times G$ is σ compact), and we may apply Fubini's theorem on the countably many components (the countable union of σ compact sets is σ compact). The functions in $L^p(G)$, for $1 \leqslant p < \infty$, also vanish outside of a σ compact subset (for if $f \in L^p(G)$, $|f|^p \in L^1(G)$ and thus vanishes outside of a σ compact set). What's more, all finite sums and products of functions from these sets (in either variable) vanish outside of σ compact subsets, so we almost never need to explicitly check the conditions for satisfying Fubini's theorem, and from now on we apply it wantonly.s

Now suppose μ and ν are both Radon measures, with $\nu \ll \mu$, and ν is σ-finite. By inner regularity, the support of ν is a σ compact set E. By inner regularity, μ restricted to E is σ finite, and so we may find a Radon

Nikodym derivative on E. This derivative can be extended to all of G because ν vanishes on G.

Finally, we note that $L^\infty(X) = L^1(X)^*$ can be made to hold if X is not σ finite, but locally compact and Hausdorff, provided we are integrating with respect to a Radon measure μ, and we modify $L^\infty(G)$ slightly. Call a set $E \subset X$ **locally Borel** if $E \cap F$ is Borel whenever F is Borel and $\mu(F) < \infty$. A locally Borel set is **locally null** if $\mu(E \cap F) = 0$ whenever $\mu(F) < \infty$ and F is Borel. We say a property holds **locally almost everywhere** if it is true except on a locally null set. $f : X \to \mathbf{C}$ is **locally measurable** if $f^{-1}(U)$ is locally Borel for every borel set $U \subset \mathbf{C}$. We now define $L^\infty(X)$ to be the space of all functions bounded except on a locally null set, modulo functions that are locally zero. That is, we define a norm

$$\|f\|_\infty = \inf\{c : |f(x)| \leqslant c \text{ locally almost everywhere}\}$$

and then $L^\infty(X)$ consists of the functions that have finite norm. It then follows that if $f \in L^\infty(X)$ and $g \in L^1(X)$, then g vanishes outside of a σ-finite set Y, so $fg \in L^1(X)$, and if we let $Y_1 \subset Y_2 \subset \dots \to Y$ be an increasing subsequence such that $\mu(Y_i) < \infty$, then $|f(x)| \leqslant \|f\|_\infty$ almost everywhere for $x \in Y_i$, and so by the monotone convergence theorem

$$\int |fg| d\mu = \lim_{Y_i \to \infty} \int_{Y_i} |fg| d\mu \leqslant \|f\|_\infty \int_{Y_i} |g| d\mu \leqslant \|f\|_\infty \|g\|_1$$

Thus the map $g \mapsto \int fg d\mu$ is a well defined, continuous linear functional with norm $\|f\|_\infty$. That $L^1(X)^* = L^\infty(X)$ follows from the decomposibility of the Carathéodory extension of μ, a fact we leave to the general measure theorists.

7.2 Unimodularity

We have thus defined a left invariant measure, but make sure to note that such a function is not right invariant. We call a group who's left Haar measure is also right invariant **unimodular**. Obviously all abelian groups are unimodular.

Given a fixed y, the measure $\mu_y(A) = \mu(Ay)$ is a new Haar measure on the space, hence there is a constant $\Delta(y) > 0$ depending only on y such that $\mu(Ay) = \Delta(y)\mu(A)$ for all measurable A. Since $\mu(Axy) = \Delta(y)\mu(Ay) =$

$\Delta(x)\Delta(y)\mu(A)$, we find that $\Delta(xy) = \Delta(x)\Delta(y)$, so Δ is a homomorphism from G to the multiplicative group of real numbers. For any $f \in L^1(\mu)$, we have

$$\int f(xy)d\mu(x) = \Delta(y^{-1}) \int f(x)d\mu(x)$$

If $y_i \to e$, and $f \in C_c(G)$, then $\|R_{y_i}f - f\|_\infty \to 0$, so

$$\Delta(y_i^{-1}) \int f(x)d\mu = \int f(xy_i)d\mu \to \int f(x)d\mu$$

Hence $\Delta(y_i^{-1}) \to 1$. This implies Δ, known as the unimodular function, is a continuous homomorphism from G to the real numbers. Note that Δ is trivial if and only if G is unimodular.

Theorem 7.4. *Any compact group is unimodular.*

Proof. $\Delta : G \to \mathbf{R}^*$ is a continuous homomorphism, hence $\Delta(G)$ is compact. But the only compact subgroup of $\mathbf{R}$ is trivial, hence Δ is trivial. □

Let G^c be the smallest closed subgroup of G containing the commutators $[x,y] = xyx^{-1}y^{-1}$. It is verified to be a normal subgroup of G by simple algebras.

Theorem 7.5. *If G/G^c is compact, then G is unimodular.*

Proof. Δ factors through G/G^c since it is abelian. But if Δ is trivial on G/G^c, it must also be trivial on G. □

The modular function relates right multiplication to left multiplcation in the group. In particular, if $d\mu$ is a Left Haar measure, then $\Delta^{-1}d\mu$ is a right Haar measure. Hence any right Haar measure is a constant multiple of $\Delta^{-1}d\mu$. Hence the measure $\nu(A) = \mu(A^{-1})$ has a value c such that for any function f,

$$\int \frac{f(x)}{\Delta(x)}d\mu(x) = c\int f(x)d\nu(x) = c\int f(x^{-1})d\mu$$

If $c \neq 1$, pick a symmetric neighbourhood U such that for $x \in U$, $|\Delta(x) - 1| \leqslant \varepsilon|c-1|$. Then if $f > 0$

$$|c-1|\mu(U) = |c\mu(U^{-1}) - \mu(U)| = \left|\int_U [\Delta(x^{-1}) - 1]d\mu(x)\right| \leqslant \varepsilon\mu(U)|c-1|$$

A contradiction if $\varepsilon < 1$. Thus we have

$$\int f(x^{-1}) d\mu(x) = \int \frac{f(x)}{\Delta(x)} d\mu(x)$$

A useful integration trick. When Δ is unbounded, then it follows that $L^p(\mu)$ and $L^p(\nu)$ do not consist of the same functions. There are two ways of mapping the sets isomorphically onto one another – the map $f(x) \mapsto f(x^{-1})$, and the map $f(x) \mapsto \Delta(x)^{1/p} f(x)$.

From now on, we assume a left invariant Haar measure is fixed over an entire group. Since a Haar measure is uniquely determined up to a constant, this is no loss of generality, and we might as well denote our integration factors $d\mu(x)$ and $d\mu(y)$ as dx and dy, where it is assumed that this integration is over the Lebesgue measure.

7.3 Convolution

If G is a topological group, then $C(G)$ does not contain enough algebraic structure to identify G – for instance, if G is a discrete group, then $C(G)$ is defined solely by the cardinality of G. The algebras we wish to study over G is the space $M(G)$ of all complex valued Radon measures over G and the space $L^1(G)$ of integrable functions with respect to the Haar measure, because here we can place a Banach algebra structure with an involution. We note that $L^1(G)$ can be isometrically identified as the space of all measures $\mu \in M(G)$ which are absolutely continuous with respect to the Haar measure. Given $\mu, \nu \in M(G)$, we define the convolution measure

$$\int \phi d(\mu * \nu) = \int \phi(xy) d\mu(x) d\nu(y)$$

The measure is well defined, for if $\phi \in C_c^+(X)$ is supported on a compact set K, then

$$\begin{aligned} \left| \int \phi(xy) d\mu(x) d\nu(y) \right| &\leqslant \int_G \int_G \phi(xy) d|\mu|(x) d|\nu|(y) \\ &\leqslant \|\mu\| \|\nu\| \|\phi\|_\infty \end{aligned}$$

This defines an operation on $M(G)$ which is associative, since, by applying the associativity of G and Fubini's theorem.

$$\begin{aligned}\int \phi d((\mu*\nu)*\lambda) &= \int\int \phi(xz)d(\mu*\nu)(x)d\lambda(z)\\ &= \int\int\int \phi((xy)z)d\mu(x)d\nu(y)d\lambda(z)\\ &= \int\int\int \phi(x(yz))d\mu(x)d\nu(y)d\lambda(z)\\ &= \int\int \phi(xz)d\mu(x)d(\nu*\lambda)(z)\\ &= \int \phi d(\mu*(\nu*\lambda))\end{aligned}$$

Thus we begin to see how the structure of G gives us structure on $M(G)$. Another example is that convolution is commutative if and only if G is commutative. We have the estimate $\|\mu*\nu\| \leqslant \|\mu\|\|\nu\|$, because of the bound we placed on the integrals above. $M(G)$ is therefore an involutive Banach algebra, which has a unit, the dirac delta measure at the identity.

As a remark, we note that involutive Banach algebras have nowhere as near a nice of a theory than that of C^* algebras. $M(G)$ cannot be renormed to be a C^* algebra, since every weakly convergent Cauchy sequence converges, which is impossible in a C^* algebra, except in the finite dimensional case.

A **discrete measure** on G is a measure in $M(G)$ which vanishes outside a countable set of points, and the set of all such measures is denoted $M_d(G)$. A **continuous measure** on G is a measure μ such that $\mu(\{x\}) = 0$ for all $x \in G$. We then have a decomposition $M(G) = M_d(G) \oplus M_c(G)$, for if μ is any measure, then $\mu(\{x\}) \neq 0$ for at most countably many points x, for

$$\|\mu\| \geqslant \sum_{x\in G} |\mu|(x)$$

This gives rise to a discrete measure ν, and $\mu - \nu$ is continuous. If we had another decomposition, $\mu = \psi + \phi$, then $\mu(\{x\}) = \psi(\{x\}) = \nu(\{x\})$, so $\psi = \nu$ by discreteness, and we then conclude $\phi = \mu - \nu$. $M_c(G)$ is actually a closed subspace of $M(G)$, since if $\mu_i \to \mu$, and $\mu_i \in M_c(G)$, and $\|\mu_i - \mu\| < \varepsilon$, then for any $x \in G$,

$$\varepsilon > \|\mu - \mu_i\| \geqslant |(\mu_i - \mu)(\{x\})| = |\mu(\{x\})|$$

Letting $\varepsilon \to 0$ shows continuity.

The convolution on $M(G)$ gives rise to a convolution on $L^1(G)$, where

$$(f * g)(x) = \int f(y)g(y^{-1}x)dy$$

which satisfies $\|f * g\|_1 \leqslant \|f\|_1\|g\|_1$. This is induced by the identification of f with $f(x)dx$, because then

$$\begin{aligned}\int \phi(f(x)dx * g(x)dx) &= \int\int \phi(yx)f(y)g(x)dydx \\ &= \int \phi(y)\left(\int f(y)g(y^{-1}x)dx\right)dy\end{aligned}$$

Hence $fd\mu * gd\mu = (f * g)d\mu$. What's more,

$$\|f\|_1 = \|fd\mu\|$$

If $\nu \in M(G)$, then we can still define $\nu * f \in L^1(G)$

$$(\nu * f)(x) = \int f(y^{-1}x)d\mu(y)$$

which holds since

$$\int \phi d(\nu * f\mu) = \int \phi(yx)f(x)d\nu(y)d\mu(x) = \int \phi(x)f(y^{-1}x)d\nu(y)d\mu(x)$$

If G is unimodular, then we also find

$$\int \phi d(f\mu * \nu) = \int \phi(yx)f(y)d\mu(y)d\nu(x) = \int \phi(x)f(y)d\mu(y)d\nu(y^{-1}x)$$

So we let $f * \mu(x) = \int f(y)d\mu(y^{-1}x)$.

Theorem 7.6. *$L^1(G)$ and $M_c(G)$ are closed ideals in $M(G)$, and $M_d(G)$ is a closed subalgebra.*

Proof. If $\mu_i \to \mu$, and each μ_i is discrete, the μ is discrete, because there is a countable set K such that all μ_i are equal to zero outside of K, so μ must also vanish outside of K (here we have used the fact that $M(G)$ is a Banach space, so that we need only consider sequences). Thus $M_d(G)$ is closed, and is easily verified to be subalgebra, essentially because $\delta_x * \delta_y = \delta_{xy}$. If $\mu_i \to \mu$, then $\mu_i(\{x\}) \to \mu(\{x\})$, so that $M_c(G)$ is closed in $M(G)$. If ν is an arbitrary measure, and μ is continuous, then

$$(\mu * \nu)(\{x\}) = \int_G \mu(\{y\}) d\nu(y^{-1}x) = 0$$

$$(\nu * \mu)(\{x\}) = \int_G \mu(\{y\}) d\nu(xy^{-1}) = 0$$

so $M_c(G)$ is an ideal. Finally, we verify $L^1(G)$ is closed, because it is complete, and if $\nu \in M(G)$ is arbitrary, and if U has null Haar measure, then

$$(f dx * \nu)(U) = \int \chi_U(xy) f(x) dx\, d\nu(y) = \int_G \int_{y^{-1}U} f(x) dx d\nu(y) = 0$$

$$(\nu * f dx)(U) = \int \chi_U(xy) d\nu(x) f(y) dy = \int_G \int_{Ux^{-1}} f(y) dy d\nu(x) = 0$$

So $L^1(G)$ is a two-sided ideal. □

If we wish to integrate by right multiplication instead of left multiplication, we find by the substitution $y \mapsto xy$ that

$$\begin{aligned}(f * g)(x) &= \int f(y) g(y^{-1}x) dy \\ &= \int\int f(xy) g(y^{-1}) dy \\ &= \int\int \frac{f(xy^{-1}) g(y)}{\Delta(y)} dy\end{aligned}$$

Observe that

$$f * g = \int f(y) L_{y^{-1}} g\, dy$$

which can be interpreted as a vector valued integral, since for $\phi \in L^\infty(\mu)$,

$$\int (f * g)(x) \phi(x) dx = \int f(y) g(y^{-1}x) \phi(x) dx dy$$

so we can see convolution as a generalized 'averaging' of translate of g with respect to the values of f. If G is commutative, this is the same as the averaging of translates of f, but not in the noncommutative case. It then easily follows from operator computations $L_z(f*g) = (L_z f)*g$, and $R_z(f*g) = f*(R_z g)$, or from the fact that

$$(f*g)(zx) = \int f(y)g(y^{-1}zx)dy = \int f(zy)g(y^{-1}x)dy = [(L_z f)*g](x)$$

$$(f*g)(xz) = \int f(y)g(y^{-1}xz)dy = [f*(R_z g)](x)$$

Convolution can also be applied to the other L^p spaces, but we have to be a bit more careful with our integration.

Theorem 7.7. *If $f \in L^1(G)$ and $g \in L^p(G)$, then $f*g$ is defined for almost all x, $f*g \in L^p(G)$, and $\|f*g\|_p \leqslant \|f\|\|g\|_p$. If G is unimodular, then the same results hold for $g*f$, or if G is not unimodular and f has compact support.*

Proof. We use Minkowski's inequality to find

$$\begin{aligned}
\|f*g\|_p &= \left(\int \left|\int f(y)|g(y^{-1}x)dy\right|^p dx\right)^{1/p} \\
&\leqslant \int |f(y)| \left(\int |g(y^{-1}x)|^p dx\right)^{1/p} dy \\
&= \|f\|_1 \|g\|_p
\end{aligned}$$

If G is unimodular, then

$$\|g*f\|_p = \left(\int \left|\int g(xy^{-1})f(y)dy\right|^p dx\right)^{1/p}$$

and we may apply the same trick as used before.

If f has compact support K, then $1/\Delta$ is bounded above by $M > 0$ on K and

$$\begin{aligned}
\|g*f\|_p &= \left(\int \left|\int \frac{g(xy^{-1})f(y)}{\Delta(y)}dy\right|^p dx\right)^{1/p} \\
&\leqslant \int \left(\int \left|\frac{g(xy^{-1})f(y)}{\Delta(y)}\right|^p dx\right)^{1/p} dy \\
&= \|g\|_p \int_K \frac{|f(y)|}{\Delta(y)} d\mu(y) \\
&\leqslant M\|g\|_p\|f\|_1
\end{aligned}$$

which shows that $g * f$ is defined almost everywhere. $\square$

Theorem 7.8. *If G is unimodular, $f \in L^p(G)$, $g \in L^q(G)$, and $p = q^*$, then $f * g \in C_0(G)$ and $\|f * g\|_\infty \leqslant \|f\|_p \|g\|_q$.*

Proof. First, note that

$$\begin{aligned}|(f * g)(x)| &\leqslant \int |f(y)||g(y^{-1}x)|dy \\ &\leqslant \|f\|_p \left(\int |g(y^{-1}x)|^q dy \right)^{1/q} \\ &= \|f\|_p \|g\|_q\end{aligned}$$

For each x and y, applying Hölder's inequality, we find

$$\begin{aligned}|(f * g)(x) - (f * g)(y)| &\leqslant \int |f(z)||g(z^{-1}x) - g(z^{-1}y)|dz \\ &\leqslant \|f\|_p \left(\int |g(z^{-1}x) - g(z^{-1}y)|^q dz \right)^{1/q} \\ &= \|f\|_p \left(\int |g(z) - g(zx^{-1}y)|^q dz \right)^{1/q} \\ &= \|f\|_p \|g - R_{x^{-1}y} g\|_q\end{aligned}$$

Thus to prove continuity (and in fact uniform continuity), we need only prove that $\|g - R_x g\|_q \to 0$ for $q \neq \infty$ as $x \to \infty$ or $x \to 0$. This is the content of the next lemma. $\square$

We now show that the map $x \mapsto L_x$ is a continuous operation from G to the weak $*$ topology on the L_p spaces, for $p \neq \infty$. It is easily verified that translation is not continuous on L_∞, by taking a suitable bumpy function.

Theorem 7.9. *If $p \neq \infty$, then $\|g - R_x g\|_p \to 0$ and $\|g - L_x g\|_p \to 0$ as $x \to 0$.*

Proof. If $g \in C_c(G)$, then one verifies the theorem by using left and right uniform continuity. In general, we let $g_i \in C_c(G)$ be a sequence of functions converging to g in the L_p norm, and we then find

$$\|g - L_x g\|_p \leqslant \|g - g_i\|_p + \|g_i - L_x g_i\|_p + \|L_x(g_i - g)\|_p = 2\|g - g_i\|_p + \|g_i - L_x g_i\|_p$$

Taking i large enough, x small enough, we find $\|g - L_x g\|_p \to 0$. The only problem for right translation is the appearance of the modular function

$$\|R_x(g - g_i)\|_p = \frac{\|g - g_i\|_p}{\Delta(x)^{1/p}}$$

If we assume our x values range only over a compact neighbourhood K of the origin, we find that $\Delta(x)$ is bounded below, and hence $\|R_x(g-g_i)\|_p \to 0$, which effectively removes the problems in the proof. □

Since the map is linear, we have verified that the map $x \mapsto L_x f$ is uniformly continuous in L^p for each $f \in L^p$. In the case where $p = \infty$, the same theorem cannot hold, but we have even better conditions that do not even require unimodularity.

Theorem 7.10. *If $f \in L^1(G)$ and $g \in L^\infty(G)$, then $f * g$ is left uniformly continuous, and $g * f$ is right uniformly continuous.*

Proof. We have

$$\|L_z(f * g) - (f * g)\|_\infty = \|(L_z f - f) * g\|_\infty \leqslant \|L_z f - f\|_1 \|g\|_\infty$$

$$\|R_z(g * f) - (g * f)\|_\infty = \|g * (R_z f - f)\|_\infty \leqslant \|g\|_\infty \|R_z f - f\|_1$$

and both integrals converge to zero as $z \to 1$. □

The passage from $M(G)$ to $L^1(G)$ removes an identity from the Banach algebra in question (except if G is discrete), but there is always a way to approximate an identity.

Theorem 7.11. *For each neighbourhood U of the origin, pick a function $f_U \in (L^1)^+(G)$, with $\int \phi_U = 1$, $supp(f_U) \subset U$. Then if g is any function in $L^p(G)$,*

$$\|f_U * g - g\|_p \to 0$$

where we assume g is left uniformly continuous if $p = \infty$, and if f_U is viewed as a net with neighbourhoods ordered by inclusion. If in addition $f_U(x) = f_U(x^{-1})$, then $\|g * f_U - g\|_p \to 0$, where g is right uniformly continuous for $p = \infty$.

Proof. Let us first prove the theorem for $p \neq \infty$. If $g \in C_c(G)$ is supported on a compact K, and if U is small enough that $|g(y^{-1}x) - g(x)| < \varepsilon$ for $y \in U$, then because $\int_U f_U(y) = 1$, and by applying Minkowski's inequality, we find

$$\begin{aligned}
\|f_U * g - g\|_p &= \left(\int \left| \int f_U(y)[g(y^{-1}x) - g(x)] dy \right|^p dx \right)^{1/p} \\
&\leqslant \int f_U(y) \left(\int |g(y^{-1}x) - g(x)|^p dx \right)^{1/p} dy \\
&\leqslant 2\mu(K)\varepsilon \int f_U(y) dy \leqslant 2\mu(K)\varepsilon
\end{aligned}$$

Results are then found for all of L^p by taking limits. If g is left uniformly continuous, then we may find U such that $|g(y^{-1}x) - g(x)| < \varepsilon$ for $y \in U$ then

$$|(f_U * g - g)(x)| = \left| \int f_U(y)[g(y^{-1}x) - g(x)] \right| \leqslant \varepsilon$$

For right convolution, we find that for $g \in C_c(G)$, where $|g(xy) - g(x)| < \varepsilon$ for $y \in U$, then

$$\begin{aligned}
\|g * f_U - g\|_p &= \left(\int \left| \int g(y) f_U(y^{-1}x) - g(x) dy \right|^p dx \right)^{1/p} \\
&= \left(\int \left| \int [g(xy) - g(x)] f_U(y) dy \right|^p dx \right)^{1/p} \\
&\leqslant \int \left(\int |g(xy) - g(x)|^p dx \right)^{1/p} f_U(y) dy \\
&\leqslant \mu(K)\varepsilon \int f_U(y)(1 + \Delta(y)) dy \\
&= \mu(K)\varepsilon + \mu(K)\varepsilon \int f_U(y)\Delta(y) dy
\end{aligned}$$

We may always choose U small enough that $\Delta(y) < \varepsilon$ for $y \in U$, so we obtain a complete estimate $\mu(K)(\varepsilon + \varepsilon^2)$. If g is right uniformly continuous, then choosing U for which $|g(xy) - g(x)| < \varepsilon$, then

$$|(g * f_U - g)(x)| = \left| \int [g(xy) - g(x)] f_U(y) dy \right| \leqslant \varepsilon$$

We will always assume from hereon out that the approximate identities in $L^1(G)$ are of this form. □

We have already obtained enough information to characterize the closed ideals of $L^1(G)$.

Theorem 7.12. *If V is a closed subspace of $L^1(G)$, then V is a left ideal if and only if it is closed under left translations, and a right ideal if and only if it is closed under right translations.*

Proof. If V is a closed left ideal, and f_U is an approximate identity at the origin, then for any g,

$$\|(L_z f_U) * g - L_z g\|_1 = \|L_z(f_U * g - g)\|_1 = \|f_U * g - g\| \to 0$$

so $L_z g \in V$. Conversely, if V is closed under left translations, $g \in L^1(G)$, and $f \in V$, then

$$g * f = \int g(y) L_{y^{-1}} f \, dy$$

which is in the closed linear space of the translates of f. Right translation is verified very similarily. □

7.4 The Riesz Thorin Theorem

We finalize our basic discussion by looking at convolutions of functions in $L^p * L^q$. Certainly $L^p * L^1 \subset L^p$, and $L^p * L^q \subset L^\infty$ for $q = p^*$. To prove general results, we require a foundational interpolation result.

Theorem 7.13. *For any $0 < \theta < 1$, and $0 < p, q \leqslant \infty$. If we define*

$$1/r_\theta = (1-\theta)/p + \theta/q$$

to be the inverse interpolation of the two numbers. Then

$$\|f\|_{r_\theta} \leqslant \|f\|_p^{1-\theta} \|f\|_q^\theta$$

Proof. We apply Hölder's inequality to find

$$\|f\|_{r_\theta} \leqslant \|f\|_{p/(1-\theta)} \|f\|_{q/\theta} = \left(\int |f|^{p/(1-\theta)} \right)^{(1-\theta)/p} \left(\int |f|^{q/\theta} \right)^{\theta/q}$$

so it suffices to prove $\|f\|_{p/(1-\theta)} \leqslant \|f\|_p^{1-\theta}$, $\|f\|_{q/\theta} \leqslant \|f\|_q^\theta$.

The map $x \mapsto x^p$ is concave for $0 < p < 1$, so we may apply Jensen's inequality in reverse to conclude

$$\left(\int |f|^{p/(1-\theta)} \right)^{(1-\theta)/p} \leqslant \left(\int |f|^p \right)^{1/p}$$

□

The Riesz Thorin interpolation theorem then implies $L^p * L^q \subset L^r$, for $p^{-1} + q^{-1} = 1 + r^{-1}$. However, these estimates only guarantee $L^1(G)$ is closed under convolution. If G is compact, then $L_p(G)$ is closed under convolution for all p (TODO). The L_p conjecture says that this is true if and only if G is compact. This was only resolved in 1990.

7.5 Homogenous Spaces and Haar Measures

The natural way for a locally compact topological group G to act on a locally compact Hausdorff space X is via a representation of G in the homeomorphisms of X. We assume the action is transitive on X. The standard example are the action of G on G/H, where H is a closed subspace. These are effectively all examples, because if we fix $x \in X$, then the map $y \mapsto yx$ induces a continuous bijection from G/H to X, where H is the set of all y for which $yx = x$. If G is a σ compact space, then this map is a homeomorphism.

Theorem 7.14. *If a σ compact topological group G has a transitive topological action on X, and $x \in X$, then the continuous bijection from G/G_x to X is a homeomorphism.*

Proof. It suffices to show that the map $\phi : G \to X$ is open, and we need only verify this for the neighbourhood basis of compact neighbourhoods V of the origin by properties of the action. G is covered by countably many translates $y_1 V, y_2 V, \dots$, and since each $\phi(y_k V) = y_k \phi(V)$ is closed (compactness), we conclude that $y_k \phi(V)$ has non-empty interior for some y_k, and hence $\phi(V)$ has a non-empty interior point $\phi(y_0)$. But then for any $y \in V$, y is in the interior of $\phi(y V y_0^{-1}) \subset \phi(V V y_0^{-1})$, so if we fix a compact U, and find V with $V^3 \subset U$, we have shown $\phi(U)$ is open in X. □

We shall say a space X is homogenous if it is homeomorphic to G/H for some group action of G over X. The H depends on our choice of basepoint x, but only up to conjugation, for if if we switch to a new basepoint y, and c maps x to y, then $ax = x$ holds if and only if $cac^{-1}y = y$. The question here is to determine whether we have a G-invariant measure on X. This is certainly not always possible. If we had a measure on $\mathbf{R}$ invariant under the affine maps $ax + b$, then it would be equal to the Haar measure by uniqueness, but the Haar measure is not invariant under dilation $x \mapsto ax$.

Let G and H have left Haar measures μ and ν respectively, denote the projection of G onto G/H as $\pi : G \to G/H$, and let Δ_G and Δ_H be the respective modular functions. Define a map $P : C_c(G) \to C_c(G/H)$ by

$$(Pf)(Hx) = \int_H f(xy) d\nu(y) = \int_H$$

this is well defined by the invariance properties of ν. Pf is obviously continuous, and $\operatorname{supp}(Pf) \subset \pi(\operatorname{supp}(f))$. Moreover, if $\phi \in C(G/H)$ we have

$$P((\phi \circ \pi) \cdot f)(Hx) = \phi(xH) \int_H f(xy) d\nu(y)$$

so $P((\phi \circ \pi) \cdot f) = \phi P(f)$.

Lemma 7.15. *If E is a compact subset of G/H, there is a compact $K \subset G$ with $\pi(K) = E$.*

Proof. Let V be a compact neighbourhood of the origin, and cover E by finitely many translates of $\pi(V)$. We conclude that $\pi^{-1}(E)$ is covered by finitely many of the translates, and taking the intersections of these translates with $\pi^{-1}(E)$ gives us the desired K. □

Lemma 7.16. *A compact $F \subset G/H$ gives rise to a function $f \geqslant 0$ in $C_c(G)$ such that $Pf = 1$ on E.*

Proof. Let E be a compact neighbourhood containing F, and if $\pi(K) = E$, there is a function $g \in C_c(G)$ with $g > 0$ on K, and $\phi \in C_c(G/H)$ is supported on E and $\phi(x) = 1$ for $x \in F$, let

$$f = \frac{\phi \circ \pi}{Pg \circ \pi} g$$

Hence

$$Pf = \frac{\phi}{Pg} Pg = \phi$$

$\square$

Lemma 7.17. *If $\phi \in C_c(G/H)$, there is $f \in C_c(G)$ with $Pf = \phi$, and $\pi(\mathrm{supp} f) = \mathrm{supp}(\phi)$, and also $f \geqslant 0$ if $\phi \geqslant 0$.*

Proof. There exists $g \geqslant 0$ in $C_c(G/H)$ with $Pg = 1$ on $\mathrm{supp}(\phi)$, and then $f = (\phi \circ \pi) g$ satisfies the properties of the theorem. $\square$

We can now provide conditions on the existence of a measure on G/H.

Theorem 7.18. *There is a G invariant measure ψ on G/H if and only if $\Delta_G = \Delta_H$ when restricted to H. In this case, the measure is unique up to a common factor, and if the factor is chosen, we have*

$$\int_G f \, d\mu = \int_{G/H} Pf \, d\psi = \int_{G/H} \int_H f(xy) d\nu(y) d\psi(xH)$$

Proof. Suppose ψ existed. Then $f \mapsto \int Pf \, d\psi$ is a non-zero left invariant positive linear functional on G/H, so $\int Pf \, d\psi = c \int f \, d\mu$ for some $c > 0$. Since $P(C_c(G)) = C_c(G/H)$, we find that ψ is determined up to a constant factor. We then compute, for $y \in H$,

$$\begin{aligned}
\Delta_G(y) \int f(x) d\mu(x) &= \int f(xy^{-1}) d\mu(x) \\
&= \int_{G/H} \int_H f(xzy^{-1}) d\nu(z) d\psi(xH) \\
&= \Delta_H(y) \int_{G/H} \int_H f(xz) d\nu(z) d\psi(xH) \\
&= \Delta_H(y) \int f(x) d\mu(x)
\end{aligned}$$

Hence $\Delta_G = \Delta_H$. Conversely, suppose $\Delta_G = \Delta_H$. First, we claim if $f \in C_c(G)$ and $Pf = 0$, then $\int f \, d\mu = 0$. Indeed if $P\phi = 1$ on $\pi(\mathrm{supp} f)$ then

$$0 = Pf(xH) = \int_H f(xy) d\nu(y) = \Delta_G(y^{-1}) \int_H f(xy^{-1}) d\nu(y)$$

so

$$
\begin{aligned}
0 &= \int_G \int_H \Delta_G(y^{-1})\phi(x) f(xy^{-1}) d\nu(y) d\mu(x) \\
&= \int_H \int_G \phi(xy) f(x) d\mu(x) d\nu(y) \\
&= \int_G P\phi(xH) f(x) d\mu(x) \\
&= \int_G f(x) d\mu(x)
\end{aligned}
$$

This implies that if $Pf = Pg$, then $\int_G f = \int_G g$. Thus the map $Pf \mapsto \int_G f$ is a well defined G invariant positive linear functional on $C_c(G/H)$, and we obtain a Radon measure from the Riesz representation theorem. □

If H is compact, then Δ_G and Δ_H are both continuous homomorphisms from H to $\mathbf{R}^+$, so Δ_G and Δ_H are both trivial, and we conclude a G invariant measure exists on G/H.

7.6 Function Spaces In Harmonic Analysis

There are a couple other function spaces that are interesting in Harmonic analysis. We define $\text{AP}(G)$ to be the set of all almost periodic functions, functions $f \in L^\infty(G)$ such that $\{L_x f : x \in G\}$ is relatively compact in $L^\infty(G)$. If this is true, then $\{R_x f : x \in G\}$ is also relatively compact, a rather deep theorem. If we define $\text{WAP}(G)$ to be the space of weakly almost periodic functions (the translates are relatively compact in the weak topology). It is a deep fact that $\text{WAP}(G)$ contains $C_0(G)$, but $\text{AP}(G)$ can be quite small. The reason these function spaces are almost periodic is that in the real dimensional case, $\text{AP}(\mathbf{R})$ is just the closure of the set of all trigonometric polynomials.

Chapter 8

The Character Space

Let G be a locally compact group. A character on G is a *continuous* homomorphism from G to $\mathbf{T}$. The space of all characters of a group will be denoted $\Gamma(G)$.

Example. *Determining the characters of* $\mathbf{T}$ *involves much of classical Fourier analysis. Let* $f : \mathbf{T} \to \mathbf{T}$ *be an arbitrary continuous character. For each* $w \in \mathbf{T}$*, consider the function* $g(z) = f(zw) = f(z)f(w)$*. We know the Fourier series acts nicely under translation, telling us that*

$$\hat{g}(n) = w^n \hat{f}(n)$$

Conversely, since $g(z) = f(z)f(w)$*,*

$$\hat{g}(n) = f(w)\hat{f}(n)$$

Thus $(w^n - f(w))\hat{f}(n) = 0$ *for all* $w \in \mathbf{T}$*,* $n \in \mathbf{Z}$*. Fixing* n*, we either have* $f(w) = w^n$ *for all* w*, or* $\hat{f}(n) = 0$*. This implies that if* $f \neq 0$*, then* f *is just a power map for some* $n \in \mathbf{Z}$*.*

Example. *The characters of* $\mathbf{R}$ *are of the form* $t \mapsto e^{ti\xi}$*, for* $\xi \in \mathbf{R}$*. To see this, let* $e : \mathbf{R} \to \mathbf{T}$ *be an arbitrary character. Define*

$$F(x) = \int_0^x e(t)dt$$

Then $F'(x) = e(x)$*. Since* $e(0) = 1$*, for suitably small* δ *we have*

$$F(\delta) = \int_0^\delta e(t)dt = c > 0$$

and then it follows that

$$F(x+\delta)-F(x)=\int_x^{x+\delta} e(t)dt=\int_0^{\delta} e(x+t)dt=ce(x)$$

As a function of x, F is differentiable, and by the fundamental theorem of calculus,

$$\frac{dF(x+\delta)-F(x)}{dt}=F'(x+\delta)-F'(x)=e(x+\delta)-e(x)$$

This implies the right side of the above equation is differentiable, and so

$$ce'(x)=e(x+\delta)-e(x)=e(x)[e(\delta)-1]$$

Implying $e'(x)=Ae(x)$ for some $A\in\mathbf{C}$, so $e(x)=e^{Ax}$. We require that $e(x)\in\mathbf{T}$ for all x, so $A=\xi i$ for some $\xi\in\mathbf{R}$.

Example. *Consider the group $\mathbf{R}^+$ of positive real numbers under multiplication. The map $x\mapsto\log x$ is an isomorphism from $\mathbf{R}^+$ and $\mathbf{R}$, so that every character on $\mathbf{R}^+$ is of the form $e^{is\log(x)}=x^{is}$, for some $s\in\mathbf{R}$. The character group is then $\mathbf{R}$, since $x^{is}x^{is'}=x^{i(s+s')}$.*

There is a connection between characters on G and characters on $L^1(G)$ that is invaluable to the generalization of Fourier analysis to arbitrary groups.

Theorem 8.1. *For any character $\phi: G\to\mathbf{C}$, the map*

$$\varphi(f)=\int\frac{f(x)}{\phi(x)}dx$$

is a non-zero character on the convolution algebra $L^1(G)$, and all characters arise this way.

Proof. The induced map is certainly linear, and

$$\begin{aligned}\varphi(f*g)&=\int\int\frac{f(y)g(y^{-1}x)}{\phi(x)}dydx\\&=\int\int\frac{f(y)g(x)}{\phi(y)\phi(x)}dydx\\&=\int\frac{f(y)}{\phi(y)}dy\int\frac{g(x)}{\phi(x)}dx\end{aligned}$$

Since ϕ is continuous, there is a compact subset K of G where $\phi > \varepsilon$ for some $\varepsilon > 0$, and we may then choose a positive f supported on K in such a way that $\varphi(f)$ is non-zero.

The converse results from applying the duality theory of the L^p spaces. Any character on $L^1(G)$ is a linear functional, hence is of the form

$$f \mapsto \int f(x)\phi(x)dx$$

for some $\phi \in L^\infty(G)$. Now

$$\begin{aligned}\int\int f(y)g(x)\phi(yx)dydx &= \int\int f(y)g(y^{-1}x)\phi(x)dydx \\ &= \int f(x)\phi(x)dx \int g(y)\phi(y)dy \\ &= \int f(x)g(y)\phi(x)\phi(y)dxdy\end{aligned}$$

Since this holds for all functions f and g in $L^1(G)$, we must have $\phi(yx) = \phi(x)\phi(y)$ almost everywhere. Also

$$\begin{aligned}\int \varphi(f)g(y)\phi(y)dy &= \varphi(f * g) \\ &= \int\int g(y)f(y^{-1}x)\phi(x)dydx \\ &= \int\int (L_{y^{-1}}f)(x)g(y)\phi(x)dydx \\ &= \int \varphi(L_{y^{-1}}f)g(y)dy\end{aligned}$$

which implies $\varphi(f)\phi(y) = \varphi(L_{y^{-1}}f)$ almost everywhere. Since the map $\varphi(L_{y^{-1}}f)/\varphi(f)$ is a uniformly continuous function of y, ϕ is continuous almost everywhere, and we might as well assume ϕ is continuous. We then conclude $\phi(xy) = \phi(x)\phi(y)$. Since $\|\phi\|_\infty = 1$ (this is the norm of any character operator on $L^1(G)$), we find ϕ maps into $\mathbf{T}$, for if $\|\phi(x)\| < 1$ for any particular x, $\|\phi(x^{-1})\| > 1$. $\square$

Thus there is a one-to-one correspondence with $\Gamma(G)$ and $\Gamma(L^1(G))$, which implies a connection with the Gelfand theory and the character theory of locally compact groups. This also gives us a locally compact topological structure on $\Gamma(G)$, induced by the Gelfand representation on $\Gamma(L^1(G))$. A sequence $\phi_i \to \phi$ if and only if

$$\int \frac{f(x)}{\phi_i(x)} dx \to \int \frac{f(x)}{\phi(x)} dx$$

for all functions $f \in L^1(G)$. This actually makes the map

$$(f, \phi) \mapsto \int \frac{f(x)}{\phi(x)} dx$$

a jointly continuous map, because as we verified in the proof above,

$$\widehat{f}(\phi)\phi(y) = \widehat{L_y f}(\phi)$$

And the map $y \mapsto L_y f$ is a continuous map into $L^1(G)$. If $K \subset G$ and $C \subset \Gamma(G)$ are compact, this allows us to find open sets in G and $\Gamma(G)$ of the form

$$\{\gamma : \|1 - \gamma(x)\| < \varepsilon \text{ for all } x \in K\} \qquad \{x : \|1 - \gamma(x)\| < \varepsilon \text{ for all } \gamma \in C\}$$

And these sets actually form a base for the topology on $\Gamma(G)$.

Theorem 8.2. *If G is discrete, $\Gamma(G)$ is compact, and if G is compact, $\Gamma(G)$ is discrete.*

Proof. If G is discrete, then $L^1(G)$ contains an identity, so $\Gamma(G) = \Gamma(L^1(G))$ is compact. Conversely, if G is compact, then it contains the constant 1 function, and

$$\widehat{1}(\phi) = \int \frac{dx}{\phi(x)}$$

And

$$\frac{1}{\phi(y)} \widehat{1}(\phi) = \int \frac{dx}{\phi(yx)} = \int \frac{dx}{\phi(x)} = \hat{1}(\phi)$$

So either $\phi(y) = 1$ for all y, and it is then verified by calculation that $\widehat{1}(\phi) = 1$, or $\widehat{1}(\phi) = 0$. Since $\widehat{1}$ is continuous, the trivial character must be an open set by itself, and hence $\Gamma(G)$ is discrete. ∎

Given a function $f \in L^1(G)$, we may take the Gelfand transform, obtaining a function on $C_0(\Gamma(L^1(G)))$. The identification then gives us a function on $C_0(\Gamma(G))$, if we give $\Gamma(G)$ the topology induced by the correspondence (which also makes $\Gamma(G)$ into a topological group). The formula is

$$\widehat{f}(\phi) = \phi(f) = \int \frac{f(x)}{\phi(x)}$$

This gives us the classical correspondence between $L^1(\mathbf{T})$ and $C_0(\mathbf{Z})$, and $L^1(\mathbf{R})$ and $C_0(\mathbf{R})$, which is just the Fourier transform. Thus we see the Gelfand representation as a natural generalization of the Fourier transform. We shall also denote the Fourier transform by $\mathcal{F}$, especially when we try and understand it's properties as an operator. Gelfand's theory (and some basic computation) tells us instantly that

- $\widehat{f * g} = \widehat{f}\widehat{g}$ (The transform is a homomorphism).
- $\mathcal{F}$ is norm decreasing and therefore continuous: $\|\widehat{f}\|_\infty \leqslant \|f\|_1$.
- If G is unimodular, and $\gamma \in \Gamma(G)$, then $(f * \gamma)(x) = \gamma(x)\widehat{f}(\gamma)$.

Whenever we integrate a function with respect to the Haar measure, there is a natural generalization of the concept to the space of all measures on G. Thus, for $\mu \in M(G)$, we define

$$\widehat{\mu}(\phi) = \int \frac{dx}{\phi(x)}$$

which we call the **Fourier-Stieltjes transform** on G. It is essentially an extension of the Gelfand representation on $L^1(G)$ to $M(G)$. Each $\widehat{\mu}$ is a bounded, uniformly continuous function on $\Gamma(G)$, because the transform is still contracting, i.e.

$$\left| \int \frac{d\mu(x)}{\phi(x)} dx \right| \leqslant \|\mu\|$$

It is uniformly continuous, because

$$(L_\nu \widehat{\mu} - \widehat{\mu})(\phi) = \int \frac{1 - \nu(x)}{\nu(x)\phi(x)} d\mu(x)$$

The regularity of μ implies that there is a compact set K such that $|\mu|(K^c) < \varepsilon$. If $\nu_i \to 0$, then eventually we must have $|\nu_i(x) - 1| < \varepsilon$ for all $x \in K$, and then

$$|(L_\nu \widehat{\mu} - \widehat{\mu})(\phi)| \leqslant 2|\mu|(K^c) + \varepsilon\|\mu\| \leqslant \varepsilon(2 + \|\mu\|)$$

Which implies uniform continuity.

Let us consider why it is natural to generalize operators on $L^1(G)$ to $M(G)$. The first reason is due to the intuition of physicists; most of classical Fourier analysis emerged from physical considerations, and it is in this field that $L^1(G)$ is often confused with $M(G)$. Take, for instance, the determination of the electric charge at a point in space. To determine this experimentally, we take the ratio of the charge over some region in space to the volume of the region, and then we limit the size of the region to zero. This is the historical way to obtain the density of a measure with respect to the Lebesgue measure, so that the function we obtain can be integrated to find the charge over a region. However, it is more natural to avoid taking limits, and to just think of charge as an element of $M(\mathbf{R}^3)$. If we consider a finite number of discrete charges, then we obtain a discrete measure, whose density with respect to the Lebesgue measure does not exist. This doesn't prevent physicists from trying, so they think of the density obtained as shooting off to infinity at points. Essentially, we obtain the Dirac Delta function as a 'generalized function'. This is fine for intuition, but things seem to get less intuitive when we consider the charge on a subsurface of $\mathbf{R}^3$, where the 'density' is 'dirac'-esque near the function, where as measure theoretically we just obtain a density with respect to the two-dimensional Hausdorff measure on the surface. Thus, when physicists discuss quantities as functions, they are really thinking of measures, and trying to take densities, where really they may not exist.

There is a more austere explanation, which results from the fact that, with respect to integration, $L^1(G)$ is essentially equivalent to $M(G)$. Notice that if $\mu_i \to \mu$ in the weak-$*$ topology, then $\widehat{\mu_i} \to \widehat{\mu}$ pointwise, because

$$\int \frac{d\mu_i(x)}{\phi(x)} \to \int \frac{d\mu(x)}{\phi(x)}$$

(This makes sense, because weak-$*$ convergence is essentially pointwise convergence in $M(G)$). Thus the Fourier-Stietjes transform is continuous with respect to these topologies. It is the unique continuous extension of the Fourier transform, because

Theorem 8.3. *$L^1(G)$ is weak-$*$ dense in $M(G)$.*

Proof. First, note that the Dirac delta function can be weak-$*$ approximated by elements of $L^1(G)$, since we have an approximate identity in the space.

First, note that if $\mu_i \to \mu$, then $\mu_i * \nu \to \mu * \nu$, because

$$\int f d(\mu_i * \nu) = \int \int f(xy) d\mu_i(x) d\nu(y)$$

The functions $y \mapsto \int f(xy) d\mu_i(x)$ converge pointwise to $\int f(xy) d\mu(y)$. Since

$$\left| \int f(xy) d\mu_i(x) \right| \leqslant \|f\|_1 \|\mu_i\|$$

If i is taken large enough that □

If $\phi_\alpha \to \phi$, in the sense that $\phi_\alpha(x) \to \phi(x)$ for all $x \in G$, then, because $\|\phi_\alpha(x)\| = 1$ for all x, we can apply the dominated convergence theorem on any compact subset K of G to conclude

$$\int_K \frac{d\mu(x)}{\phi_\alpha(x)} \to \int_K \frac{d\mu(x)}{\phi(x)}$$

It is immediately verified to be a map into $L^1(\Gamma(G))$, because

$$\int \left| \int \frac{d\mu(x)}{\phi(x)} \right| d\phi \leqslant \int \int \|\mu\|$$

The formula above immediately suggests a generalization to a transform on $M(G)$. For $\nu \in M(G)$, we define

$$\mathcal{F}(\nu)(\phi) = \int \frac{d\nu}{\phi}$$

If $\mathcal{G} : L^1(G) \to C_0(\Gamma(G))$ is the Gelfand transform, then the transform induces a map $\mathcal{G}^* : M(\Gamma(G)) \to L^\infty(G)$.

The duality in class-ical Fourier analysis is shown through the inversion formulas. That is, we have inversion functions

$$\mathcal{F}^{-1}(\{a_k\}) = \sum a_k e^{kit} \qquad \mathcal{F}^{-1}(f)(x) = \int f(t) e^{2\pi i x t}$$

which reverses the fourier transform on $\mathbf{T}$ and $\mathbf{R}$ respectively, on a certain subclass of L^1. One of the challenges of Harmonic analysis is trying to find where this holds for the general class of measurable functions.

The first problem is to determine surjectivity. We denote by $A(G)$ the space of all continuous functions which can be represented as the fourier transform of some function in $L^1(G)$. It is to even determine $A(\mathbf{T})$, the most basic example. $A(G)$ always separates the points of $\Gamma(G)$, by Gelfand theory, and if G is unimdoular, then it is closed under conjugation. If we let $g(x) = \overline{f(x^{-1})}$, we find

$$\mathcal{F}(g)(\phi) = \int \frac{g(x)}{\phi(x)} dx = \overline{\int \frac{f(x^{-1})}{\phi(x^{-1})} dx} = \int \frac{f(x)}{\phi(x)} dx = \overline{\mathcal{F}(f)(\phi)}$$

so that by the Stone Weirstrass theorem $A(G)$ is dense in $C_0(\Gamma(L^1(G)))$.

Chapter 9

Banach Algebra Techniques

In the mid 20th century, it was realized that much of the analytic information about a topological group can be captured in various C^* algebras related to the group. For instance, consider the Gelfand space of $L^1(\mathbf{Z})$ is $\mathbf{T}$, which represents the fact that one can represent functions over $\mathbf{T}$ as sequences of numbers. Similarily, we find the characters of $L^1(\mathbf{R})$ are the maps $f \mapsto \widehat{f}(x)$, so that the Gelfand space of $\mathbf{R}$ is $\mathbf{R}$, and the Gelfand transform is the Fourier transform on this space. For a general G, we may hope to find a generalized Fourier transform by understanding the Gelfand transform on $L^1(G)$. We can also generalize results by extending our understanding to the class $M(G)$ of regular, Borel measures on G.

Chapter 10

Vector Spaces

If $\mathbf{K}$ is a closed, multiplicative subgroup of the complex numbers, then $\mathbf{K}$ is also a locally compact abelian group, and we can therefore understand $\mathbf{K}$ by looking at its dual group $\mathbf{K}^*$. The map $\langle x, y \rangle = xy$ is bilinear, in the set that it is a homomorphism in the variable y for each fixed x, and a homomorphism in the variable x for each y.

If $\mathbf{K}$ is a subfield of the complex numbers, then $\mathbf{K}$ is also an abelian group under addition, and we can consider the dual group $\mathbf{K}^*$. The inner product $\langle x, y \rangle = xy$ gives a continuous bilinear map $\mathbf{K} \times \mathbf{K} \to \mathbf{C}$, and therefore we can define $x^* \in \mathbf{K}^*$ by $x^*(y) = \langle x, y \rangle$. If $x^*(y) = xy = 0$ for all y, then in particular $x^*(1) = x$, so $x = 0$. This means that the homomorphism $\mathbf{K} \to \mathbf{K}^*$ is injective.