MA342A 2017–18

Chapter 1: Examples for Fourier analysis

Contents

1.1 Review: Fourier series

We will take the point of view that Fourier analysis makes sense for *Locally compact abelian groups* but first we will consider a few examples which will be developed more fully later.

The basic idea of Fourier was that all functions could be made up from a superposition of sines and cosines. The easiest situation to describe is where we deal with periodic functions and the case of a general period L can be reduced to the that have period 1 by a change of variable that re-scales the independent variable. However, it is also common to use $L = 2\pi$ and we review that first.

1.1.1 Fourier sine and cosine series

The idea of Fourier was that all functions $f(\theta)$ could be built up from cosines and sine. We will look at the case or periodic functions where we come up with Fourier series (but there is also a theory of Fourier transforms applicable to many [non periodic] functions $f\colon \mathbb{R} \to \mathbb{R}$).

So the idea is that we can build up $f(\theta)$ by combining $\cos(n\theta)$ and $\sin(n\theta)$ for different $n \in \mathbb{Z}$. Because $\cos$ is even $\cos(-n\theta) = \cos(n\theta)$ and so there is nothing to be gained by allowing $n < 0$. For $n = 0$ we get $\cos 0 = 1$, the constant function and we do keep that. As $\sin$ is odd $\sin(-n\theta) = -\sin(n\theta)$ and again we stick to n positive. Now for $n = 0$ we get $\sin 0 \equiv 0$ and we don't keep that.

So the idea of Fourier was to consider sums

$$A_0 + \sum_{n=1}^{\infty}(A_n \cos(n\theta) + B_n \sin(n\theta)) \tag{1.1.1}$$

We will gloss over lots of details for the moment and arrive at a plausible strategy for finding suitable A_n's and B_n's starting from f. Later we will return to investigate whether (on in what circumstances) the sum (1.1.1) converges in any sense and whether it give the right answer $f(\theta)$.

As all the terms we are looking at in (1.1.1) are 2π-periodic, it seems we should restrict ourselves to functions f that satisfy $f(\theta + 2\pi) \equiv f(\theta)$. We can then consider and inner product for functions

$$\langle f, g\rangle = \int_0^{2\pi} f(\theta) g(\theta)\, d\theta$$

[Warning: we will change to dealing with $\mathbb{C}$-valued functions very soon and then the right thing will be $\langle f, g\rangle = \int_0^{2\pi} f(\theta)\overline{g(\theta)}\, d\theta$.] It turns out that most of the inner products between the building block functions $\cos(n\theta)$ and $\sin(n\theta)$ are zero. For instance, if $n \neq m$

$$\begin{aligned}
\langle \cos(n\theta), \cos(m\theta)\rangle &= \int_0^{2\pi} \cos(n\theta)\cos(m\theta)\, d\theta \\
&= \int_0^{2\pi} \frac{1}{2}(\cos((n+m)\theta) + \cos((n-m)\theta))\, d\theta \\
&= \left[\frac{1}{2}\left(\frac{1}{n+m}\sin((n+m)\theta) + \frac{1}{n-m}\sin((n-m)\theta)\right)\right]_0^{2\pi} \\
&= 0
\end{aligned}$$

On the other hand, for $n = m > 0$

$$\begin{aligned}
\langle \cos(n\theta), \cos(n\theta)\rangle &= \int_0^{2\pi} \cos^2(n\theta)\, d\theta \\
&= \int_0^{2\pi} \frac{1}{2}(\cos(2n\theta) + 1)\, d\theta \\
&= \left[\frac{1}{2}\left(\frac{1}{2n}\sin(2n\theta) + \theta\right)\right]_0^{2\pi} = \pi
\end{aligned}$$

nut for $n = 0$ we get $\langle 1, 1\rangle = 2\pi$. For the $\sin$ inner products we get $\langle \sin(n\theta), \sin(m\theta)\rangle = 0$ if $n \geq m$ but π if $n = m > 0$ and the mixed inner products $\langle \sin(n\theta), \cos(m\theta)\rangle = 0$ for $n > 0$, $m \geq 0$.

If we use these facts and take a cavalier attitude to any technical difficulties that might arise, if $f(\theta)$ is equal to the series (1.1.1), then $\langle f(\theta), \cos(m\theta)\rangle$ should be

$$A_0\langle 1, \cos(m\theta)\rangle + \sum_{n=1}^{\infty}(A_n\langle\cos(n\theta), \cos(m\theta)\rangle + B_n\langle\sin(n\theta)), \cos(m\theta)\rangle$$

So $\langle f(\theta), \cos(m\theta)\rangle = A_m\pi$ if $m > 0$ and $= a_0(2\pi)$ if $m = 0$. Similarly $\langle f(\theta), \sin\cos(m\theta)\rangle = B_m\pi$ for $m > 0$.

What we can do then is define the Fourier $\cos$ - $\sin$ series for f to be (1.1.1) with

$$A_n = \begin{cases} \frac{1}{\pi}\int_0^{2\pi} f(\theta)\cos(n\theta)\, d\theta & \text{for } n > 0 \\ \frac{1}{2\pi}\int_0^{2\pi} f(\theta) d\theta & \text{for } n = 0 \end{cases}$$

and

$$B_n = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\sin(n\theta)\,d\theta \qquad (n > 0).$$

Then our next **question** could be whether the series (1.1.1) now adds up to f is some reasonable sense?

Before that, we should maybe take care of what we mean by those integrals that give the coefficients A_n and B_n. One approach is to take $f\colon [0, 2\pi] \to \mathbb{R}$ to be continuous and we would probably insist also that $f(0) = f(2\pi)$. With $f(0) = f(2\pi)$ we can repeat the values $f(\theta)$ for $0 \leq \theta \leq 2\pi$ on each interval $2(n-1)\theta \leq \theta \leq 2n\theta$ $(n \in \mathbb{Z})$ and then we have $f\colon \mathbb{R} \to \mathbb{R}$ continuous and 2π-periodic. For these continuous f we can use the Riemann integral.

More generally we could consider Lebesgue integrable f, meaning ones which are Lebesgue measurable on $[0, 2\pi]$ and have $\int_0^{2\pi} |f(\theta)|\,d\theta < \infty$. [Recall $f\colon [0, 2\pi] \to \mathbb{R}$ is Lebesgue measurable if $\{\theta \in [0, 2\pi] : f(\theta) \leq a\}$ is always a Lebesgue measurable set, for each $a \in \mathbb{R}$. Also it is quite hard to come across non-measurable functions so that this measurability assumption is not usually a difficulty. If f is measurable, then $|f(\theta)|$ is measurable and never negative. So it is possible to define $\int_0^{2\pi} |f(\theta)|\,d\theta$ always if we allow ∞ as a value. To get $\int_0^{2\pi} f(\theta)\,d\theta$ we take $\int_0^{2\pi} f^+(\theta)\,d\theta - \int_0^{2\pi} f^-(\theta)\,d\theta$ and we need at least one of $\int_0^{2\pi} f^\pm(\theta)\,d\theta$ to be finite for this to make any sense. We say f is integrable if both are finite and then $\int_0^{2\pi} f(\theta)\,d\theta \in \mathbb{R}$.]

If f is integrable then so are $f(\theta)\cos(n\theta)$ for all n because they are certainly measurable and $|f(\theta)\cos(n\theta)| \leq |f(\theta)| \Rightarrow \int_0^{2\pi} |f(\theta)\cos(n\theta)|\,d\theta \leq \int_0^{2\pi} |f(\theta)|\,d\theta < \infty$. Similarly the $f(\theta)\sin(n\theta)$ are integrable and we are not in any trouble defining the A_n and B_n.

Since changing integrands on sets of θ's of measure zero does not affect the Lebesgue integral, it does not really make sense to have $f(0) = f(2\,pi)$ as a restriction any more. Or, we could have it but it can't really help.

1.1.2 Complex Fourier series

From now on, we will consider a different approach because I find it more agreeable. We will take period $L = 1$ instead of 2π and use complex exponentials instead of $\cos$ and $\sin$. We can arrive at $L = 1$ by a change of variable $\theta = 2\pi x$. Then we would have $f\colon [0, 1] \to \mathbb{R}$, and replace (1.1.1) by

$$A_0 + \sum_{n=1}^{\infty}(A_n\cos(2\pi nx) + B_n\sin(2\pi nx)) \tag{1.1.2}$$

where now

$$A_n = \begin{cases} 2\int_0^1 f(x)\cos(2\pi nx)\,dx & \text{for } n > 0\\ \int_0^1 f(x)dx & \text{for } n = 0,\end{cases}$$

$$B_n = 2\int_0^1 f(x)\sin(2\pi nx)\,dx \qquad (n > 0).$$

If we are considering continuous f we would usually insist that $f(0) = f(1)$ or alternatively consider $f\colon \mathbb{R} \to \mathbb{R}$ that satisfies $f(x+1) = f(x)$.

So far a small change and only a slight improvement as there is a different formula for A_0 still. Our next step is to move to complex exponentials

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

(de Moivre's theorem), or

$$e^{2\pi inx} = \cos(2\pi nx) + i\sin(2\pi nx)$$

Since we can add that to

$$e^{-2\pi inx} = \cos(2\pi nx) - i\sin(2\pi nx)$$

to get

$$e^{2\pi inx} + e^{-2\pi inx} = 2\cos(2\pi nx) \Rightarrow \cos(2\pi nx) = \frac{e^{2\pi inx} + e^{-2\pi inx}}{2}$$

and subtract to get

$$\sin(2\pi nx) = \frac{e^{2\pi inx} - e^{-2\pi inx}}{2i},$$

it follows that everything we can express in terms of the cosines and sines ($\cos(2\pi nx)$ $(n \geq 0)$ and $\sin(2\pi nx)$ $(n > 0)$) can equally well be expressed in terms of the building blocks

$$\phi_n(x) = e^{2\pi inx} \qquad (n \in \mathbb{Z}).$$

We get then to this definition

1.1.2.1 Definition (Subject to future explanation)**.** For $f\colon [0,1] \to \mathbb{C}$ integrable, we define the *Fourier coefficients* (complex) of f to be

$$\hat{f}(n) = \int_0^1 f(x)e^{-2\pi inx}\,dx \qquad (n \in \mathbb{Z})$$

and the (complex) *Fourier series* for f to be

$$\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2\pi inx} \tag{1.1.3}$$

We will then be left with these **questions**:

Q1. How to make sense of integrals of complex functions? (This part is rather easy.)

Q2. Does the series converge in any sensible way? and is the sum f?

1.1.2.2 Example (Exercise). (i) Express $f(x) = 3\cos(4\pi x) + 7\cos(6\pi x) - \sin(2\pi x)$ in terms of the complex exponentials $\phi_n(x)$ $(n \in \mathbb{Z})$.

(ii) Express $f(x) = \phi_0(x) + (3+2i)\phi_1(x) + (3-2i)\phi_{-1}(x)$ in terms of sines and cosines.

1.1.3 Complex integrals

As mentioned above, the interpretation of integrals of $\mathbb{C}$-valued integrands $f\colon [0,1,] \to \mathbb{C}$ is not difficult. We take $u(x) = \operatorname{Re} f(x)$ and $v(x) = \operatorname{Im} f(x)$ so that $f(x) = u(x) + iv(x)$ and then we define

$$\int_0^1 f(x)\,dx = \int_0^1 u(x)\,dx + i \int_0^1 v(x)\,dx.$$

But that leaves us with two possible ways to interpret these integrals of $\mathbb{R}$-valued functions:

1. We can use the Riemann integral and require u and v (or equivalently f) to be continuous.

2. We can use the Lebesgue integral and require u and v to be Lebesgue integrable.

   As discussed above, u then has toe be Lebesgue measurable and $\int_0^1 |u(x)|\,dx < \infty$. Same for v. If u and v are measurable then so is $|f|$ as $|f(x)| = \sqrt{(u(x))^2 + (v(x))^2}$. So $\int_0^1 |f(x)|\,dx$ makes sense but might be ∞. However, as $|f(x)| \le |u(x)| + |v(x)|$ we must have $\int_0^1 |f(x)|\,dx < \infty$ if u and v are integrable. As $|u(x)| \le |f(x)|$ and $|v(x)| \le |f(x)|$, then we can summarise what is needed as

   - u and v measurable
   - $\int_0^1 |f(x)|\,dx < \infty$

For calculations we will usually fall back on the Riemann theory and the fact that the Lebesgue integral agrees with the Lebesgue integral for continuous integrands. To compute the Riemann integral $\int_0^1 u(x)\,dx$ we want an antiderivative $U(x)$ (with $U'(x) \equiv u(x)$) and of course $\int_0^1 u(x)\,dx = [U(x)]_0^1 = U(1) - U(0)$. Same for v and an antiderivative V.

We can make lief easier if we look at derivatives $F'(x)$ of $\mathbb{C}$-valued functions of a real variable. We can either

- say that $F(x) = U(x) + iV(x)$ is differentiable if $U = \operatorname{Re} F$ and $V = \operatorname{Im} F$ are both differentiable and define

  $$F'(x) = U(x) + iV(x)$$

  or

- we can define

  $$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h}$$

  (if the limit exists). [We need to modify at the end points of the domain.] (Aside: this is not what you study in complex analysis — there h is complex but here it is real.)

Whichever way we do it we get the usual basic rules for differentiation of $\mathbb{C}$-valued functions of x:

- derivative of a sum = the sum of the derivatives

- constant multiples
$$\frac{d}{dx}(\lambda F(x)) = \lambda F'(x)$$
if $\lambda \in \mathbb{C}$ is constant and $F'(x)$ exists

- Product rule
$$\frac{d}{dx}(F(x)G(x)) = F'(x)G(x) + F(x)G'(x)$$

We can check these either by writing both sides out in terms of real and imaginary parts, or by running through the usual proofs for the $\mathbb{R}$-valued case. (Perhaps we can get away without a chain rule. If we compose with a real function $x = x(t)$ and consider $f(x) = f(x(t))$ we are ok, but for a composition $g(f(x)$ we need complex differentiablity of g to get a normal chain rule, or to use partial derivatives of g.)

From the product rule (by integrating both sides) we get integration by parts:
$$\int_0^1 u\,dv = [uv]_0^1 - \int_0^1 v\,du$$
where now u and v have become $\mathbb{C}$-valued differentiable functions.

We can also say that if $F'(x) = f(x)$ and f is continuous, then
$$\int_0^1 f(x)\,dx = [F(x)]_0^1 = F(1) - F(0)$$
as in the $\mathbb{R}$-valued case (because this is just combining the U and V remarks above about $\int_0^1 u(x)\,dx$ and $\int_0^1 v(x)\,dx$ into one complex calculation).

1.1.3.1 Lemma.
$$\frac{d}{dx}e^{ax} = ae^{ax}$$
for $a \in \mathbb{C}$*.*

Proof. Say $a = \alpha + i\beta$ in terms of its real and imaginary parts. Then
$$e^{ax} = e^{\alpha x + i\beta x} = e^{\alpha x}e^{i\beta x}$$
and we will use the product rule.

First look at
$$e^{i\beta x} = \cos(\beta x) + i\sin(\beta x)$$
so that
$$\frac{d}{dx}e^{i\beta x} = -\beta\sin(\beta x) + i\beta\cos(\beta x) = i\beta(\cos(\beta x) + i\sin(\beta x)) = i\beta e^{i\beta x}$$
(the result we want when a is purely imaginary).

Then, from the product rule,
$$\frac{d}{dx}e^{ax} = \frac{d}{dx}(e^{\alpha x}e^{i\beta x}) = \alpha e^{\alpha x}e^{i\beta x} + i\beta e^{\alpha x}e^{i\beta x} = (\alpha + i\beta)e^{ax} = ae^{ax}.$$
□

1.1.3.2 Example. For $\phi_n(x) = e^{2\pi i n x}$, we have

$$\langle \phi_n, \phi_m \rangle = \int_0^1 \phi_n(x)\overline{\phi_m(x)}\,dx = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

To verify this we compute

$$\int_0^1 \phi_n(x)\overline{\phi_m(x)}\,dx = \int_0^1 e^{2\pi i n x} e^{-2\pi i m x}\,dx = \int_0^1 e^{2\pi i (n-m)x}\,dx$$

If $n = m$, this is just $\int_0^1 1\,dx = 1$ and otherwise we have

$$\left[\frac{1}{2\pi i(n-m)} e^{2\pi i(n-m)x}\right]_0^1 = 0$$

1.1.4 Formal definitions

When we consider vector spaces, they will often be vector spaces of functions defined on some set S. $V = \{f \colon S \to \mathbb{C}\}$ is a vector space when we define the vector space operations

V-Op1 sum

$(f+g)(s) = f(s) + g(s)$ for $s \in S$, $f, g \in V$

V-Op2 scalar multiplication

$(\lambda f)(s) = \lambda f(s)$ for $s \in S$, $\lambda \in \mathbb{C}$, $f \in V$

The zero function is the function $f(s) \equiv 0$. [Fancy notation would be $\mathbb{C}^S$ for this V.]

1.1.4.1 Definition. $C[0,1] = \{f \colon [0,1] \to \mathbb{C} : f \text{ continuous}\}$.
$CP[0,1] = \{f \colon [0,1] \to \mathbb{C} : f \text{ continuous and } f(0) = f(1)\}$.

(Aside: C for continuous and P for periodic.)

1.1.4.2 Lemma. *$C[0,1]$ and $CP[0,1]$ are vector spaces (subspaces of $\mathbb{C}^{[0,1]}$, or of V as above with $S = [0,1]$).*

1.1.4.3 Definition. An *inner product space* (also known as a pre-Hilbert space) is a vector space V over $\mathbb{K}$ (= $\mathbb{R}$ or $\mathbb{C}$) together with a map

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{K}$$

satisfying (for $u, v, w \in V$ and $\lambda \in \mathbb{K}$):

(i) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$

(ii) $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$

(iii) $\langle v, u\rangle = \overline{\langle u, v\rangle}$

(iv) $\langle u, u\rangle \geq 0$

(v) $\langle u, u\rangle = 0 \Rightarrow u = 0$

Note that it follows from the first 3 properties that:

(i)' $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$

(ii)' $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$

If we have all but property (v) we call it a semi-inner product space.

1.1.4.4 Lemma (Cauchy-Schwarz)**.** *If $(V, \langle\cdot,\cdot\rangle)$ is an inner product space (or even a semi-inner product space) and $u, v \in V$, then*

$$|\langle u, v\rangle| \leq \sqrt{\langle u, u\rangle\langle v, v\rangle}$$

Proof. The usual proof starts from $\langle u + \lambda v, u + \lambda v\rangle \geq 0$, expands that out and chooses a suitable λ to get the result. □

1.1.4.5 Definition. A *norm* on a vector space V over $\mathbb{K}$ (where $\mathbb{K}$ can be $\mathbb{C}$ or $\mathbb{R}$) is a map $\|\cdot\|: V \to \mathbb{R}$ satisfying

(i) $\|v\| \geq 0$ for $v \in V$

(ii) $\|u + v\| \leq \|u\| + \|v\|$ for $u, v \in V$ (*triangle inequality*)

(iii) $\|\lambda v\| = |\lambda|\|v\|$ for $\lambda \in \mathbb{K}$ and $v \in V$

(iv) $\|v\| = 0 \Rightarrow v = 0$

If we have the properties (i), (ii) and (ii), we say we have a seminorm.

1.1.4.6 Lemma. *If $(V, \langle\cdot,\cdot\rangle)$ is an inner product space, then we can define a norm on V by*

$$\|v\| = \sqrt{\langle v, v\rangle} \qquad (v \in V).$$

(If we just have a semi-inner product we get a seminorm.)

Proof. The only complicated part to prove is the triangle inequality and that relies on Cauchy-Schwarz. □

1.1.4.7 Example. On $C[0,1]$ or $CP[0,1]$ we can define an inner product by

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

There are 3 norms on $C[0,1]$ or $CP[0,1]$ that we will be considering:

a. $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$

b. $\|f\|_2 = \sqrt{\langle f, f\rangle} = \sqrt{\int_0^1 |f(x)|^2\,dx}$

c. $\|f\|_1 = \int_0^1 |f(x)|\,dx$

It is not difficult to check that the above gives an inner product, and then $\|\cdot\|_2$ is the norm arising from the inner product. It is in fact easy to check that $\|\cdot\|_\infty$ and $\|\cdot\|_1$ are norms.

In some sense $\|\cdot\|_\infty$ is the appropriate norm to use on $C[0,1]$ or $CP[0,1]$ because it makes the space complete (or a Banach space). In the other norms, we can very quickly get into trouble because the space is not complete.

1.1.4.8 Definition. If V is a vector space, then an infinite subset $S \subseteq V$ is called *linearly independent* if each finite subset of S is linearly independent, that is if for each choice $s_1, s_2, \ldots, s_n$ of n distinct elements of S (and each possible $n \geq 1$) the only linear combination

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = 0$$

is the trivial combination where $a_j = 0$ for $1 \leq j \leq n$.

1.1.4.9 Example. The function $\phi_n(x) = e^{2\pi i n x}$ ($n \in \mathbb{Z}$) are linearly independent in $CP[0,1]$.

Proof. We take finitely many n's, say $n_1, n_2, \ldots, n_k$ are distinct integers. Suppose

$$\lambda_1 \phi_{n_1} + \lambda_2 \phi_{n_2} + \cdots + \lambda_k \phi_{n_k} = 0.$$

We can choose N large enough that $-N \leq n_j \leq N$ for $1 \leq j \leq k$ and define

$$a_n = \begin{cases} \lambda_j & \text{if } n = n_j \text{ some } j \\ 0 & \text{otherwise} \end{cases}$$

then we have

$$a_{-N}\phi_{-N} + a_{-N+1}\phi_{-N+1} + \cdots + a_n \phi_N = 0.$$

We can write that more succinctly as

$$\sum_{n=-N}^{N} a_n \phi_n = 0.$$

We want to show $a_n = 0$ must hold for all n.

Method using inner products. Take inner product with a fixed ϕ_m to get

$$\left\langle \sum_{n=-N}^{N} a_n \phi_n, \phi_m \right\rangle = \langle 0, \phi_m \rangle = 0.$$

That gives

$$\sum_{n=-N}^{N} a_n \langle \phi_n, \phi_m \rangle = 0$$

By Example 1.1.3.2, that simplifies to

$$a_m 1 + 0 = 0 \Rightarrow a_m = 0.$$

As this hold for all m, we are done.

Method using direct linear algebra. To say that $\sum_{n=-N}^{N} a_n \phi_n = 0$ means that $\sum_{n=-N}^{N} a_n \phi_n(x) = 0$ for each $x \in [0,1]$. Notice that $\phi_n(x) = (\phi_1(x))^n$ (because $e^{2\pi i n x} = (e^{2\pi i x})^n$) and so we have

$$\sum_{n=-N}^{N} a_n (\phi_1(x))^n = 0$$

for all $x \in [0,1]$. If we pick $2N+1$ different $x \in [0,1)$, say $x_j = j/(2N+1)$ for $j = 0, 1, \ldots, 2N$ and put $\zeta_j = \phi_1(x_j)$ we have a system of $2N+1$ linear equations

$$\sum_{n=-N}^{N} a_n \zeta_j^n = 0 \qquad (0 \le j \le 2N)$$

in $2N+1$ unknowns a_n $(-N \le n \le N)$. In matrix form this can be written

$$\begin{bmatrix} \zeta_0^{-N} & \zeta_0^{-(N-1)} & \cdots & \zeta_0^N \\ \zeta_1^{-N} & \zeta_1^{-(N-1)} & \cdots & \zeta_N^N \\ \vdots & & & \\ \zeta_{2N}^{-N} & \zeta_{2N}^{-(N-1)} & \cdots & \zeta_{2N}^N \end{bmatrix} \begin{bmatrix} a_{-N} \\ a_{-N+1} \\ \vdots \\ a_N \end{bmatrix} = 0$$

The matrix is invertible, or has nonzero determinant because (if we factor ζ_j^{-N} from each row $j+1$)

$$\det \begin{bmatrix} \zeta_0^{-N} & \zeta_0^{-(N-1)} & \cdots & \zeta_0^N \\ \zeta_1^{-N} & \zeta_1^{-(N-1)} & \cdots & \zeta_N^N \\ \vdots & & & \\ \zeta_{2N}^{-N} & \zeta_{2N}^{-(N-1)} & \cdots & \zeta_{2N}^N \end{bmatrix} = (\zeta_0 \zeta_1 \cdots \zeta_{2N})^{-N} \det \begin{bmatrix} 1 & \zeta_0 & \zeta_0^2 & \cdots & \zeta_0^{2N+1} \\ 1 & \zeta_1 & \zeta_1^2 & \cdots & \zeta_N^N \\ \vdots & & & & \\ 1 & \zeta_{2N} & \zeta_{2N}^2 & \cdots & \zeta_{2N}^{2N+1} \end{bmatrix}$$

and the latter is what is called a Vandermonde determinant, well-known to be nonzero provided the numbers $\zeta_0, \zeta_1, \ldots, \zeta_{2N}$ are all distinct (which they are in our case). □

1.1.4.10 Definition. If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space, then $u, v \in V$ are called *orthogonal* if $\langle u, v \rangle = 0$ (and we may write $u \perp v$).

A subset $S \subseteq V$ is called *orthogonal* if $u \perp v$ whenever $u, v \in S$ and $u \neq v$.

A subset $S \subseteq V$ is called *orthonormal* if it is orthogonal and satisfies $\langle u, u \rangle = 1$ for each $u \in S$ (or $\|u\| = 1$).

1.1.4.11 Proposition. *Orthonormal subsets S of inner product spaces V are always linearly independent.*

Proof. This is essentially one of the proofs we just gave for Example 1.1.4.9.

If S is finite, so that $S = \{s_1, s_2, \dots, s_n\}$ for some $n \geq 0$, then we consider a linear combination

$$v = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = 0$$

Compute $\langle v, s_j \rangle$ for any j with $1 \leq j \leq n$ and get $0 = \langle v, s_j \rangle$ for any j with $1 \leq j \leq n$ and get $0 = \langle v, s_j \rangle = a_j$. Thus $a_j = 0$ for $1 \leq j \leq n$.

In the case of infinite S, the argument is essentially the same except that $s_1, s_2, \dots, s_n$ are now any n distinct elements of S (rather than the whole of S). □

1.1.4.12 Remark. If you examine the argument carefully, $n = 0$ was allowed in the finite case. The empty set $S = \emptyset$ of vectors is linearly independent in any vector space, either by convention or by a literal reading of the definition. It is also orthonormal in any inner product space.

1.1.5 Integrable functions

For $f \in CP[0,1]$ we have a neat connection

$$\hat{f}(n) = \langle f, \phi_n \rangle = \int_0^1 f(x) e^{-2\pi i n x} \, dx$$

between Fourier coefficient and inner products and that seems to suggest strongly that we should consider Fourier series in the context of inner product spaces. While there is a lot to be said for that idea (and we will see that later), the most general context where we can make sense of $\hat{f}(n)$ is for integrable f, as in Definition 1.1.2.1. This leads us to consider this space:

1.1.5.1 Definition. We define

$$\mathcal{L}^1([0,1]) = \{f \colon [0,1] \to \mathbb{C} : f \text{ measurable and } \int_0^1 |f| \, dx < \infty\}.$$

This space is called the space of Lebesgue integrable functions (on $[0,1]$).

(We could define an $\mathbb{R}$-valued version but we will concentrate on the $\mathbb{C}$-valued case.)

We also define the magnitude of $f \in \mathcal{L}^1([0,1])$ as

$$\|f\|_1 = \int_0^1 |f| \, d\mu$$

1.1.5.2 Lemma. *With the usual vector space operations on functions $\mathcal{L}^1([0,1])$ is a vector space over $\mathbb{C}$.*

Also $\|\cdot\|_1$ defines a semi-norm *on $\mathcal{L}^1([0,1])$.*

This semi-norm is not a norm*, that is it does not satisfy $\|f\|_1 = 0$ only when $f = 0$. In fact all we can say is that $\|f\|_1 = 0 \iff f(x) = 0$ holds for almost every $x \in [0,1]$ (or $\iff \mu\{x \in [0,1] : f(x) \neq 0\} = 0$).*

Proof. The proof of this is fairly straightforward.

The semi-norm properties are very easy to check and the statement about $\|f\|_1 = 0$ being equivalent to $f(x) = 0$ almost everywhere was discussed in MA2224. □

1.1.5.3 Lemma (Properties of Fourier coefficients)**.**

(i) Definition 1.1.2.1 is valid for $f \in \mathcal{L}^1([0,1])$ *(that is,* $f(x)e^{-2\pi inx}$ *is integrable if* f *is).*

(ii) For each n*, the* n^{th} *Fourier coefficient map*

$$f \mapsto \hat{f}(n) \colon \mathcal{L}^1([0,1]) \to \mathbb{C}$$

is a linear transformation (called a linear functional since the values are in the vector space of scalars).

(iii)

$$\sup_{n\in\mathbb{Z}} |\hat{f}(n)| \le \|f\|_1$$

holds for $f \in \mathcal{L}^1([0,1])$

(iv) If $f, g \in \mathcal{L}^1([0,1])$ *have* $f(x) = g(x)$ *almost everywhere, then* $\hat{f}(n) = \hat{g}(n)$ *for all* $n \in \mathbb{Z}$ *(that is,* f *and* g *have the same Fourier series).*

Proof. (i) We need to know that products of measurable functions are measurable, even in the $\mathbb{C}$-valued case (where when we say $f(x)$ is measurable we mean that $\operatorname{Re} f(x)$ and $\operatorname{Im} f(x)$ are both measurable. This not complicated to show as

$$\begin{aligned} f(x)g(x) &= (\operatorname{Re} f(x) + i \operatorname{Im} f(x))(\operatorname{Re} g(x) + i \operatorname{Im} g(x)) \\ &= (\operatorname{Re} f(x) \operatorname{Re} g(x) - \operatorname{Im} f(x) \operatorname{Im} g(x)) + i(\operatorname{Re} f(x) \operatorname{Im} g(x) + \operatorname{Im} f(x) \operatorname{Re} f(x)) \end{aligned}$$

Next $|f(x)e^{-2\pi inx}| = |f(x)|$ and so $f(x)e^{-2\pi inx}$ is integrable.

(ii) This means that $\widehat{(f+g)}(n) = \hat{f}(n) + \hat{g}(n)$ and $\widehat{(\lambda f)}(n) = \lambda \hat{f}(n)$ (for $f, g \in \mathcal{L}^1([0,1])$ and $\lambda \in \mathbb{C}$), both of which are easy to see from the integrals.

(iii)

$$|\hat{f}(n)| = \left| \int_{[0,1]} f(x)e^{-2\pi inx} \, d\mu(x) \right| \le \int_{[0,1]} |f(x)e^{-2\pi inx}| \, d\mu(x) = \|f\|_1$$

for each n. That does rely on the triangle inequality in the form $\left|\int_0^1 f(x)\,dx\right| \le \int_0^1 |f(x)|\,dx$ (valid for integrable f). To get the complex version from the real one we can choose $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ and

$$\lambda \int_0^1 f(x)\,dx$$

real and positive. Then

$$\left|\int_0^1 f(x)\,dx\right| = \left|\lambda \int_0^1 f(x)\,dx\right| = \lambda \int_0^1 f(x)\,dx = \int_0^1 \lambda f(x)\,dx = \int_0^1 \mathrm{Re}(\lambda f(x))\,dx$$

The latter integral is then

$$\leq \int_0^1 |\lambda f(x)|\,dx = \int_0^1 |f(x)|\,dx.$$

From $|\hat{f}(n)| \leq \|f\|_1$ for each $n \in \mathbb{Z}$ we get

$$\sup_{n\in\mathbb{Z}} |\hat{f}(n)| \leq \|f\|_1$$

(iv) If $f(x) = g(x)$ almost everywhere, then $f - g$ is zero almost everywhere and so $\|f - g\|_1 = 0$ and that implies by the above $\widehat{(f - g)}(n) = 0$. By linearity, this gives $\hat{f}(n) - \hat{g}(n) = 0 \Rightarrow \hat{f}(n) = \hat{g}(n)$ (for each $n \in \mathbb{Z}$).

□

1.1.5.4 Remark. Property (iii) is much less than we would need to show any kind of convergence of the Fourier series $\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2\pi inx}$

1.2 The unit circle

We take a slightly different approach now and consider $CP[0,1]$ as continuous functions on the unit circle. This will be an important step for us because the circle is a group under multiplication of complex scalars, an abelian group. It is also compact in $\mathbb{C}$.

We will use this example to introduce some of the more abstract ideas we will develop later. When we get to Fourier series on the circle (in Definition 1.3.13) we will be only slightly restating the classical definition we already recalled in Definition 1.1.2.1.

1.2.1 Notation. $\mathbb{T}$ will denote the unit circle $\{\zeta \in \mathbb{C} : |\zeta| = 1\}$.

(Some books may use S^1 instead of $\mathbb{T}$.)

$\mathbb{T}$ is a *metric space* with the (restriction) of the usual absolute value distance $d(\zeta, \eta) = |\zeta - \eta|$ from $\mathbb{C}$.

$C(\mathbb{T})$ means the space $\{F : \mathbb{T} \to \mathbb{C} : F \text{ continuous}\}$ of continuous complex-valued functions on $\mathbb{T}$. [We try to use F for functions on the circle, lower case f for functions on line.]

1.2.2 Definition. A *group* is a set G with a binary operation $G \times G \to G$ satisfying various axioms. If we write the binary operation as a product gh for $g, h \in G$, the axioms are

(G1) $gh \in G$ for $g, h \in G$

(G2) (associative law) $(gh)k = g(hk)$ for $g, h, k \in G$

(G3) (identity element) there is $e \in G$ such that $eg = g$ and $ge = g$ hold for each $g \in G$

(G4) (existence of inverses) For each $g \in G$ there is an inverse element $h \in G$ such that $gh = e = hg$.

The inverse of $g \in G$ is unique and written g^{-1}.

A group G is called *abelian* if $gh = hg$ holds for each $g, h \in G$. (Abelian groups are sometimes written using additive notation, so that $g + h$ replaces gh, the identity element is written 0 and $-g$ replaces g^{-1}.)

A subset $H \subseteq G$ of a group G is called a *subgroup* if it forms a group using the same operation as G (restricted to H). Since H cannot be empty (as it has to have its own identity) we can show that a subset $H \subseteq G$ is a subgroup if and only if it satisfies

(SG1) $H \neq \emptyset$

(SG2) $h, k \in H \Rightarrow hk^{-1} \in H$.

(We can also write (SG1) as $e \in H$ because (SG2) tells us that $h \in H \Rightarrow e = hh^{-1} \in H$ (and (SG1) says $\exists h \in H$). And (SG2) can be written in a more long winded way as $h, k \in H \Rightarrow hk \in H$ and $k \in H \Rightarrow k^{-1} \in H$.)

The unit interval $[0, 1]$ and the circle $\mathbb{T}$ have in common that they are compact metric spaces. Unlike $[0, 1]$, $\mathbb{T}$ is also a group.

All our topologies will come from metrics. For completeness we recall what the metric topology is on a metric space (X, d) (see MA2223).

1.2.3 Definition. Given any set X of points and a function $d \colon X \times X \to [0, \infty) \subset \mathbb{R}$ with these 3 properties:

(M1) $d(z, w) \geq 0$ with equality if and only if $z = w$;

(M2) $d(z, w) = d(w, z)$;

(M3) $d(z, w) \leq d(z, v) + d(v, w)$ (triangle inequality),

we say that d is a *metric on the space* X and we call the combination (X, d) a *metric space*.

1.2.4 Notation. In any metric space (X, d) we define *open balls* as follows. Fix any point $x_0 \in X$ (which we think of as the centre) and any $r > 0$. Then the *open ball* of radius r centre x_0 is

$$B(x_0, r) = \{x \in X : d(x, x_0) < r\}.$$

The *closed ball* with the same centre and radius is

$$\bar{B}(x_0, r) = \{x \in X : d(x, x_0) \leq r\}.$$

1.2.5 Definition (Open sets in a metric space)**.** For a metric space (X, d) and a subset $G \subseteq X$ and a point $x \in G$, we say that x is *an interior point of* G if there is a ball $B(x, r)$ of some positive radius $r > 0$ centred at x so that $B(x, r) \subset G$. We write G° for the set of interior points of G.

A set $G \subseteq X$ is called *open* if each $x \in G$ is an interior point of G (that is, if $G^\circ = G$).

The *metric topology* on (X, d) is the collection $\mathscr{T}_d$ of all subsets $G \subseteq X$ that are open.

A set $S \subseteq X$ is called *closed* if its complement $X \setminus S$ is open.

1.2.6 Definition. For a metric spaces (X, d) a subset $K \subseteq X$ is called *compact* if **every** sequence $(k_n)_{n=1}^\infty$ in K has **some** convergent subsequence with a limit in K, that is a subsequence $(k_{n_j})_j$ such that $\exists \lim_{j\to\infty} k_{n_j} \in K$.

X is compact if $K = X$ is a compact subset.

There is a more general definition for topological spaces (involving open covers) but we will manage without using that.

1.2.7 Lemma. *(i) $\mathbb{T}$ is an abelian group with the operation of multiplication of complex numbers (and identity element $1 \in \mathbb{T}$).*

(ii) $\mathbb{T}$ is a compact metric space.

(iii) $C(\mathbb{T})$ is a vector space over $\mathbb{C}$ with addition and multiplication by complex scalars defined pointwise.

(iv) There is a vector space isomorphism

$$T\colon C(\mathbb{T}) \to CP[0, 1]$$

given by

$$TF(x) = F(e^{2\pi i x})$$

Proof. (i) It is well known (and follows from the axioms for a field) that $\mathbb{C}_* = \{z \in \mathbb{C} : z \neq 0\}$ (the nonzero elements in $\mathbb{C}$) is an abelian group under multiplication (with identity element 1) and since $|\zeta\eta| = |\zeta||\eta|$ and $|1/\zeta| = 1/|\zeta|$ and $1 \in \mathbb{T}$, it is easy to see that $\mathbb{T}$ is a subgroup of $\mathbb{C}_*$.

(ii) $\mathbb{T}$ is compact because of the Heine-Borel theorem (as it is closed and bounded in $\mathbb{C} = \mathbb{R}^2$).

(iii) To show that $C(\mathbb{T})$ is a vector space, we could check that it is a vector subspace of the vector space of all functions $f\colon \mathbb{T} \to \mathbb{C}$. That means showing that (the constant function) $0 \in C(\mathbb{T})$ and that $f + \lambda g$ is continuous on $\mathbb{T}$ if $f, g \in C(\mathbb{T})$ and $\lambda \in \mathbb{C}$.

(iv) Since $\phi_1(x) = e^{2\pi i x}$ is a continuous function from $[0, 1]$ to $\mathbb{T}$, and our definition of TF (or $T(F)$ if you prefer that) is

$$TF = F \circ \phi_1,$$

TF will be continuous on $[0, 1]$ for each $F \in C(\mathbb{T})$. Also $\phi_1(0) = 1 = \phi_1(1)$ and so $TF(0) = TF(1)$, which is the additional property we need for $TF \in CP[0, 1]$.

It is easy to see that $T\colon C(\mathbb{T}) \to CP[0,1]$ is a linear transformation.

To show it is an isomorphism of vector spaces, what we need is to show T is surjective. If $f \in CP[0,1]$ we want to define $F \in C(\mathbb{T})$ so that $TF = f$ and that forces us to take $F(e^{2\pi i x}) = f(x)$. Each $\zeta \in \mathbb{T}$ can be written as $\zeta = e^{2\pi i x}$ for $x \in [0,1]$ and in fact x is uniquely determined by ζ except that for $\zeta = 1$ we have two choices $x = 0$ and $x = 1$. Since we have $f(0) = f(1)$ we have $F(1)$ defined unambiguously. To show that F is continuous, we can rely on complex logarithms $\log \zeta$ and $x = 1/(2\pi i) \log \zeta$, provided we exclude $\zeta = 1$ and choose the branch of $\log$ so that $\log \zeta \in i(0, 2\pi)$. This shows $F(\zeta) = f(1/(2\pi i) \log \zeta)$ is continuous on $\mathbb{T} \setminus \{1\}$. To deal with continuity at $\zeta = 1$ we could exclude $\zeta = -1$ and choose another branch of $\log \zeta \in i(-\pi, \pi)$. We also need to extend $f \in CP[01]$ to be (continuous and) periodic on $\mathbb{R}$ and then again $F(\zeta) = f(1/(2\pi i) \log \zeta)$ is continuous on $\mathbb{T} \setminus \{-1\}$. So $F \in C(\mathbb{T})$ and $TF = f$.

□

1.2.8 Remark. Under the isomorphism T above, the functions on $\mathbb{T}$ that correspond to $\phi_n(x) = e^{2\pi i n x} \in CP[0,1]$ are $\chi_n(\zeta) = \zeta^n$. (That is $T\chi_n = \phi_n$ because $\phi_n(x) = (e^{2\pi i x})^n$.)

Each such χ_n is a group homomorphism $\chi_n\colon \mathbb{T} \to \mathbb{T}$ and is furthermore continuous. We will move now to more general groups G than $G = \mathbb{T}$ and a key role in Fourier analysis of functions $F\colon G \to \mathbb{C}$ will be played by characters. Characters will be defined as continuous group homomorphisms $\chi\colon G \to \mathbb{T}$.

1.3 Topological abelian groups

1.3.1 Definition. By a *metrizable space* X we mean a set X with a collection $\mathscr{T}$ of subsets of X such that there exists some metric d on X with $\mathscr{T} = \mathscr{T}_d$.

Two metrics d and d' on X are called *equivalent* if they both give the same open sets, that is if $\mathscr{T}_d = \mathscr{T}_{d'}$.

1.3.2 Remark. Since continuity of functions $f\colon X \to Y$ between metric spaces (X, d) and (Y, ρ) can be expressed solely using open sets, we take the view that the metric is not so important to us. But we stop short of allowing arbitrary topological spaces because some things are simpler for metric spaces (such as the definition of compactness — recall remarks in the proof of Lemma 1.2.7). Also limits of sequences can be used to describe continuity and closures (but do not suffice in general topological spaces).

However, we may want to switch from one metric to another at times. For instance if we have one metric d on X, then

$$d'(x,y) = \frac{d(x,y)}{1 + d(x,y)}$$

defines an equivalent metric d', different from d. If, for example, $X = \mathbb{R}$ and $d(x,y) = |x - y|$, then d' has the (strange but sometimes useful) property that $d'(x,y) < 1$ always.

1.3.3 Definition. A metric space (X, d) is called *locally compact* if for each point $x_0 \in X$ there is some $r > 0$ (depending on x_0) such that the closed ball $\bar{B}(x_0, r)$ is compact.

We will only use the term locally compact for metrizable spaces.

1.3.4 Example. $\mathbb{R}$ with the usual absolute value metric $d(x,y) = |x-y|$ is locally compact because if we take $r=1$, then $\bar{B}(x_0,r) = [x_0-1, x_0+1]$ is compact.

$\mathbb{T}$ is locally compact because it is a compact metric space and so it follows that every closed ball $\bar{B}(\zeta_0,r)$ with $r>0$ is compact (or if we take $r=2$, we have $\bar{B}(\zeta_0,r) = \mathbb{T}$ compact (any $\zeta_0 \in \mathbb{T}$)).

1.3.5 Definition. By a *topological group* we will mean a group G which is also a metrizable space where multiplication and inversion are continuous.

In more detail, we suppose d is a metric on G that gives rise to the topology, and then we defined a metric on $G \times G$ by $\rho((g_1,g_2),(h_1,h_2)) = d(g_1,h_1) + d(g_2,h_2)$. Then $G \times G$ is a metrizable space. We insist then that the maps

(TG1) multiplication : $G \times G \to G$ $((g,h) \mapsto gh)$

(TG2) inversion : $G \to G$ $(g \mapsto g^{-1})$

are each continuous.

1.3.6 Examples.

(a) $\mathbb{T}$ is a topological group

Proof. To prove that multiplication is continuous from $\mathbb{T} \times \mathbb{T} \to \mathbb{T}$ fix $(\zeta_0,\eta_0) \in \mathbb{T} \times \mathbb{T}$ and $\varepsilon > 0$.

To show continuity at (ζ_0,η_0) we show that it is possible to find $\delta > 0$ such that

$$(\zeta,\eta) \in \mathbb{T} \times \mathbb{T}, \rho((\zeta,\eta),(\zeta_0,\eta_0)) = |\zeta-\zeta_0| + |\eta-\eta_0| < \delta \Rightarrow |\zeta\eta - \zeta_0\eta_0| < \varepsilon$$

Looking at

$$\begin{aligned} |\zeta\eta - \zeta_0\eta_0| &= |\zeta(\eta-\eta_0) + (\zeta-\zeta_0)\eta_0| \\ &\leq |\zeta||\eta-\eta_0| + |\zeta-\zeta_0||\eta_0| \\ &= |\eta-\eta_0| + |\zeta-\zeta_0| \end{aligned}$$

we can see that $\delta = \varepsilon$ will work.

Being continuous at each point $(\zeta_0,\eta_0) \in \mathbb{T} \times \mathbb{T}$, we have that multiplication is continuous. (Another proofs could use sequences tending to (ζ_0,η_0).)

The argument for inversion is also quite elementary. □

(b) $(\mathbb{R},+)$, by which we mean $\mathbb{R}$ with addition as the group operation, is a topological group (using the usual metric topology on $\mathbb{R}$)

Proof. So now we have to prove that addition : $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ $((x,y) \mapsto x+y)$ and negation : $\mathbb{R} \to \mathbb{R}$ $(x \mapsto -x)$ are continuous.

We know that these are continuous, but they are not at all hard to check. □

(c) $(\mathbb{Z}, +)$ is a topological group. (On $\mathbb{Z}$ we take the usual absolute value distance but then $B(n, 1) = \{n\}$ for each $n \in \mathbb{Z}$ and so every subset of $\mathbb{Z}$ is open. It is called a discrete space.)

Proof. Every subset of $\mathbb{Z} \times \mathbb{Z}$ is also open (or $\mathbb{Z} \times \mathbb{Z}$ is discrete) and so every function from $\mathbb{Z} \times \mathbb{Z}$ to any other metric space must be continuous. In particular addition is continuous.

Similarly, negation $n \mapsto -n$ is continuous on $\mathbb{Z}$. □

(d) If G is a finite abelian group (for example the cyclic group of order n for some $n > 1$) then G is a topological group with the discrete topology (or with the metric where $d(g, h) = 1$ when $g \neq h$ and $d(g, g) = 0$).

The above examples will be our main concrete examples of topological groups. They are all abelian groups and are locally compact metrizable spaces.

1.3.7 Definition (Characters)**.** If G is a an abelian topological group, then a *character* of G is a continuous group homomorphism $\chi \colon G \to \mathbb{T}$.

1.3.8 Remark. If the group operation on G is written as multiplication, then we need χ to be a continuous map and to satisfy the homomorphism rule

$$\chi(gh) = \chi(g)\chi(h) \qquad (g, h \in G)$$

If $e \in G$ is the identity element, we must have $\chi(e) = 1$ (because $\chi(e) = \chi(ee) = \chi(e)^2 \in \mathbb{T}$).

So, in any group G we can take the trivial character $\chi_0(g) = 1$ (constant map) as one example.

If $G = \mathbb{T}$, examples are $\chi_n(\zeta) = \zeta^n$ for $n \in \mathbb{Z}$. (It is not a coincidence that this ties in with $\phi_n(x) = e^{2\pi i n x} = (e^{2\pi i x})^n$ used in our Fourier series.)

If G is $\mathbb{R}$ with addition as the group operation, we want $\chi \colon \mathbb{R} \to \mathbb{T}$ to satisfy

$$\chi(x + y) = \chi(x)\chi(y) \qquad (x, y \in \mathbb{R})$$

(and that forces $\chi(0) = 1$). Examples are $\chi_t(x) = e^{2\pi i x t}$ (for $t \in \mathbb{R}$).

[Maybe I should warn you that there is a different use of the word ‘character’ for traces of representations of groups (often finite groups), something you might encounter in a different module on algebra.]

1.3.9 Definition (Dual group)**.** If G is an abelian topological group, we define $\hat{G}$ to be the set of all characters $\chi \colon G \to \mathbb{T}$ and introduce a multiplication rule for characters χ_1, χ_1 by

$$(\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g)$$

1.3.10 Proposition. *If G is an abelian topological group, then $\hat{G}$ is again an abelian group with the above multiplication rule and identity element the trivial character χ_0 (given by $\chi_0(g) \equiv 1$).*

Proof. If $\chi_1, \chi_2 \in \hat{G}$, then the product function $\chi_1\chi_2$ is certainly continuous (as a product of two $\mathbb{C}$-valued continuous functions). It is also a character because $|(\chi_1\chi_2)(g)| = |\chi_1(g)||\chi_2(g)| = 1$ ($\forall g \in G$) and

$$\begin{aligned}(\chi_1\chi_2)(gh) &= \chi_1(gh)\chi_2(gh)\\ &= \chi_1(g)\chi_1(h)\chi_2(g)\chi_2(h)\\ &= \chi_1(g)\chi_2(g)\chi_1(h)\chi_2(h)\\ &= (\chi_1\chi_2)(g)(\chi_1\chi_2)(h)\end{aligned}$$

(for $g, h \in G$).

The trivial character $\chi_0(g) \equiv 1$ is an identity for this multiplication on $\hat{G}$ because for $\chi \in \hat{G}$

$$(\chi_0\chi)(g) = \chi_0(g)\chi(g) = 1\chi(g) = \chi(g)$$

and so $\chi_0\chi = \chi$. Similarly $\chi\chi_0 = \chi$ but in fact the multiplication of functions is commutative: $\chi_1\chi_2 = \chi_2\chi_1$ holds for $\chi_1, \chi_2 \in \hat{G}$.

Finally there is an inverse $1/\chi$ for each $\chi \in \hat{G}$. For this we have to check that $1/\chi$ is a group homomorphism from G to $\mathbb{T}$ (so $1/\chi \in \hat{G}$) and $\chi(1/\chi) = \chi_0 = (1/\chi)\chi$ is easy to check.

So $\hat{G}$ is an abelian group. □

Next we compute $\hat{G}$ for some examples.

1.3.11 Proposition. *(a) The characters of the additive group $(\mathbb{R}, +)$ are all of the form $\chi_t(x) = e^{2\pi ixt}$ for a unique $t \in \mathbb{R}$ and $\hat{\mathbb{R}}$ is isomorphic to $\mathbb{R}$ via $\chi_t \mapsto t$.*

(b) The characters of the group $\mathbb{T}$ are all of the form $\chi_n(\zeta) = \zeta^n$ for a unique $n \in \mathbb{Z}$ and $\hat{\mathbb{T}}$ is is isomorphic to $\mathbb{Z}$ via $\chi_n \mapsto n$.

(c) The characters of the additive group $(\mathbb{Z}, +)$ are all of the form $\chi_\zeta(n) = \zeta^n$ for a unique $\zeta \in \mathbb{T}$ and $\hat{\mathbb{Z}}$ is isomorphic to $\mathbb{T}$ via $\chi_\zeta \mapsto \zeta$.

Proof. (a) We can check easily that $\chi_t \in \hat{\mathbb{R}}$ for each $t \in \mathbb{R}$.

Let $\chi\colon \mathbb{R} \to \mathbb{T}$ be a character. By continuity there is $\delta > 0$ so that $|x| < \delta \Rightarrow |\chi(x) - \chi(0)| = |\chi(x) - 1| < 1$. So for $|x| < \delta$, $\operatorname{Re}\chi(x) > 0$ and there is a unique $\theta(x) \in (-\pi/2, \pi/2)$ with $e^{i\theta(x)} = \chi(x)$. If $|x| < \delta$, then

$$\chi(x) = \chi\left(\frac{x}{2} + \frac{x}{2}\right) = \chi(x/2)\chi(x/2) = \chi(x/2)^2$$

Thus $\chi(x) = e^{i\theta(x)} = (e^{i\theta(x)/2})^2 = (\chi(x/2))^2$ and so $\chi(x/2) = \pm e^{i\theta(x)/2}$. Since $-e^{i\theta(x)/2}$ has negative real part, $\chi(x/2) = e^{i\theta(x)/2}$ and $\theta(x/2) = \theta(x)/2$. By induction $\theta(x/2^n) = \theta(x)/2^n$ for $n \in \mathbb{N}$.

Fix $x_0 = \delta/2$ and put $t = \theta(x_0)/(2\pi x_0)$. We have

$$\chi(x_0/2^n) = e^{i\theta(x_0)/2^n} = e^{2\pi i x_0 t/2^n} = e^{2\pi i(x_0/2^n)t}$$

By induction on $k \in \mathbb{N}$, we get

$$\chi\left(\frac{kx_0}{2^n}\right) = \chi\left(\frac{x_0}{2^n} + \cdots + \frac{x_0}{2^n}\right) = (\chi(x_0/2^n))^k = e^{2\pi i(kx_0/2^n)t}$$

and since $\chi(x)\chi(-x) = \chi(0) = 1$, this also holds for negative $k \in \mathbb{Z}$ (and for $k = 0$). Since numbers $x = kx_0/2^n$ are dense in $\mathbb{R}$ and both sides are continuous, we conclude

$$\chi(x) = e^{2\pi ixt}$$

for all $x \in \mathbb{R}$. Thus $\chi = \chi_t$.

If $t \neq t'$, then $\chi_t \neq \chi_{t'}$ because $\chi_t(x) \neq \chi_{t'}(x)$ for small $x > 0$.

So the map $t \mapsto \zeta_t$ is a bijection from $\mathbb{R}$ to $\hat{\mathbb{R}}$. Since $\chi_t(x)\chi_{t'}(x) = e^{2\pi ixt}e^{2\pi ixt'} = e^{2\pi ix(t+t')} = \zeta_{t+t'}(x)$, the map is a group homomorphism from $(\mathbb{R}, +)$ to $\hat{\mathbb{R}}$. Being a bijection, the map is an isomorphism.

(b) We can check easily that $\zeta \mapsto \zeta^n$ is a character of $\mathbb{T}$ for each $n \in \mathbb{Z}$.

If $\chi\colon \mathbb{T} \to \mathbb{T}$ is a character, then, then $\mu(x) = \chi(e^{2\pi ix})$ is a character $\mu \in \hat{\mathbb{R}}$ (because it a composition of the continuous homomorphism $x \mapsto e^{2\pi ix}$ from $\mathbb{R}$ to $\mathbb{T}$ and the continuous homomorphism $\zeta\colon \mathbb{T} \to \mathbb{T}$). Also $\mu(1) = \chi(e^{2\pi i}) = \chi(1) = 1$.

By the (a), there is $t \in \mathbb{R}$ with $\mu = \chi_t$ in the notation from (a), that is $\mu(x) = e^{2\pi ixt}$. Since $\mu(1) = 1$, we have $t \in \mathbb{Z}$ and we write n instead of t. Now

$$\chi(e^{2\pi ix}) = \mu(x) = e^{2\pi ixn} = (e^{2\pi ix})^n$$

for all $x \in \mathbb{R}$, or $\chi(\zeta) = \zeta^n$ for all $\zeta \in \mathbb{T}$. Hence $\chi = \chi_n$ in the notation from (b).

If $n \neq n'$, then if $x \in \mathbb{R}$ is close enough to 0 and $\zeta = e^{2\pi ix}$, $\zeta^n/\zeta^{n'} = e^{2\pi inx}/e^{2\pi in'x} = e^{2\pi i(n-n')x} \neq 1$. Hence $\chi_n \neq \chi_{n'}$.

Thus the map $n \mapsto \chi_n$ is a bijection from $\mathbb{Z}$ to $\hat{\mathbb{T}}$.

Since $\zeta^{n+n'} = \zeta^n\zeta^{n'}$ for $\zeta \in \mathbb{T}$ and $n, n' \in \mathbb{Z}$ we have $\chi_{n+n'} = \chi_n\chi_{n'}$ and this is what is needed to show that the map $n \mapsto \chi_n$ is a group homomorphism. Being a bijection, it is an isomorphism.

(c) If $\chi\colon \mathbb{Z} \to \mathbb{T}$ is a character, put $\zeta = \chi(1)$. Then $\chi(2) = \chi(1+1) = \chi(1)\chi(1) = \zeta^2$ and by induction $\chi(n) = \zeta^n$ for $n \in \mathbb{N}$. Since $\chi(-n) = (\chi(n))^{-1} = \zeta^{-n}$ and $\chi(0) = 1$, we have $\chi(n) = \zeta^n$ for each $n \in \mathbb{Z}$. So $\chi = \chi_\zeta$.

Note that $\chi_\zeta \in \hat{\mathbb{Z}}$ for each $\zeta \in \mathbb{T}$ and $\chi_\zeta = \chi_{\zeta'} \Rightarrow \zeta = \chi_\zeta(1) = \chi_{\zeta'}(1) = \zeta'$. So $\zeta \mapsto \chi_\zeta$ is a bijection from $\mathbb{T}$ to $\hat{\mathbb{Z}}$.

Finally the bijection has $\chi_\zeta\chi_{\zeta'} = \chi_{\zeta\zeta'}$ for $\zeta, \zeta' \in \mathbb{T}$ (because $\chi_\zeta(n)\chi_{\zeta'}(n) = \zeta^n(\zeta')^n = (\zeta\zeta')^n = \chi_{\zeta\zeta'}(n)$ for $n \in \mathbb{Z}$) and so it is a group homomorphism. As a bijection, it is an isomorphism.

□

1.3.12 Remark. It is actually the case that these examples reveal some facts that hold quite often. We will discuss some more technicalities before we get to the results, but for each of the 3 abelian groups G we have $\hat{G}$ is another locally compact abelian group and the dual of $\hat{G}$, that is $\hat{\hat{G}}$, turns out to be G again.

Our approach to Fourier theory is that the right general setting for Fourier's ideas on reconstructing functions from sines and cosines is first to take complex exponentials (as we have done already) and then to consider them as members of an appropriate dual group $\hat{G}$. So the general idea is that all functions (satisfying some restrictions to be specified) from G to $\mathbb{C}$ can be reconstructed by "adding up" linear combinations of characters (in some way to be specified later).

Our first example will be $G = \mathbb{T}$ and functions $F \colon \mathbb{T} \to \mathbb{C}$.

1.3.13 Definition (Fourer series/transform on $\mathbb{T}$)**.** If $F \in C(\mathbb{T})$ we define the Fourier coefficients $\hat{F}(n)$ for $n \in \mathbb{Z}$ by considering the associated $f \in CP[0,1]$ given by

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

and taking

$$\hat{F}(n) = \hat{f}(n) = \int_0^1 f(x) e^{-2\pi i n x}\,dx = \int_0^1 F(e^{2\pi i x})\overline{(e^{2\pi i x})^n}\,dx \tag{1.3.1}$$

Recalling from Examples 1.3.11 (b) that $\hat{\mathbb{T}} = \{\chi_n : n \in \mathbb{Z}\}$ can be identified with $\mathbb{Z}$, we define the *Fourier transform* of F to be the function $\hat{F} \colon \hat{\mathbb{T}} \to \mathbb{C}$ (re-using the same name) given by

$$\hat{F}(\chi_n) = \hat{F}(n).$$

1.3.14 Remark. It would seem better to not have to go to the parametrization $\zeta = e^{2\pi i x}$ of $\mathbb{T}$ in the formula (1.3.1) and instead to say that we are integrating $F(\zeta)\overline{\zeta^n}$ over $\mathbb{T}$, or integrating $F(\zeta)\overline{\chi_n(\zeta)}$ over $\mathbb{T}$ to get $\hat{F}(n)$. In fact the formula (1.3.1) is in a way more practical, but an integral over $\mathbb{T}$ seems more elegant.

We will soon discuss a quick fix remedy for this in an approach to integration on $\mathbb{T}$.

Before that, it might be instructive to compare Definition 1.3.13 with Definition 1.1.2.1. For $f \in CP[0,1]$ we wrote down a Fourier series (1.1.3) although we have yet to discuss any sense in which the series might be summed except when there are only finitely many nonzero terms (leading to a trigonometric polynomial as the sum). We could write down a series for $F \in C(\mathbb{T})$. It would be

$$\sum_{\chi \in \hat{\mathbb{T}}} \hat{F}(\chi)\chi = \sum_{n \in \mathbb{Z}} \hat{F}(\chi_n)\chi_n \tag{1.3.2}$$

So the terms $\hat{F}(\chi)\chi$ are multiples of functions (characters) $\chi \in C(\mathbb{T})$ by coefficients $\hat{F}(\chi)$. The function $\hat{F}(\chi)\chi$ evaluated at $\zeta \in \mathbb{T}$ gives $\hat{F}(\chi)\chi(\zeta)$. Again we postpone how this series might be summed and whether the sum is actually F.

The difference in Definition 1.3.13 is that we are heading for a group-theoretic approach, in the language of locally compact abelian groups G and their duals $\hat{G}$.

So $\mathbb{T}$ is a group but the unit interval $[0,1]$ is not. However we use $x \in [0,1]$ to parametrize $\mathbb{T}$ via $\zeta = e^{2\pi i x}$ and then the thing to note is that $x = 0$ and $x = 1$ both give $\zeta = 1$. So we lose the distinction between the endpoints of $[0,1]$, but in a way we already lost that when we restricted to periodic $f \in CP[0,1]$, those with $f(0) = f(1)$. We can make $[0,1)$ into a group either by addition modulo 1 (so the sum $x+y$ of $x, y \in [0,1)$ is defined as $x+y$ if $x+y < 1$ or $x+y-1$ if $x+y \geq 1$) or by multiplying

$$e^{2\pi i x}e^{2\pi i y} = e^{2\pi i(x+y)}$$

and expressing the result as $e^{2\pi i z}$ where $z \in [0,1)$.

Another way, essentially the same, is to consider the subgroup $\mathbb{Z}$ of the (additive) group $\mathbb{R}$ and to consider the quotient group $\mathbb{R}/\mathbb{Z}$. Elements of the quotient group are cosets $x + \mathbb{Z}$, and $x + \mathbb{Z} = x' + \mathbb{Z} \iff x - x' \in \mathbb{Z}$. So every coset $x + \mathbb{Z}$ can be expressed uniquely with $x \in [0,1)$ and then addition of cosets $(x + \mathbb{Z}) + (x' + \mathbb{Z}) = (x + x') + \mathbb{Z}$ comes down to addition modulo 1. The quotient group $\mathbb{R}/\mathbb{Z}$ is isomorphic to $\mathbb{T}$ (because the surjective group homomorphism $x \mapsto e^{2\pi i x}$ from $(\mathbb{R}, +)$ to $\mathbb{T}$ has kernel $\mathbb{Z}$). The periodic functions $f(x)$ on $\mathbb{R}$ with 1 as a period (those with $f(x+1) \equiv f(x)$) are exactly those that make sense on the quotient $\mathbb{R}/\mathbb{Z}$ via $f(x+\mathbb{Z}) = f(x)$. When we showed $\hat{\mathbb{T}} = \mathbb{Z}$, we basically showed that $\widehat{(\mathbb{R}/\mathbb{Z})}$ is the set of characters $x + \mathbb{Z} \mapsto e^{2\pi i n x}$ with $n \in \mathbb{Z}$ (and that multiplication of characters corresponds to addition of the n's).

1.3.15 Definition (Riemann integration on $\mathbb{T}$)**.** If $g \colon \mathbb{T} \to \mathbb{C}$ is continuous we can define the integral of g over $\mathbb{T}$ (with respect to normalized arc-length) to be

$$\int_{\mathbb{T}} g(\zeta)\,|d\zeta|/(2\pi) = \int_{\zeta \in \mathbb{T}} g(\zeta)\,|d\zeta|/(2\pi) \stackrel{\text{def}}{=} \int_0^1 g(e^{2\pi i x})\,dx$$

1.3.16 Proposition (Alternative formula for Fourier coefficients)**.** *If $F \in C(\mathbb{T})$, then*

$$\hat{F}(\chi) = \int_{\mathbb{T}} F(\zeta)\overline{\chi(\zeta)}\,|d\zeta|/(2\pi) \qquad (\chi \in \hat{\mathbb{T}}).$$

Proof. Combine the notation in Definition 1.3.15 with Definition 1.3.13. □

1.3.17 Remark (Measurable functions on $\mathbb{T}$). In order to generalize Definition 1.3.13 from continuous F to more general functions as we did when we included $f \in \mathcal{L}^1[0,1]$ in Definition 1.1.2.1, we should start with a measure on $\mathbb{T}$ and then define measurable functions on $\mathbb{T}$ and finally integrable functions.

A short-cut is to say that $g \colon \mathbb{T} \to \mathbb{C}$ is measurable if $f \colon [0,1] \to \mathbb{C}$ given by $f(x) = g(e^{2\pi i x})$ is Lebesgue measurable, that g is integrable if and only if f is and then that the integral of g is just the Lebesgue integral $\int_{[0,1]} f\,d\mu$.

The usual and more satisfactory (but in the end equivalent) approach is to set up a σ-algebra of subsets of $\mathbb{T}$ and a measure λ on the σ algebra that gives normalized arclength of subsets, then use the same sort of approach as used in MA2224 to define integrals over the real line. That would end up with $\int_{\mathbb{T}} g\,d\lambda$. It would be somehow better because it maps out how to do things when $\mathbb{T}$ is replaced by a general (compact) group. But in the end we would have the same value for the integral as with our short-cut. We will use the notation $\int_{\mathbb{T}} g\,d\lambda$ even though we have not really explained λ.

1.3.18 Definition. For $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$ we define

$$\mathcal{L}^1(\mathbb{T}) = \{g \colon \mathbb{T} \to \mathbb{K} : g \text{ measurable and } \int_{\mathbb{T}} |g| \, d\lambda < \infty\}.$$

This space is called the space of Lebesgue integrable functions (on $\mathbb{T}$) and λ refers to normalized arclength measure.

We also define the magnitude of $g \in \mathcal{L}^1(\mathbb{T})$ as

$$\|g\|_1 = \int_{\mathbb{T}} |g| \, d\lambda$$

1.3.19 Remark. We could define $\lambda(E)$ for $E \subseteq \mathbb{T}$ by saying

$$\lambda(E) = \mu(\{x \in [0,1] : e^{2\pi i x} \in E\})$$

and that this is defined only for those E where the subset of $[0,1]$ is a Lebesgue measurable subset. Again a kind of short-cut.

1.3.20 Definition (more general definition that Definition 1.3.13)**.** If $F \in \mathcal{L}^1(\mathbb{T})$ we define the *Fourier transform* of F to be the function $\hat{F} \colon \hat{\mathbb{T}} \to \mathbb{C}$ given by

$$\hat{F}(\chi) = \int_{\mathbb{T}} F \bar{\chi} \, d\lambda$$

(defined with the Lebesgue integral).

(The integrand is the function with value at $\zeta \in \mathbb{T}$ given by $F(\zeta)\overline{\chi(\zeta)}$.)

1.3.21 Remark. We could now state and prove a version of Lemma 1.1.5.3 in our new context, but in fact it would not really be a different result, just using different notation.

1.4 Finite abelian groups

In this section we discuss what the Fourier theory says for functions $f \colon G \to \mathbb{C}$ on a finite abelian group G.

We will write G additively, so $(G, +)$. As mentioned before (in Examples 1.3.6 (d)), finite groups can get the discrete topology and then the continuous functions $f \colon G \to \mathbb{C}$ are just all functions. If we list the elements $G = \{g_0 = 0, g_1, \ldots, g_{N-1}\}$ then a function $f \colon G \to \mathbb{C}$ is the same as an N-tuple $(f(g_0), f(g_1), \ldots, f(g_N)\}$ of values, or a point in $\mathbb{C}^N$ (with N the number of elements in G, usually referred to as the *order of* G).

The dual group $\hat{G}$ is the set of characters $\chi \colon G \to \mathbb{T}$, functions that satisfy $\chi(g + h) = \chi(g)\chi(h)$ for $g, h \in G$. We should say that these χ have to be continuous also, but all functions on G are continuous in this context, and so continuity is not a concern.

1.4.1 Examples (Simple examples of finite abelian groups).

(i) (Cyclic groups)

Written with multiplicative notation, the cyclic group of order N has one generator g with $g^N = 1$, and all the distinct elements of G are $1 = g^0, g, g^2, \ldots, g^{N-1}$. Multiplication $g^n g^m$ gives the exponent $n + m$ reduced modulo N.

If we write the operation additively, powers become multiples, that is g^n becomes ng.

We can take the cyclic group of order N to be $\mathbb{Z}/(N\mathbb{Z})$, and each coset of the subgroup $N\mathbb{Z}$ has a unique representative among $0, 1, \ldots, N-1$. We write $\mathbb{Z}_N$ for $\mathbb{Z}/(N\mathbb{Z})$ or $\{0, 1, \ldots, N-1\}$ with addition modulo N.

The characters are then the homomorphisms $\chi\colon \mathbb{Z}_N \to \mathbb{T}$. These are determined completely by the value $\zeta = \chi(1)$, because $\chi(0) = 1$, $\chi(2) = \chi(1+1) = \chi(1)\chi(1) = \zeta^2$ and in general $\chi(n) = \zeta^n$ for $0 \leq n < N$. But we have to have $\zeta^N = 1$, or ζ an N^{th} root of 1. The N^{th} roots of 1 (in $\mathbb{T}$) are of the form $e^{2\pi i k/N}$ for $k = 0, 1, \ldots, N-1$ and form a (multiplicative) cyclic group with generator $e^{2\pi i/N}$.

So in a way the dual group of $\mathbb{Z}_N$ is $\mathbb{Z}_N$ again.

(ii) (direct sums of cyclic groups)

If G_1 and G_2 are abelian groups (we write them additively now) then their direct sum $G_1 \oplus G_2$ is the Cartesian product $G_1 \times G_2$ with coordinatewise addition. That is the operation is given by

$$(g_1, g_2) + (h_1, h_2) = (g_1 + h_1, g_2 + h_2).$$

The identity element is $(0, 0)$.

If $\chi\colon G_1 \oplus G_2 \to \mathbb{T}$ is a character, then $\chi_1(g_1) = \chi(g_1, 0)$ gives a character of G_1 and $\chi_2(g_2) = \chi(0, g_2)$ gives a character of G_2. In fact we can recover χ from χ_1 and χ_2 because

$$\chi(g_1, g_2) = \chi((g_1, 0) + (0, g_2)) = \chi(g_1, 0)\chi(0, g_2) = \chi_1(g_1)\chi_2(g_2)$$

From this we can also see that if $\chi_1 \in \hat{G}_1$ and $\chi_2 \in \hat{G}_2$, then the above formula gives $\chi \in \widehat{G_1 \oplus G_2}$ and also that χ is uniquely determined by (χ_1, χ_2). It follows that we can identify

$$\widehat{G_1 \oplus G_2} \cong \hat{G}_1 \times \hat{G}_2$$

and we can also see that the group structure on $\widehat{G_1 \oplus G_2}$ corresponds to the direct product group operation on $\hat{G}_1 \times \hat{G}_2$, the one where

$$(\chi_1, \chi_2) \cdot (\chi_1', \chi_2') = (\chi_1\chi_1', \chi_2\chi_2')$$

This is actually the same as the direct sum of $\hat{G}_1$ and $\hat{G}_2$ but the direct sum notation is best used for groups written additively.

In fact the above two examples allow us to deal with all finite abelian groups, because of the following result which we quote without proof.

1.4.2 Theorem (Structure theorem for finite abelian groups)**.** *If $(G,+)$ is a finite abelian group, then there are cyclic subgroups $G_1, G_2, \ldots, G_k$ such that*

$$G = G_1 \oplus G_2 \oplus \cdots \oplus G_k$$

It is in fact possible to take the groups G_j to be each of prime power order and then the orders of the G_j are uniquely determined by G apart from the the possibility of permuting the summands.

1.4.3 Remark. There is a more general theorem for finitely generated abelian groups G, those such that there is a finite subset F of G such that no proper subgroup of G contains F. In that case the result allows for finitely many infinite cyclic summands, or finitely many copies of $\mathbb{Z}$, in addition to the finite cyclic summands as above. The direct sum of ℓ copies of $\mathbb{Z}$ is $\mathbb{Z}^\ell$.

The (nonzero) finitely generated abelian groups with no elements of finite order (apart from the identity) are those isomorphic to $\mathbb{Z}^\ell$ for some $\ell \in \mathbb{N}$.

[Recall the the *order of an element* of a group is the order of the subgroup generated by that element.]

1.4.4 Corollary. *If G is a finite abelian group, then $\hat{G}$ is a finite abelian group of the same order.*

Proof. We can use Theorem 1.4.2 together with Examples 1.4.1 to prove this by induction on the number of cyclic summands required to give G. □

1.4.5 Remark. For a finite abelian group G, we could use the notation $C(G)$ for the continuous functions $f\colon G \to \mathbb{C}$, but all functions are continuous in this context and so the $C(G)$ notation seems out of place, though it is analogous to $C(\mathbb{T})$. As noted at the start of this discussion, if G has order N and $f\colon G \to \mathbb{C}$, we can consider f as an N-tuple of complex numbers, with coordinates corresponding to the values of f.

Either way, these functions $f\colon G \to \mathbb{C}$ form a vector space over $\mathbb{C}$ of dimension N.

1.4.6 Lemma. *If G is a cyclic group of order N, then the characters $\chi \in \hat{G}$ are linearly independent functions on G.*

Proof. We can prove this using Vandermonde determinants in a way very similar to what we did in the proof of Example 1.1.4.9. However, we will also redo this shortly via inner products.

Say $G = \mathbb{Z}_N$ and then the dual group $\hat{G}$ is generated by the basic character χ_1 given by $\chi_1(n) = e^{2\pi i n/N}$. That is $\hat{G} = \{1, \chi_1, \chi_1^2, \ldots, \chi_1^{N-1}\}$ (see Examples 1.4.1). To show they are linearly independent as functions write each χ_1^j as an N-tuple

$$(\chi_1^j(0), \chi_1^j(1), \ldots, \chi_1^j(N-1)) = (1, (e^{2\pi i}/N)^j, \ldots, (e^{2\pi i(N-1)}/N)^j)$$

$(j = 0, 1, \ldots, N-1)$. Assembling these N vectors into a matrix, we get an $N \times N$ matrix and its determinant is a Vandermonde determinant, nonzero because the N numbers $(\chi_1(0), \chi_1(1), \ldots, \chi_1(N-1))$ are distinct. □

1.4.7 Remark. This means that every function $f\colon \mathbb{Z}_N \to \mathbb{C}$ can be expressed as a linear combination of the characters $\chi \in \widehat{\mathbb{Z}_N}$. This means a Fourier decomposition valid for all functions, and so it is somewhat more complete information than we had for Fourier series of $f \in CP[0,1]$ or $f \in C(\mathbb{T})$ (where we have no result beyond the case of trigonometric polynomials). However, in the context of $\mathbb{Z}_N$ we do not yet have a concrete formula for the coefficients of f in an expansion with characters.

We can get such a formula by using inner products. In the case of $C(\mathbb{T})$ we got the appropriate inner product by integration but in the case of a finite group this is replaced by summation.

1.4.8 Definition. If G is a finite abelian group, we define an inner product of functions $f, h\colon G \to \mathbb{C}$ as

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

1.4.9 Lemma. *Definition 1.4.8 gives an inner product on the space of $\mathbb{C}$-valued functions on a finite abelian group G.*

Proof. If we list the elements of G in some order $G = \{g_0, g_1, \ldots, g_{N-1}\}$, then the inner product we have defined corresponds to the usual inner product on $\mathbb{C}^n$ (applied to the N-tuples of values of the functions)

$$\langle f, h\rangle = \sum_{j=0}^{N-1} f(g_j)\overline{h(g_j)} = \langle (f(g_0), \ldots, f(g_{N-1})), (h(g_0), \ldots, h(g_{N-1}))\rangle$$

Hence we know it satisfies the conditions of Definition 1.1.4.3. □

1.4.10 Proposition. *If G is a finite abelian group of order N, then the characters of G form an orthogonal set of functions on G (in the inner product of Definition 1.4.8) and $\langle \chi, \chi\rangle = N$ for each $\chi \in \hat{G}$.*

Proof. We prove this first for the case of a cyclic group G, which we may take to be $G = \mathbb{Z}_N$. If $\chi, \chi' \in \hat{G}$, then we have χ and χ' given by two N^{th} roots of unity $\chi(1)$ and $\chi'(1)$. Thus

$$\langle \chi, \chi'\rangle = \sum_{j=0}^{N-1} \chi(j)\overline{\chi'(j)} = \sum_{j=0}^{N-1} \chi(1)^j\overline{\chi'(1)^j} = \sum_{j=0}^{N-1} (\chi(1)/\chi'(1))^j$$

Now $\chi(1)/\chi'(1)$ is also an N^{th} root of unity. If it is 1, then $\chi = \chi'$ and we get $\langle \chi, \chi'\rangle = \langle \chi, \chi\rangle = N$. If $\chi \neq \chi'$ we get

$$\langle \chi, \chi'\rangle = \frac{1 - (\chi(1)/\chi'(1))^N}{1 - \chi(1)/\chi'(1)} = 0$$

from the formula for the sum of a geometric series (or maybe you are familiar with this standard fact about sums of roots of unity). This completes the proof in the cyclic case.

For a general finite abelian group, we rely on the structure theorem (1.4.2) and induction on the number of cyclic summands. If G is not cyclic the theorem says we can write $G = G_1 \oplus G_2$

where (say) G_1 is cyclic and G_2 has (one) fewer cyclic summands than G. By the first part of the proof we have the result for G_1 and by the inductive hypothesis we have it for G_2. Let N_1 be the order of G_1, N_2 the order of G_2 and $\chi, \chi' \in \hat{G}$. So then χ has the form

$$\chi(g_1 + g_2) = \chi_1(g_1)\chi_2(g_2) \qquad (g_1 \in G_1, g_2 \in G_2)$$

for $\chi_1 \in \hat{G}_1$, $\chi_2 \in \hat{G}_2$. Similarly $\chi'(g_1 + g_2) = \chi_1'(g_1)\chi_2'(g_2)$. Then we have

$$\begin{aligned}\langle \chi, \chi' \rangle &= \sum_{g_1 \in G_1, g_2 \in G_2} \chi(g_1 + g_2)\overline{\chi'(g_1 + g_2)} \\ &= \sum_{g_1 \in G_1, g_2 \in G_2} \chi_1(g_1)\chi_2(g_2)\overline{\chi_1'(g_1)\chi_2'(g_2)} \\ &= \left(\sum_{g_1 \in G_1} \chi_1(g_1)\overline{\chi_2'(g_1)}\right)\left(\sum_{g_2 \in G_2} \chi_2(g_2)\overline{\chi_1'(g_2)}\right)\end{aligned}$$

If $\chi_1 \neq \chi_1'$ then the first of the latter two sums vanishes, while if $\chi_2 \neq \chi_2'$ we also get 0 by the inductive hypothesis. If, on the other hand, $\chi = \chi'$, then we get $N_1 N_2 = N$.

That completes the inductive step. □

1.4.11 Corollary (Fourier expansion for finite abelian groups)**.** *If G is a finite abelian group of order N, then*

$$\{\chi/\sqrt{N} : \chi \in \hat{G}\}$$

is an orthonormal basis for $C(G) = \{f : G \to \mathbb{C}\}$ (in the inner product of Definition 1.4.8).

It follows that each $f \in C(G)$ has a representation

$$f = \sum_{\chi \in \hat{G}} \langle f, \chi/\sqrt{N} \rangle \frac{\chi}{\sqrt{N}}$$

1.4.12 Definition. If G is a finite abelian group of order N and $f : G \to \mathbb{C}$, then the *Fourier transform* of f is the function $\hat{f} : \hat{G} \to \mathbb{C}$ given by

$$\hat{f}(\chi) = \langle f, \chi/\sqrt{N} \rangle = \frac{1}{\sqrt{N}} \sum_{g \in G} f(g)\overline{\chi(g)}$$

1.4.13 Corollary (Inverse Fourier transform formula)**.** *If G is a finite abelian group of order and $f : G \to \mathbb{C}$, then*

$$f = \frac{1}{\sqrt{N}} \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi$$

or

$$f(g) = \frac{1}{\sqrt{N}} \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi(g) \qquad (g \in G).$$

1.4.14 Remark. Note then that the inverse Fourier transform formula is similar to the Fourier transform, exchanging G and $\hat{G}$ and also omitting the complex conjugation.

Comparing the Fourier transform for finite abelian G with Definition 1.3.20 for $f \in \mathcal{L}^1(\mathbb{T})$ we note that on $\mathbb{T}$ we integrated with respect to a normalized measure on $\mathbb{T}$. In fact it is a rotation-invariant measure, or a measure λ on $\mathbb{T}$ invariant under translation by elements of $\mathbb{T}$ and with total mass $\lambda(\mathbb{T}) = 1$. (The invariance property is $\lambda(\zeta E) = \lambda(E)$ for $E \subseteq \mathbb{T}$ measurable, where $\zeta E = \{\zeta z : z \in E\}$ is the rotation of E by ζ.)

The natural way to get a measure on a finite set like G is to assign a mass $m_g \geq 0$ to each $g \in G$ and then to define the measure or total mass of a subset $E \subset G$ by $m(E) = \sum_{g \in E} m_g$. Matters of measurability are not normally relevant on finite sets and for any function $f \colon G \to \mathbb{C}$ we have an integral

$$\int_G f\, dm = \sum_{g \in G} f(g) m_g$$

which is actually a finite sum. If we developed an abstract approach to measures and integrals (including the Lebesgue integral as a special case), it would reduce for finite sets to the above kind of sum.

In the case of $\mathbb{T}$ we normalized so that $\lambda(\mathbb{T}) = 1$. For a finite group G, the translation-invariance requirement would imply that m_g is constant, independent of $g \in G$. Then there are two rather obvious choices, one to have $m_g = 1$ for each $g \in G$ (and then m would be called counting measure, $m(E) =$ the number of elements of E) and the other would be to insist $m(G) = 1$. The latter would force $m_g = 1/N$ with N the order of G.

The formulae we have do not use either of these. You could argue that we take the half-way house $m_g = 1/\sqrt{N}$ instead, so that $\hat{f}(\chi) = \int_G f \bar{\chi}\, dm$. The inverse formula (in Corollary 1.4.13), where we sum over $\chi \in \hat{G}$, could also be written as an integral over $\hat{G}$, but again with a measure that assigns mass $1/\sqrt{N}$ to each $\chi \in \hat{G}$.

1.5 The real line

We consider now the case $G = \mathbb{R}$. Again the idea is to write $f \colon \mathbb{R} \to \mathbb{C}$ in terms of characters $\chi \in \hat{\mathbb{R}}$, which are identified with points of $\mathbb{R}$ again by Proposition 1.3.11 (a). Recall $\chi_t(x) = e^{2\pi i x t}$ determines all characters as t varies over $t \in \mathbb{R}$.

There are issues here in identifying a class of f where we can take a transform and there is no obvious way in which the characters are orthogonal. However, the $\mathcal{L}^1$ approach works out fine.

1.5.1 Definition. For $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$ we define

$$\mathcal{L}^1(\mathbb{R}) = \{f \colon \mathbb{R} \to \mathbb{K} : f \text{ measurable and } \int_{\mathbb{R}} |f|\, d\mu < \infty\}.$$

This space is called the space of Lebesgue integrable functions (on $\mathbb{R}$) and μ refers to Lebesgue length measure on the real line.

On this space we define addition $f+g$ of functions $f, g \in \mathcal{L}^1(\mathbb{R})$ and scalar multiples λf ($\lambda \in \mathbb{K}$ and $f \in \mathcal{L}^1(\mathbb{R})$) in the same way as for previous $\mathcal{L}^1$ definitions, by $(f+g)(x) = f(x) + g(x)$ and $(\lambda f)(x) = \lambda f(x)$.

We also define the magnitude of $f \in \mathcal{L}^1(\mathbb{R})$ as

$$\|f\|_1 = \int_{\mathbb{R}} |f| \, d\mu$$

1.5.2 Definition. If $f \in \mathcal{L}^1(\mathbb{R})$ we define the *Fourier transform* of f to be the function $\hat{f} \colon \hat{\mathbb{R}} \to \mathbb{C}$ given by

$$\hat{f}(\chi) = \int_{\mathbb{R}} f \bar{\chi} \, d\mu$$

(defined with the Lebesgue integral).

(The integrand is the function with value at $x \in \mathbb{R}$ given by $f(x)\overline{\chi(x)}$.)

Identifying $\hat{\mathbb{R}}$ with $\mathbb{R}$, this corresponds to

$$\hat{f}(\chi_t) = \int_{\mathbb{R}} f(x)\overline{e^{2\pi ixt}} \, d\mu(x) = \int_{\mathbb{R}} f(x) e^{-2\pi ixt} \, d\mu(x)$$

and we may sometimes write $\hat{f}(t)$ in place of $\hat{f}(\chi_t)$.

1.5.3 Examples. Examples of $f \in \mathcal{L}^1(\mathbb{R})$ can be found by considering continuous $f \colon \mathbb{R} \to \mathbb{R}$ that tend to 0 fast enough as $x \to \pm\infty$. In particular, if there is $\alpha > 1$ such that

$$|f(x)| \leq \frac{C}{|x|^\alpha} \qquad (|x| \geq 1),$$

that is enough.

An even stronger condition is to assume that f is a *compactly supported* continuous function, where compactly supported means that

$$\{x \in \mathbb{R} : f(x) \neq 0\}$$

has compact closure in $\mathbb{R}$. Since the Heine Borel theorem says that the compact subsets of $\mathbb{R}$ are those that are closed and bounded, we could also simply require that $\{x \in \mathbb{R} : f(x) \neq 0\}$ is bounded, or contained in some finite interval $[a, b]$ $(-\infty < a \leq b < \infty)$. In that case $f \in \mathcal{L}^1(\mathbb{R})$ because

$$\int_{\mathbb{R}} |f| \, d\mu = \int_{[a,b]} |f| \, d\mu$$

is finite (as a Riemann integral).

The notation $C_c(\mathbb{R})$ is used for the space of compactly supported continuous functions on $\mathbb{R}$. We are saying $C_c(\mathbb{R}) \subset \mathcal{L}^1(\mathbb{R})$.

As in the case of $C(\mathbb{T})$ and $CP[0, 1]$, or their $\mathcal{L}^1$ relatives, we have not yet stated any result about recovering f from $\hat{f}$. So far we have such a result only for finite abelian G. And recall the constraints identified in Lemma 1.1.5.3 (iv).

We can't fall back on trigonometric polynomials here, because characters have constant absolute value 1 and are therefore not in $\mathcal{L}^1(\mathbb{R})$.

A Structure theorem for finite abelian groups

Here we indicate a proof (more or less an old one due apparently to Kronecker) of the bare bones of the theorem. We did not cover this in lectures and it won't be examined. (There could be a simpler way to express the argument.)

A.1 Theorem (Structure theorem for finite abelian groups, Theorem 1.4.2 simplified)**.** *If $(G, +)$ is a finite abelian group, then there are cyclic subgroups $G_1, G_2, \ldots, G_k$ such that*

$$G = G_1 \oplus G_2 \oplus \cdots \oplus G_k$$

More precisely, there are elements $g_1, g_2, \ldots, g_k \in G$ with g_j of order N_j (for $1 \leq j \leq k$), such that N_{j+1} divides N_j for $1 \leq j < k$, and each $g \in G$ can be written uniquely as

$$g = m_1 g_1 + m_2 g_2 + \cdots + m_k g_k \text{ with } 0 \leq m_j < N_j \text{ for } 1 \leq j \leq k$$

(and we take $G_j = \mathbb{Z} g_j$ for $1 \leq j \leq k$).

Proof. As usual for additive groups we write 0 for the identity element of G and define ng for $n \in \mathbb{Z}$ in the usual way ($0g = 0$, $1g = g$, $(n+1)g = ng + g$ for $n \geq 0$, $(-n)g = -(ng) =$ the additive inverse of ng). The *order* of $g \in G$ is the smallest $n \geq 0$ with $ng = 0$. In a general group there might be no such n but for G finite, every $g \in G$ has order at most $|G| =$ the order of G (the number of elements of G). The only $m \in \mathbb{N}$ with $mg = 0$ are then those where n divides m. Also we know n divides $|G|$ by Lagrange's theorem.

If $g \in G$, the subgroup generated by g is $\{0, g, 2g, \ldots\} = \{0, g, 2g, \ldots, (n_g - 1)g\}$ where n_g is the order of g. We write $\mathbb{Z}g$ for this subgroup. As G is abelian, all subgroups are normal subgroups and so the quotient group $G/\mathbb{Z}g$ makes sense.

If $g_1, g_2 \in G$, then G has an element of order the least common multiple of their orders. [Write n_i for the order of g_i ($i = 1, 2$). Let h be the highest common factor of n_1 and n_2. Then $g_3 = hg_2 \in G$ has order $n_3 = n_2/h$ and we can check that $g_1 + g_3$ has order $n_1 n_3 = n_1 n_2 / h = \operatorname{lcm}(n_1, n_2)$.

There is a largest possible order for an element of G. Say N_1 is that largest order and g_1 has order N_1. It follows from the remark in the previous paragraph that every $g \in G$ has order dividing N_1, or $N_1 g = 0 (\forall g \in G)$. Now consider $G' = G/\mathbb{Z}g_1$.

If G' has order 1, then $G = \mathbb{Z}g_1$ is cyclic and we are finished proving what we want. Otherwise we can set up an induction on the order of the group and assume we have already proved the theorem for groups of any order smaller than $|G|$, in particular for G'.

So we can find $g_2', \ldots, g_k' \in G/\mathbb{Z}g_1$ so that each g_j' has order N_j and every $g' \in G/\mathbb{Z}g_1$ can be expressed uniquely as

$$g' = m_2 g_2' + \cdots + m_k g_k'.$$

Also each N_{j+1} divides N_j for $2 \leq j < k$ by the inductive hypothesis (and divides N_1 too). Recall that elements of $G/\mathbb{Z}g_1$ are cosets $g + \mathbb{Z}g_1$. Express $g_j' = \tilde{g}_j + \mathbb{Z}g_1$ for $\tilde{g}_j \in G$ and so our conclusion is that for $g \in G$ there are unique $m_2, \ldots, m_k$ with $0 \leq m_j < N_j$ for $2 \leq j \leq k$ and

$$g + \mathbb{Z}g_1 = m_2(\tilde{g}_2 + \mathbb{Z}g_1) + \cdots + m_k(\tilde{g}_k + \mathbb{Z}g_1).$$

Since $N_j g_j' = 0$, it means that $N_j(\tilde{g}_j + \mathbb{Z}g_1) = 0 + \mathbb{Z}g_1$ or $N_j\tilde{g}_j \in \mathbb{Z}g_1$.

Now

$$\frac{N_1}{N_j} N_j \tilde{g}_j = N_1 g_j = 0$$

and so $N_j\tilde{g}_j \in \{0, N_j g_1, 2N_j g_1, \ldots, \left(\frac{N_1}{N_j} - 1\right) N_j g_1\} = \{x \in \mathbb{Z}g_1 : (N_1/N_j)x = 0\}$. It follows that there is $g_j \in \tilde{g}_j + \mathbb{Z}g_1$ with $N_j g_j = 0$. This g_j has $g_j' = g_j + \mathbb{Z}g_1$, and then we have that each $g' = g + \mathbb{Z}g_1 \in G/\mathbb{Z}g_1$ can be expressed uniquely as

$$g' = g + \mathbb{Z}g_1 = m_2(g_2 + \mathbb{Z}g_1) + \cdots + m_k(g_k + \mathbb{Z}g_1) = (m_2 g_2 + \cdots + m_k g_k) + \mathbb{Z}g_1.$$

For each $g \in G$ we can conclude that we can express

$$g = m_1 g_1 + m_2 g_2 + \cdots + m_k g_k$$

with $0 \leq m_j < N_j$ always. As the order of G is $|G| = N_1|G'| = N_1 N_2 \cdots N_k$, the numbers $m_1, m_2, \ldots, m_k$ have to be uniquely determined by g (as otherwise we would not have enough choices to represent each $g \in G$). □

17 April 2018: Fix typo in §1.1.3.
Richard M. Timoney (April 17, 2018)