

Harmonic Analysis

Jacob Denson

September 21, 2016

Table Of Contents

I	Classical Fourier Analysis	2
1	Springs, Strings, and Symmetry	4
1.1	The Heat Equation	8
1.2	Exponentials and Euler	11
2	Introductory Results	15
2.1	Methods of Summation	18
2.2	A Continuous Function with Divergent Fourier Series . . .	20
3	The Fourier Transform	22
4	Applications	26
4.1	The Wirtinger Inequality on an Interval	26
4.2	Energy Preservation in the String equation	27
II	A More Sophisticated Viewpoint	28
III	Abstract Harmonic Analysis	29
5	Finite Character Theory	31
5.1	Fourier analysis on Cyclic Groups	33
5.2	The Fast Fourier Transform	35
5.3	An Arbitrary Finite Abelian Group	36
5.4	Convolutions	39
5.5	Dirichlet's Theorem	40

6	Topological Groups	42
6.1	Basic Results	42
6.2	Ordered Groups	47
6.3	Topological Groups arising from Normal subgroups	49

Part I

Classical Fourier Analysis

Deep mathematical knowledge is often found by noticing a subtle symmetry in the phenomena studied. Nowhere is this more clear than in the foundations of harmonic analysis, where we attempt to understand ‘oscillating’ mathematical objects, in one form or another. The resulting ideas offer beautiful insights into many mathematical problems, and the mathematical ideas form the foundations of most of modern analysis. Here we approach the subject from the classical viewpoint, exploring the convergence of Fourier series, and elementary properties of the Fourier transform.

Chapter 1

Springs, Strings, and Symmetry

It is almost compulsory to begin our study by looking at the most basic example of oscillating behaviour. We shall see it's dynamics reflected in all that we shall study. In fact, almost every example of real world oscillation follows this behaviour – be it the pendulum of a grandfather clock or a compressed spring. It is in these systems that force on objects resists motion away from an equilibrium position. Regardless of the equation describing the force, the motion can be linearly approximated by the differential equation

$$\ddot{x} = -k^2x$$

It is well known that solutions of this equation take the form

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

where the two representations are connected by the trigonometric equality

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

No serious effort was required on our part to produce these equations – they were known to Newton, and to Hooke before him, and require only the basic methods of calculus. But this equation is very important; it forebodes that trigonometric functions will occur over and over again in the study of oscillatory behaviour.

Fourier analysis took form from a more substantial problem. Consider a tethered string vibrating under the influence of tension. 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 – we must describe the motion over a line of infinitely many points. However, as we know from physics, a string is really just a collection of atoms, strung together by certain physical forces. Therefore we should be able to understand the motion of the string by looking at only finitely many particles at a time. This shall turn out to be the key in a great many problems we shall find in the field of Fourier analysis. Infinite dimensional problems are best understood via approximation from the finite dimensional.

Since the horizontal tension is uniform in a string of uniform thickness and mass, there is no horizontal motion in the string. Thus we can describe the position of a string at time t by a real-valued function $u(t, x)$, defined on $\mathbf{R} \times [0, \pi]$. Since the string is tethered down at both ends, $u(t, 0) = u(t, \pi) = 0$. To obtain an equation describing u , we approximate our system by imagining a system of n evenly distributed particles on the string. We may see the dynamics of the system, as if adjacent particles are strung together by springs, since the tension between adjacent particles can be linearly approximated. Physical experience tells us that as we make the springs smaller, the tension in the spring increases. Thus we assume there is a number K such that the spring constant k of the string segments is equal to nK . If M is the total mass of the spring, then the smaller masses have mass $m = M/n$. We can then combine the spring forces to the left and right of each force, and conclude (by Newton's law) that

$$\frac{M}{n} \frac{\partial^2 u}{\partial t^2} \approx -nK \left[u \left(t, x + \frac{\pi}{n} \right) - u(t, x) \right] - nK \left[u \left(t, x - \frac{\pi}{n} \right) - u(t, x) \right]$$

Taking this approximation to its limits, and assuming $u \in C^2[0, \pi]$, we find

$$\frac{\partial^2 u}{\partial t^2} = \lim_{n \rightarrow \infty} -n^2 \frac{K}{M} \left[u \left(t, x + \frac{\pi}{n} \right) + u \left(t, x - \frac{\pi}{n} \right) - 2u(t, x) \right] = \frac{K}{M} \frac{\partial^2 u}{\partial x^2}$$

The one dimensional wave equation is the partial differential equation

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

a normalized version of the equation describing the string.

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

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

where v is periodic. As a start to solving the wave equation, let us find all standing wave solutions to the wave equation. Such an equation must satisfy

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

or

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

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 equations¹

$$\psi''(x) = \lambda\psi(x) \quad v''(t) = \lambda v(t)$$

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

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

¹One small problem with this argument is that ψ and v could be zero at certain points, and it is not clear that the region in which ψ and v is non-zero is connected, so that this argument works. Technically, this doesn't cause problems, since the method 'works' (it gives us solutions to the wave equation), but for complete understanding, we should be a bit more careful.

First assume $u \neq 0$. Then there is (t, x) such that $u(t, x) = \psi(x)v(t) \neq 0$. Here it is possible to consider the relation

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

By first varying t , and then varying x , we conclude that v''/v and ψ''/ψ are equal to some constant λ whenever v and ψ do not equal zero. Thus $v''(t) = \lambda v(t)$, except possibly when $v(t) = 0$. We shall show that this relation does hold everywhere, by breaking our argument into two cases. First, suppose there is a sequence t_1, t_2, \dots converging to t such that $v(t_k) \neq 0$. Then, assuming $v \in C^2(\mathbf{R})$, $v''(t_k) \rightarrow v''(t)$, but also $v''(t_k) = \lambda v(t_k) \rightarrow \lambda v(t)$, so $v''(t) = \lambda v(t)$. The second case occurs when t is contained in an interval (a, b) on which $v = 0$. In this case, $v'' = 0$ (since v is constant here), so certainly $v''(t) = \lambda v(t)$. The same argument shows that the relation holds for ψ everywhere as well.

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, we obtain a 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 $n \rightarrow \infty$. Now given the initial conditions $u(0, x) = f(x)$, we find

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

and given the initial velocity $\frac{\partial u}{\partial t}(0, x) = g(x)$, if there's any justice in the world, we should have

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

Finding the expansion of u in terms of harmonic frequencies reduces to decomposing 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 convergence?

The first question can also be approached by formal calculation. Suppose that a function f has an expansion

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

Using the fact that

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

We find that

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

And so given any function $f : [0, \pi] \rightarrow \mathbf{R}$, a reasonable candidate for each A_n is

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

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

1.1 The Heat Equation

Now we come to a quite different physical situation. Suppose we have a two-dimensional region D , upon which temperature fluctuates. What differential equation describes the change of temperature over time? Let $u(t, x, y)$ describe the temperature at a certain time and position. Consider a small square S centered at x , with sides parallel to the axis and with side lengths $2h$. Then the total heat in this square at time t is

$$H(t) = \int_S u(t, x, y) dx dy \approx 4h^2 u(t, x, y)$$

After differentiating, we find

$$H'(t) \approx 4h^2 \frac{\partial u}{\partial t}(t, x, y)$$

Newton's law of cooling tells us that the heat flow at a point is proportional to the negative of the gradient. To find the total amount of heat leaving

the square, we perform a line integral around the sides of the square. Thus

$$\begin{aligned} H'(t) &= \int_{\partial S} \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) \cdot n \, dS \\ &\approx 2h \left[\frac{\partial u}{\partial x}(t, x+h) - \frac{\partial u}{\partial x}(t, x-h) + \frac{\partial u}{\partial y}(t, x+h) - \frac{\partial u}{\partial y}(t, x-h) \right] \end{aligned}$$

Now we approximate to the extreme to obtain equality.

$$\frac{\partial u}{\partial t} = \lim_{h \rightarrow 0} \frac{1}{2h} \left[\frac{\partial u}{\partial x}(t, x+h) - \frac{\partial u}{\partial x}(t, x-h) + \frac{\partial u}{\partial y}(t, x+h) - \frac{\partial u}{\partial y}(t, x-h) \right]$$

Using the same trick as in the string equation, we find that

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

where Δ is the Laplacian operator. This partial differential equation is the heat equation.

To simplify again, we start by looking at only the steady state heat equations, those functions u satisfying $\Delta u = 0$. 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 \mathbf{D} , with 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

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

Or, separating variables,

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

In which case we find the value is fixed, and equal to λ^2 for some λ (If the constant value was negative, g wouldn't be periodic). Solving these equations tells us

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

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 (and because we want to solve the equation on all of \mathbf{D}), 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 $rf''(r) + f'(r) = 0$. here all solutions 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

$$u(r, \theta) = \sum_{m=0}^{\infty} [A_m \cos(m\theta) + B_m \sin(m\theta)]r^m$$

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. Applying a formal integration trick, given a particular function $f(\theta)$ on $[-\pi, \pi]$, the coefficients of the expansion should be

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

If we begin with a function on $[0, \pi]$, and enlarge the domain to $[-\pi, \pi]$ by making the function odd, then the cosine terms cancel out, and we end up with the same expansion as on $[0, \pi]$. Since an arbitrary function f 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 $[0, \pi]$.

1.2 Exponentials and Euler

We are working with 2π -periodic functions $f : \mathbf{R} \rightarrow \mathbf{R}$, and attempting to decompose them into summations of sines and cosines, but we have a much more elegant representation. 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 $g : \mathbf{T} \rightarrow \mathbf{R}$, define

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

Then f is 2π periodic. Conversely, a 2π periodic function f gives rise to a function g , defined by the same formula. 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

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

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

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

so a Fourier expansion is really just a power series expansion in disguise.

Now if f is real-valued, then $C_k = \overline{C_k}$, and the complex parts of the expansion cancel out. But really, in this form, we see that there is no real harm in generalizing our reach to complex-valued functions. If we are given $f(t)$, 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

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

defined on \mathbf{Z} , called the Fourier series of f . For notational simplicity, we shall define the integral

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

which essentially just calculates the average of f over the circle. The notation for the Fourier transform then becomes

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

the most austere and elegant way to write the transform. In the sequel, our core goal is to analyze the relation of \hat{f} to f , most notably the convergence of

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) z^n$$

to the function f .

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 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 can be calculated as

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

which we may rewrite as

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

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

Example. The tent function

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

has a Fourier expansion

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

for $n \neq 0$, and $\hat{f}(0) = \frac{\delta}{2\pi}$, so

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

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

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

Hence we may write

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

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

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

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 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

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

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 Fourier series need not be absolute.

Chapter 2

Introductory Results

Let's focus in on the problem we introduced in the last chapter. For each function $f : \mathbb{T} \rightarrow \mathbb{C}$, we have an associated series

$$\sum_{n=-\infty}^{\infty} a_n e^{inx}$$

known as a **trigonometric series**. We can also consider finite sums

$$\sum_{n=-N}^N a_n e^{inx}$$

which we call a **trigonometric polynomial**. The largest value of n such that $|a_n| + |a_{-n}| \neq 0$ is known as the **degree** of the polynomial. At this point, we haven't deduced a reason for the equation

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

to hold in any reasonable way. First, we define the m 'th partial sum

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

The first relation we can expect is pointwise convergence; is it true that for every x ,

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

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

To begin with, we shall begin with a brief look at the properties of the association of f and \hat{f} . For a start, since the operation of integration is linear, we know that

$$\widehat{(af + bg)}(n) = a\hat{f}(n) + b\hat{g}(n)$$

If we rotate the argument of a function, writing $g(z) = f(zw)$, for $w \in S^1$, then

$$\hat{g}(n) = \int_{S^1} f(zw)z^{-n}dz = \int_{S^1} f(z)(zw^{-1})^{-n} = w^n \hat{f}(n)$$

If we take complex conjugates, we find

$$\overline{\hat{f}}(n) = \int_{S^1} \overline{f(z)}z^{-n}dz = \overline{\int_{S^1} f(z)z^n dz} = \overline{\hat{f}(-n)}$$

The relations among the Fourier coefficients essentially holds because of the rotational, scaling, and inversion symmetry of the circle. Note that this shows us that for real valued functions, the n 'th Fourier coefficient is the conjugate of the $-n$ 'th coefficient.

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 integrals, remains the same. Nonetheless, if a function is continuous editing the function at a point will break continuity, so we may have some hope.

Theorem 2.1. *If the Fourier coefficients of a function vanishes, then the function vanishes at every point of continuity.*

Proof. We shall begin by proving this for real valued functions. Consider a function f whose Fourier coefficients vanish. Then for every trigonometric polynomial $P(x) = \sum_{n=-N}^N a_n e^{-nix}$, we have

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

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 chosen small enough that $P(x) > A > 1$ for $|x| < \delta/2$, and $P(x) < B < 1$ for $|x| \geq \delta/2$. Consider the series of trigonometric polynomials

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

For which we have

$$\begin{aligned} \left| \int_{-\pi}^{\pi} P_n(x) f(x) dx \right| &\geq \left| \int_{|x| < \delta/2} P_n(x) f(x) dx \right| - \left| \int_{|x| \geq \delta/2} P_n(x) f(x) dx \right| \\ &> \delta A^n \frac{f(0)}{2} - B^n \|f\|_{\infty} \end{aligned}$$

The left side is always equal to zero, regardless of n , whereas the right side tends to infinity as we take n to extreme values.

In general, for an arbitrary point of continuity x , we replace f with the function $g(y) = f(x + y)$. Then

$$\hat{g}(n) = e^{nix} \hat{f}(n) = 0$$

So the Fourier coefficients of g vanish at 0, hence $g(0) = f(x) = 0$. For an arbitrary function $f(x) = u(x) + iv(x)$, we have

$$\begin{aligned} \hat{u}(n) &= \frac{\hat{f}(n) + \widehat{\bar{f}}(n)}{2} = \frac{\hat{f}(n) + \overline{\hat{f}(-n)}}{2} = 0 \\ \hat{v}(n) &= \frac{\hat{f}(n) - \widehat{\bar{f}}(n)}{2i} = \frac{\hat{f}(n) - \overline{\hat{f}(-n)}}{2i} = 0 \end{aligned}$$

so $u = v = 0$. □

Corollary 2.2. *If two continuous functions f and g have the same Fourier coefficients, then $f = g$.*

Proof. Because then $f - g$ is a continuous function whose Fourier coefficients vanish, so $f - g = 0$. □

We also have a first positive result for convergence.

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

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

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

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

2.1 Methods of Summation

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

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

Similarly,

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

However, we shall see that this method of summation is not sufficient for understanding the convergence of fourier series, which we have seen manifest in the dirichlet kernel not vanishing away from the origin.

Thus we must introduce more subtle methods of convergence. The first is due to Cesaro. Rather than considering limits of partial sums, we consider limits of averages of sums. Letting $s_n = \sum_{k=0}^n a_k$, we define the Cesaro summation as

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

If the normal summation exists, then the Cesaro limit exists, and is equal to the original sum. However, the Cesaro summation is stronger, for if we consider the sequence

$$1, -1, 1, -1, 1, \dots$$

Then the partial sums do not converge, but the Cesaro sum converges to zero.

Finally, we consider Abel summation. Given a sequence $\{a_i\}$, we can consider the power series $\sum a_k r^k$. If this converges in $(-1, 1)$, we can ask if $\lim_{r \rightarrow 1} \sum a_k r^k$, which should be ‘almost like’ $\sum a_k$. If this limit exists, we call in the Abel summation.

Theorem 2.4. *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 $n \rightarrow \infty$, 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

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

Fixing N , and letting $r \rightarrow 1$, 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

To relate Abel summation to fourier analysis, we express the limit of the abel sum as a limit of convolutions. If

$$f(z) \sim \sum a_i z^i$$

then formally, we write

$$f(z) \sim \lim_{r \rightarrow 1} \sum_{i=-\infty}^{\infty} a_i r^{|i|} z^i = \left(\sum a_i z^i \right) * \left(\sum r^{|i|} z^i \right)$$

Now for $r < 1$, and $z \in S^1$,

$$\sum_{i=-\infty}^{\infty} r^{|i|} z^i = 1 + \frac{rz}{1-rz} + \frac{rz^{-1}}{1-rz^{-1}} = \frac{1-r^2}{1+r^2-2r\Re[z]}$$

We define the Poisson kernel P_r by this formula, and by convergence properties we have already established, the abel limit of the fourier series is

$$\lim_{r \rightarrow 1} (f * P_r)$$

Thankfully, we find P_r is a good kernel, and the Abel means work nicely with Fourier series.

2.2 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

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

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**

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

which results in the relation

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

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} .

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

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 $b > a$,

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

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 \hat{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 (continuous 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 ,*

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

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

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

taking limits of both sides, we find

$$\lim_{x \rightarrow \infty} \int_{g(x)}^{f(x)} h \leq \int_{-\infty}^{\infty} h$$

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

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

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

$$\lim_{x \rightarrow \infty} \int_{g(x)}^{f(x)} h \geq \int_{-\infty}^{\infty} h$$

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*

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

Chapter 4

Applications

4.1 The Wirtinger Inequality on an Interval

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

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

Proof. Consider the fourier series

$$f(t) \sim \sum a_n e^{nit} \quad f'(t) \sim \sum i n a_n e^{nit}$$

Then $a_0 = 0$, and so

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

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

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

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

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

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

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, the functions take the form of the power maps

$$z \mapsto z^k$$

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

- The functions are orthogonal to one another.
- The family can approximate functions on the space.
- $e_n(zw) = e_n(z)e_n(w)$.

The last property of these exponentials should be immediately recognizable to any student of group theory. It implies the exponentials are homomorphisms from \mathbf{T} into itself. This is the easiest of the three properties to generalize. We shall call a homomorphism from any abelian group to \mathbf{T} a **character**. For any abelian group G , we can put all characters together¹ to form the character group \hat{G} , which forms an abelian group under the rule

$$(f \cdot g)(z) = f(z) \cdot g(z)$$

It is these functions which are ‘primitive’ in the synthesis of Fourier series on a general space.

¹If G is topological, then \hat{G} shall consist of only the continuous characters.

Example. If μ_n is the set of n 'th roots of unity

$$\mu_n = \{z \in \mathbf{C} : z^n = 1\}$$

Then $\hat{\mu}_n$ consists of the power maps $z \mapsto z^k$, which are distinct for $0 \leq k \leq n-1$. If we define $\omega = e^{2\pi i/n}$, then we find that

$$\mu_n = \{1, \omega, \dots, \omega^{n-1}\}$$

Let $f : \mu_n \rightarrow \mathbf{T}$ be an arbitrary character. Then for any $z \in \mu_n$,

$$f(z)^n = f(z^n) = f(1) = 1$$

so f maps μ_n into itself. Since

$$f(\omega^k) = f(\omega)^k$$

we see that the map f is uniquely determined by where it maps ω . Since the characters $e_k : z \mapsto z^k$, map ω to any element of μ_n , we see that these are all the characters. Since $e_i \cdot e_j = e_{i+j}$, we find that the character space of μ_n is isomorphic to μ_n .

The group \mathbf{Z}_n is isomorphic to μ_n under the identification $n \mapsto \omega^n$. Thus, as with \mathbf{T} , we do not distinguish functions 'defined in terms of n ' and 'defined in terms of ω ', assuming the correspondance $n = \omega^n$. The characters of \mathbf{Z}_n are then exactly the maps $n \mapsto \omega^{kn}$. This follows from the general fact that if $f : G \rightarrow H$ is an isomorphism of abelian groups, the map $F : \phi \mapsto \phi \circ f$ is an isomorphism from \hat{H} to \hat{G} .

Example. Finding the characters of \mathbf{T} shall employ much of the Fourier analysis we've built up so far. First off, we shall only be interested in the continuous characters, and it shall only be these characters which compose the character space. Let $f : \mathbf{T} \rightarrow \mathbf{T}$ 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 either $f(w) = w^n$ for all w , for some n , or $\hat{f}(n) = 0$ for all n , implying $f = 0$ since f is continuous.

Example. The characters of \mathbf{R} are of the form $t \mapsto e^{ti\xi}$, for $\xi \in \mathbf{R}$. To see this, let $e : \mathbf{R} \rightarrow \mathbf{T}$ 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 $\xi \in \mathbf{R}$.

5.1 Fourier analysis on Cyclic Groups

We shall start our study of abstract Fourier analysis by looking at a particular case – 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.

There is a trick which we can use to obtain quick results in μ_n , but it does not generalize to arbitrary groups. Given a function $f : \mu_n \rightarrow \mathbf{C}$, consider the periodic function

$$g = \sum_{k=1}^n f(k) \chi\left(\frac{\pi(2k-1)}{n}, \frac{\pi(2k+1)}{n}\right)$$

We then have the expansion

$$g(t) \sim \sum a_m e^{mit}$$

where

$$\begin{aligned} a_m &= \frac{1}{2\pi} \int_0^\pi g(t) e^{-mit} dt = \frac{1}{2\pi} \sum_{k=1}^n f(k) \int_{\frac{\pi(2k-1)}{n}}^{\frac{\pi(2k+1)}{n}} e^{-mit} dt \\ &= \frac{i}{2\pi m} \sum_{k=1}^n f(k) e^{-\frac{2\pi imk}{n}} \left[e^{-\frac{\pi im}{n}} - e^{\frac{\pi im}{n}} \right] \\ &= \frac{\sin\left(\frac{\pi m}{n}\right)}{\pi m} \sum_{k=1}^n f(k) e^{-\frac{2\pi imk}{n}} \end{aligned}$$

and since g is differentiable at k , we have

$$\begin{aligned} f(a) &= g\left(\frac{2\pi a}{n}\right) = \frac{1}{n} \sum_{k=1}^n f(k) + \sum_{m \neq 0} \sum_{k=1}^n \frac{\sin\left(\frac{\pi m}{n}\right)}{\pi m} f(k) e^{\frac{2\pi im(a-k)}{n}} \\ &= \frac{1}{n} \sum_{k=1}^n f(k) + \sum_{m=1}^{n-1} \frac{\sin\left(\frac{\pi m}{n}\right)}{\pi} \sum_{k=1}^n f(k) e^{\frac{2\pi im(a-k)}{n}} \sum_{l \in \mathbb{Z}} \frac{(-1)^l}{(m + ln)} \end{aligned}$$

Now it takes some complex analysis to show

$$\sum_{l \in \mathbb{Z}} \frac{(-1)^l}{m + ln} = \frac{\pi}{n \sin\left(\frac{\pi m}{n}\right)}$$

This shows that we have a finite expansion

$$f(k) = \sum_{m=1}^n a_m e^{2\pi imk/n}$$

where

$$a_m = \frac{1}{n} \sum_{k=1}^n f(k) e^{-2\pi imk/n}$$

but this method is really just a cheat, for it does not generalize to arbitrary abelian groups. A more enlightening approach takes advantage of the

character theory we've developed. Perhaps more interesting is the fact that we can recover much of the theory of Fourier series on \mathbf{T} from finite fourier analysis, which I'll do the calculations for later.

Let V denote the set of all functions from μ_n to \mathbf{C} . 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 $e_k : z \mapsto z^k$ are orthonormal in this space, for

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

If $i - j = 0$, $\langle e_i, e_j \rangle = n$. 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 independent. Because V is only n dimensional by itself, the family of characters must span the space. Thus, for any function f , we have

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

which is the fourier expansion we obtained above.

5.2 The Fast Fourier Transform

The main use of the fourier series on μ_n is to approximate fourier series methods on \mathbf{T} in physical problems needing numerical values. Computing the Fourier coefficients of a function $f : \mu_n \rightarrow \mathbf{C}$ is possible in $O(n^2)$ 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 symmetry in the group μ_{2n} . Given $f : \mu_{2n} \rightarrow \mathbb{C}$, define two functions $g, h : \mu_n \rightarrow \mathbb{C}$, defined by

$$g(k) = f(2k) \quad h(k) = f(2k+1)$$

And this pair of functions encodes all the information in f . Indeed, if $v = e^{\pi i/n}$ is the generator of μ_{2n} , we have

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

Because

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

If $H(m)$ is the number of operations needed to calculate the fourier transform of a function on μ_{2^n} , then the above relation tells us

$$H(2m) = 2H(m) + 3 \cdot (2m)$$

If $G(n) = H(2^n)$, then

$$G(n) = 2H(n-1) + 3 \cdot 2^n$$

and $G(0) = 1$. Now 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 $H(m) = m(1 + 3\log_2(m)) = O(m \log m)$.

5.3 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 \mathbb{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 independent.

Theorem 5.1. *The set \hat{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. □

We see from this that $|\hat{G}| \leq |G|$. 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

$$\widehat{G \times H} \cong \hat{G} \times \hat{H}$$

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.

Theorem 5.2. Let $\{T_1, \dots, T_n\}$ be a family of commuting unitary matrices in $U_m(\mathbf{C})$. 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 trivial. 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 hypothesis 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 theorem 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 – there is a family f_1, \dots, f_n such that for each $a \in G$, $T_a f_i = \lambda_{i,a} f_i$. We may assume $f_i(1) = 1$ for each i . Then, for any $a \in G$, we have

$$f_i(a) = f_i(a \cdot 1) = \lambda_{i,a} f_i(1) = \lambda_{i,a}$$

So for any $b \in G$,

$$f_i(ab) = \lambda_{i,a} f_i(b) = f_i(a) f_i(b)$$

and each f_i is a character, completing the proof.

We summarize our arguments with the following theorem.

Theorem 5.3. Let G be a finite abelian group. Then \hat{G} has the same cardinality as G , and for any function $f : G \rightarrow \mathbf{C}$,

$$f(a) = \sum_{e \in \hat{G}} \langle f, e \rangle e(a)$$

where

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

In this context, we also have Parseval's theorem

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

$$\langle f, g \rangle = \sum_{e \in \hat{G}} \langle f, e \rangle \overline{\langle g, e \rangle}$$

5.4 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

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

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

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

which shows $f * g = g * f$. 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

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

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

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

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

5.5 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*

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

contains infinitely many prime numbers.

An exploration of this requires 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

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \leq 1 + \int_1^{\infty} \frac{1}{t^s} = 1 + \lim_{n \rightarrow \infty} \frac{1}{s-1} [1 - 1/n^{s-1}] = 1 + \frac{1}{s-1}$$

The series converges uniformly on $[1 + \varepsilon, N]$ for any $\varepsilon > 0$, so ζ is continuous on $(1, \infty)$. As $t \rightarrow 1$, $\zeta(t) \rightarrow \infty$, because $n^s \rightarrow n$ 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 $1 \leq n \leq M$, then

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

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

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, enumerating the primes to be p_1, p_2, \dots , we find

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

□

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 countable product \mathbf{T}^ω of Torii, which occurs in the theory of Dirichlet series in number theory.*

A few more pathological examples are also important for understanding the intuition of results. We detail a basic example now. More will occur in the sequel.

6.1 Basic Results

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

ical group, the maps

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

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

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

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 $u_i v_i \rightarrow x$. Since U is compact, there is a subnet u_{i_k} converging to y . Then $y \in U$, and we find

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

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 hence 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. \square

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 . \square

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.*

An advantage of the real line \mathbf{R} is because continuity can be explained in a *uniform sense*, because we can transport any topological questions about a certain point x to questions about topologies near the origin, via the map $g \mapsto x^{-1}g$. Thus it makes sense to say that a function $f : \mathbf{R} \rightarrow \mathbf{R}$ is continuous uniformly, by saying that for any $\varepsilon > 0$ there is δ such that if $|y| < \delta$, $|f(x + y) - f(x)| < \varepsilon$. Instead of having to specify a δ for every point on the domain, it works uniformly everywhere. Similarly, we say a function $f : G \rightarrow H$ 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.

Example. *Let G be any Hausdorff non-commutative topological group, with sequences x_i and y_i for which $x_i y_i \rightarrow e$, $y_i x_i \rightarrow z \neq e$. Let U be a neighbourhood of z , and V a neighbourhood of e , with U and V disjoint. Let*

$$x_j (y_j^{-1})^{-1} = x_j y_j \in V \quad (y_j^{-1})^{-1} x_j = y_j x_j \in U$$

for $j > n$. Thus

Then the left and right uniform structures are inequivalent. This shows that on most matrix groups, the left and right uniformity is 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, a theorem of logic might give us respite. We can turn a statement of the form $(\forall x, \exists y : S(x, y))$ to a statement of the form $(\exists y, \forall x : S'(x, y))$ by letting y be a function, so that $S'(x, y)$ is interpreted as $S(x, y(x))$. In the case of uniform continuity, we have the logical statement

$$(\forall U, \exists V, \forall x : f(x)^{-1} f(xV) \subset U)$$

which can be ‘skolemized’ to

$$(\forall U, \exists V : f(c)^{-1} f(cV(c)) \subset U)$$

we introduce a function V of points $g \in G$ and neighbourhoods $U \in H$, such that $f(V(g, U)) \subset U$.

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

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

The family of all L_U (resp. R_U) is known as the left (right) uniform structure on G , denoted $\mathcal{S}_l(G)$ and $\mathcal{S}_r(G)$. Fix a map $f : G \rightarrow H$, and consider the map

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

from G^2 to H^2 . Then f is uniformly continuous if and only if g is continuous with respect to the uniform structures on G and H . Since $\mathcal{S}_l(G)$ and $\mathcal{S}_r(G)$ are weaker than the product topologies on G and H . We can also consider uniform maps with respect to $\mathcal{S}_l(G)$ and $\mathcal{S}_r(H)$, and so on and so forth. We can also consider uniform continuity on functions defined on subsets 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 $\mathcal{S}_l(G)$ and $\mathcal{S}_r(G)$ define the same topologies, and so uniform continuity is invariant of the uniform structure chosen.
- Translation maps $x \mapsto axb$, for $a, b \in G$, are uniform.
- Inversion is uniformly continuous.

Theorem 6.5. 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 $f : K \rightarrow H$ 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 $U'_x{}^2 \subset U_x$. 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. \square

Corollary 6.6. *Uniform continuity on compact groups is invariant of the uniform structure chosen.*

6.2 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

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

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 $b^2 \geq a$, then $b \geq ab^{-1}$ and so

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

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 $f_i \rightarrow f$ if eventually $f_i(s) = f(s)$ for all $s < w$ simultaneously, and $f_i(w) \rightarrow f(w)$. 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 $f_i \rightarrow f$. 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 $f_i \rightarrow f$ 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 \aleph , then every set of cardinality $\leq \aleph$ 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

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

Suppose that $f_i \leq k < h < k' \leq g_j$. 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 possibility is that $f_i \rightarrow h$ or $g_i \rightarrow h$, 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.

6.3 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 $g_i \rightarrow g$ 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 $g_i h_i \in gN hN = ghN$, 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 $g_i \rightarrow g$, 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 Furstenberg, 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.