Fourier Analysis

Tom Leinster

2013–14

Contents

Chapter A

Preparing the ground

A1 The algebraist's dream

For the lecture of 13 January 2014

The algebraist thinks: 'Analysis is hard. Can we reduce it to algebra?'

Idea Use power series.

1 Many functions $f : \mathbb{R} \to \mathbb{C}$ can be written as a power series,

$$f(x) = \sum_{n=0}^{\infty} c_n x^n. \tag{A:1}$$

If we can deal with the sequence (c_n) rather than the function f, everything will be much easier (and more algebraic).

2 *If* we can express f in this way, then the coefficients c_n must be given by

$$c_n = \frac{f^{(n)}(0)}{n!}.$$

(Proof: differentiate each side of (A:1) n times, then evaluate at 0.)

3 Problem: $\sum c_n x^n$ might not converge. Or, it might converge, but not to $f(x)$. E.g. put

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

(Figure A.1). It can be shown that $f^{(n)}(0)/n! = 0$ for all n (which you might guess from the flatness of the function near 0). But of course $f(x) \neq \sum_{n=0}^{\infty} 0 \cdot x^n$ except at $x = 0$. This brings to light:

4 Another problem: different infinitely differentiable functions can have the same power series. (E.g. this is the case for the f just defined and the constant function 0.)

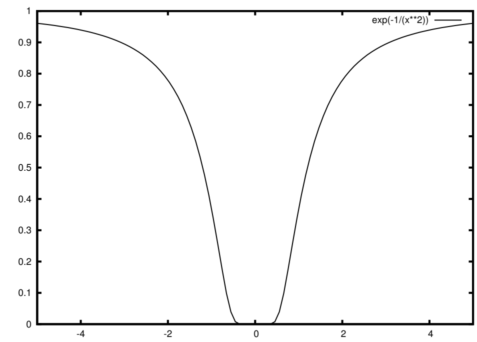


Figure A.1: The function $x \mapsto e^{-1/x^2}$

5 Yet another problem: the power series method only captures infinitely differentiable functions. (For recall that every power series is infinitely differentiable inside its disk of convergence.) However, many interesting functions are not infinitely differentiable.

6 In this power series approach to functions, the number 0 is given a special role. This isn't necessarily good or bad, but might seem a little suspicious. We could equally well consider power series $\sum c_n(x-a)^n$ centred at a, for any other value of a.

7 Also suspicious: the power series $\sum \frac{f^{(n)}(0)}{n!}x^n$ of f depends only on the values of f near 0. In the jargon, it's *locally determined.* You can't reasonably expect to predict the value of f for large $|x|$ given only the value of f for small $|x|$.

In other words, if you had two functions f and g that were the same throughout the interval $(-\delta, \delta)$, then they'd have the same power series. However, they might have very different values outside that interval.

So for various reasons, the algebraist's dream can't be realized using power series. However...

Alternative idea Use Fourier series.

1 Many 1-periodic functions $f: \mathbb{R} \to \mathbb{C}$ can be written as a Fourier series:

$$f(x) = \sum_{n=0}^{\infty} a_n \cos 2\pi nx + \sum_{n=0}^{\infty} b_n \sin 2\pi nx$$

$(a_n, b_n \in \mathbb{C})$. **1-periodic** means that $f(x+1) = f(x)$ for all $x \in \mathbb{R}$.

(Clearly any function expressible in this way *is* 1-periodic. You might be more used to dealing with 2π-periodic functions and Fourier series involving the terms $\cos nx$ and $\sin nx$, but the difference is purely cosmetic.)

Since

$$\cos\theta = (e^{i\theta} + e^{-i\theta})/2, \qquad \sin\theta = (e^{i\theta} - e^{-i\theta})/2i,$$

we can rewrite this more efficiently as

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi i k x}$$

for certain $c_k \in \mathbb{C}$. (Note that the sum starts at $-\infty$.)

2 *If* we can express f in this way, then the coefficients c_k must be given by

$$c_k = \int_0^1 f(x) e^{-2\pi i k x}\, dx.$$

(We'll see why later, or you can try proving it now.)

3 Problem: $\sum c_k e^{2\pi i k x}$ might not converge. Or, it might converge, but not to $f(x)$. (There are examples of both these phenomena.)

4 *But* unlike for power series, different continuous functions always have different Fourier series.

5 *And* unlike for power series, functions of many kinds can be captured using Fourier series. Far from having to be infinitely differentiable, even some *discontinuous* functions can be captured.

6 *And* again unlike for power series, in the definition of c_k, no point in $\mathbb{R}$ is given a special role.

7 *And* once more unlike for power series, the Fourier series $\sum c_k e^{2\pi i k x}$ of f depends on *all* of f (since $c_k = \int_0^1 f(x) e^{-2\pi i k x}\, dx$, and the periodic function f is determined by its restriction to $[0, 1]$). In the jargon, it's *globally determined*. So it's not unreasonable to expect to be able to predict the value of f for all x given only the Fourier coefficients c_k. We'll see that we often can.

The algebraist's dream can't be fully realized. But Fourier series come closer to realizing it than power series do, at least for periodic functions on $\mathbb{R}$. They're not as obvious an idea as power series, but in many ways they work better.

We'll spend a lot of the course exploring the extent to which functions can be understood in terms of their Fourier series. There are many analytic subtleties, which we'll have to think hard about.

The development of Fourier theory has been very important historically. It has been the spur for a lot of important ideas in mathematics, not all obviously connected to Fourier analysis. We'll meet some along the way.

Figure A.2: Excerpt from the index of Tom Körner's book *Fourier Analysis*

A2 Pseudo-historical overview

For the lecture of 16 January 2014

Most mathematicians are terrible historians. They can't resist recounting what *should* have happened, not what *did* happen. I can't claim to be any better—hence the 'pseudo' of the title.

This lecture is all about the 'excessive optimism' and 'excessive pessimism' mentioned in the index of Körner's book (Figure A.2).

A 1-periodic function on $\mathbb{R}$ is determined by its values on $[0, 1)$ (or any other interval of length 1). We say that a 1-periodic function is **integrable** if its restriction to $[0, 1)$ is integrable in the usual sense. Let $f : \mathbb{R} \to \mathbb{C}$ be an integrable 1-periodic function.

For $k \in \mathbb{Z}$, the kth **Fourier coefficient** of f is

$$\hat{f}(k) = \int_0^1 f(x) e^{-2\pi i k x} \, dx.$$

The **Fourier series** of f is the doubly infinite series Sf given by

$$(Sf)(x) = \sum_{k=-\infty}^{\infty} \hat{f}(k) e^{2\pi i k x}.$$

It's not entirely clear what kind of thing Sf is. For instance, it might not always converge, and it's not even obvious what convergence should *mean*. But the following definition is perfectly safe: for $n \geq 0$, the nth **Fourier partial sum** is the function $S_n f$ given by

$$(S_n f)(x) = \sum_{k=-n}^{n} \hat{f}(k) e^{2\pi i k x}.$$

Classical question: for 'nice' functions f, is it true that

$$(S_n f)(x) \to f(x) \text{ as } n \to \infty$$

for all $x \in [0, 1)$? For 'most' $x \in [0, 1)$? For *some* $x \in [0, 1)$?

This is the question of *pointwise convergence*. There are other kinds of convergence, as we'll see, some of which are important in ways that pointwise convergence is not.

Fourier himself (1768–1830) thought that yes, it's always true. But he was imprecise about this (and most other things). It's not even clear what he would have taken the word 'function' to mean.

To start with a point that was certainly clear to Fourier himself, it can't be true for *all* x and *all* f. For instance, suppose it's true for all x for some particular f. Define g to be the same as f except at one point of $[0, 1)$, where it takes some different value. Then $\hat{g}(k) = \hat{f}(k)$ for all k, so $S_n g = S_n f$ for all n, so it can't also be true that $(S_n g)(x) \to g(x)$ as $n \to \infty$ for all x.

(For instance, take g to be the 1-periodic function given by $g(x) = 1$ for $x \in \mathbb{Z}$ and $g(x) = 0$ for $x \notin \mathbb{Z}$. Then for all n, the function $S_n g$ has constant value 0, so $(S_n g)(x) \not\to g(x)$ whenever $x \in \mathbb{Z}$.)

Backing up Fourier's intuition, Dirichlet proved:

Theorem A2.1 (Dirichlet, 1829) *Let $f\colon \mathbb{R} \to \mathbb{C}$ be a 1-periodic, continuously differentiable function. Then $(S_n f)(x) \to f(x)$ as $n \to \infty$ for all $x \in [0, 1)$ (or equivalently, for all $x \in \mathbb{R}$).*

Once Dirichlet had proved that, it was generally believed that 'infinitely differentiable' could be relaxed to 'continuous'—that it was just a matter of time before someone figured out a proof that would work in this wider generality.

Most of the most prominent mathematicians of the day believed this. Dirichlet believed it, Cauchy, believed it, and Riemann, Weierstrass, Dedekind and Poisson all believed it. Cauchy even claimed he'd proved it. (Standards of rigour were lower in those days.) Dirichlet promised he'd prove it, but never did.

But then came a bombshell.

Theorem A2.2 (du Bois-Reymond, 1876) *There is a 1-periodic, continuous function $f\colon \mathbb{R} \to \mathbb{C}$ such that for some $x \in [0, 1)$, the sequence $\big((S_n f)(x)\big)_{n=0}^{\infty}$ fails to converge.*

Using this example, it takes comparatively little effort to construct a 1-periodic continuous function such that $((S_n f)(x))_{n=0}^{\infty}$ fails to converge at 10 points, or 100, or 1000:

Theorem A2.3 *Let $E \subseteq [0, 1)$ be a finite set. Then there is a 1-periodic, continuous function $f\colon \mathbb{R} \to \mathbb{C}$ such that for all $x \in E$, the sequence $\big((S_n f)(x)\big)_{n=0}^{\infty}$ fails to converge.*

Widespread pessimism ensued. The pendulum swung from the general consensus that Fourier series behave perfectly for all continuous functions to the opposite extreme. After du Bois-Reymond's theorem appeared, it came to be believed that there was some 1-periodic continuous function $f\colon \mathbb{R} \to \mathbb{C}$ with the property that for *all* $x \in [0, 1)$, the sequence $\big((S_n f)(x)\big)_{n=0}^{\infty}$ fails to converge.

This pessimism was reinforced by a further example:

Theorem A2.4 (Kolmogorov, 1926) *There exists a 1-periodic, Lebesgue integrable function $f\colon \mathbb{R} \to \mathbb{C}$ such that for all $x \in [0, 1)$, the sequence $\big((S_n f)(x)\big)_{n=0}^{\infty}$ fails to converge.*

We won't be doing Lebesgue integrability in this course, and I won't assume you know anything about it. But as general cultural background:

Remark A2.5 Every Riemann integrable function is Lebesgue integrable. (In particular, a Lebesgue integrable function need not be continuous, and Kolmogorov's example certainly wasn't.) Moreover, the Lebesgue integral of a Riemann integrable function is equal to its Riemann integral. However, there are functions that are Lebesgue integrable but not Riemann integrable. So, the theory of Lebesgue integration extends, and is more powerful than, the Riemann theory.

This situation persisted until relatively recently. For instance, one of the books on my own undergraduate reading list was Apostol's 1957 text *Mathematical Analysis*. In most respects, it is like today's textbooks, but he states in it that it's still unknown whether for a continuous function, the Fourier series has to converge at even *one* point.

The turning point came in the 1960s.

Theorem A2.6 (Carleson, 1964) *Let $f\colon \mathbb{R} \to \mathbb{C}$ be a 1-periodic function.*

- *If f is continuous then $(S_n f)(x) \to f(x)$ as $n \to \infty$ for at least one $x \in [0, 1)$.*
- *Better still, if f is Riemann integrable then $(S_n f)(x) \to f(x)$ as $n \to \infty$ for at least one $x \in [0, 1)$.*
- *Better* still*, if f is Riemann integrable then $(S_n f)(x) \to f(x)$ as $n \to \infty$ for almost all $x \in [0, 1)$. ('Almost all' will be defined later.)*
- Even *better still, if $|f|^2$ is Lebesgue integrable (which is true if f is Riemann integrable) then $(S_n f)(x) \to f(x)$ as $n \to \infty$ for almost all $x \in [0, 1)$.*

All known proofs of Carleson's theorem are hard, still far too hard for a course such as this.

(I say 'still' because most hard proofs of important facts are simplified over time. This was a big theorem that attracted a lot of attention, so lots of people have put in lots of work to simplify the proof. They *have* simplified it, but it's still well beyond our reach. This is true even for the very watered-down version of Carleson's theorem in the first bullet point: that for merely *continuous* functions, there is merely *one* point at which the Fourier series behaves well.)

Carleson's theorem is best possible, in a sense that can be made precise (Theorem A3.18). Roughly, this means that you can't do any better than 'almost all'. It almost completely answers the question of pointwise convergence of Fourier series. But there are other types of convergence too, and the behaviour of Fourier series from those points of view is interesting too, and has provoked a lot of mathematical developments—as we shall see.

A3 Integration

For the lecture of 20 January 2014

We won't be doing any *Fourier* analysis for the first couple of weeks. The course pulls together several strands of analysis, and we're going to look at them one at a time before attempting to bring them together. This has the disadvantage of postponing the moment when we first see any Fourier theory, but the advantage that we can concentrate on one thing at a time.

For the rest of this section, let $I \subseteq \mathbb{R}$ be a bounded interval.

Definition A3.1 i. A function $f : I \to \mathbb{R}$ is **integrable** if it is bounded and Riemann integrable.

ii. Let $f : I \to \mathbb{C}$ be a function. Writing $f_1, f_2 : I \to \mathbb{R}$ for its real and imaginary parts (so that

$$f(x) = f_1(x) + i f_2(x)$$

for all $x \in \mathbb{R}$), we say that f is **integrable** if f_1 and f_2 are both integrable. In that case, we define the **integral** of f by

$$\int_I f(x)\,dx = \int_I f_1(x)\,dx + i \int_I f_2(x)\,dx.$$

Part (i) is a declaration about how the word 'integrable' will be used *in this course*. Other people use it differently. For example, a more permissive meaning (which you don't have to know anything about) would be 'Lebesgue integrable'.

Examples A3.2 i. Bounded continuous functions are integrable. (If I is closed then 'bounded' follows automatically from 'continuous'.)

ii. Let $J \subseteq I$ be an interval. The **characteristic function** (or **indicator function**) of J is the function

$$\chi_J : I \to \mathbb{C}$$

defined by

$$\chi_J(x) = \begin{cases} 1 & \text{if } x \in J, \\ 0 & \text{if } x \notin J. \end{cases}$$

Then χ_J is integrable, and $\int_I \chi_J(x)\,dx = |J|$. Here $|J|$ is the **length** of the interval J, defined by $|J| = \sup J - \inf J$ (or as 0 if $J = \emptyset$). Concretely, if J is $[a, b]$ or $[a, b)$ or $(a, b]$ or (a, b) then $|J| = b - a$.

Notation A3.3 Let $f, g : I \to \mathbb{C}$. We write $f + g : I \to \mathbb{C}$ for the function defined by

$$(f + g)(x) = f(x) + g(x)$$

($x \in I$). We define $f \cdot g$ similarly, $\bar{f}$ by $\bar{f}(x) = \overline{f(x)}$ (where the bar means complex conjugate), and $|f|$ by $|f|(x) = |f(x)|$. Given $c \in \mathbb{C}$, we write $c : I \to \mathbb{C}$ for the function with constant value c. So, for instance, $c \cdot f$ (or cf) is the function $I \to \mathbb{C}$ given by $(cf)(x) = c \cdot f(x)$.

Lemma A3.4 *i. If $f, g\colon I \to \mathbb{C}$ are integrable then so is $f+g$, with*

$$\int_I (f+g)(x)\,dx = \int_I f(x)\,dx + \int_I g(x)\,dx.$$

ii. If $f\colon I \to \mathbb{C}$ is integrable and $c \in \mathbb{C}$ then cf is integrable, with

$$\int_I (cf)(x)\,dx = c\int_I f(x)\,dx.$$

iii. If $f\colon I \to \mathbb{C}$ is integrable then so is $\bar{f}$, with

$$\int_I \bar{f}(x)\,dx = \overline{\int_I f(x)\,dx}.$$

Proof Parts (i) and (ii) follow from the corresponding properties of real-valued integration, and part (iii) follows directly from the definitions. □

I'll occasionally state 'Facts', without proof. Some of these facts were proved in previous courses (such as PAA). Others weren't, and I'll be asking you to take those on trust.

Fact A3.5 *Let $f\colon I \to \mathbb{C}$ be an integrable function and $\phi\colon \mathbb{C} \to \mathbb{C}$ a continuous function. Then the composite $\phi \circ f\colon I \to \mathbb{C}$ is integrable.*

Example A3.6 If f is integrable then so is $|f|^p$, for any real $p > 0$. (Take $\phi(z) = |z|^p$.)

Lemma A3.7 *If $f, g\colon I \to \mathbb{C}$ are integrable then so is $f \cdot g\colon I \to \mathbb{C}$.*

Proof See Sheet 1, q.3. □

(If you've come across Lebesgue integration, you may be aware that both this lemma and Example A3.6 fail for Lebesgue integration. This is one respect in which the Riemann theory makes life simpler.)

Fact A3.8 *If $f, g\colon I \to \mathbb{R}$ are integrable and $f(x) \leq g(x)$ for all $x \in I$ then $\int_I f(x)\,dx \leq \int_I g(x)\,dx$.*

Note that this is for functions into $\mathbb{R}$, not $\mathbb{C}$; inequalities makes no sense in $\mathbb{C}$.

Lemma A3.9 (Triangle inequality for integration) *If $f\colon I \to \mathbb{C}$ is integrable then $|f|$ is integrable, with*

$$\left|\int_I f(x)\,dx\right| \leq \int_I |f|(x)\,dx.$$

Proof That $|f|$ is integrable is a special case of Example A3.6. Since $\int_I f(x)\,dx$ is a complex number, we may write

$$\int_I f(x)\,dx = u\left|\int_I f(x)\,dx\right|$$

for some $u \in \mathbb{C}$ with $|u| = 1$. Now

$$\begin{aligned}
\left| \int_I f(x)\,dx \right| &= \mathrm{Re} \left| \int_I f(x)\,dx \right| = \mathrm{Re}\Big(\frac{1}{u} \int_I f(x)\,dx\Big) = \mathrm{Re} \int_I \frac{1}{u} f(x)\,dx \\
&= \int_I \mathrm{Re}\Big(\frac{1}{u} f(x)\Big)\,dx \qquad \text{(A:2)} \\
&\leq \int_I \Big| \frac{1}{u} f(x) \Big|\,dx \qquad \text{(A:3)} \\
&= \int_I |f(x)|\,dx = \int_I |f|(x)\,dx.
\end{aligned}$$

Here $\mathrm{Re}(z)$ denotes the real part of a complex number z, equation (A:2) follows from the definition of complex-valued integration (Definition A3.1(ii)), the inequality (A:3) follows from Fact A3.8 and the fact that $\mathrm{Re}(z) \leq |z|$ for all $z \in \mathbb{C}$, and the rest follow either from earlier lemmas or directly from the definitions.□

We now come to a very important concept. It was introduced by Lebesgue, and forms part of his theory of integration; and although we are not studying that theory, we will need this one particular concept.

Definition A3.10 A subset E of $\mathbb{R}$ has **measure zero** (or is **null**) if for all $\varepsilon > 0$, there exists a sequence of intervals $(J_n)_{n=1}^{\infty}$ in $\mathbb{R}$ such that

$$E \subseteq \bigcup_{n=1}^{\infty} J_n \qquad \text{and} \qquad \sum_{n=1}^{\infty} |J_n| \leq \varepsilon.$$

Examples A3.11

i. Any countable set $E \subseteq \mathbb{R}$ has measure zero. This is clear for $E = \emptyset$, so assume $E \neq \emptyset$. We can choose a sequence $(x_n)_{n=1}^{\infty}$ such that $E = \{x_1, x_2, \ldots\}$. Let $\varepsilon > 0$. Put $J_n = \big(x_n - 2^{-(n+1)}\varepsilon, x_n + 2^{-(n+1)}\varepsilon\big)$. Then $E \subseteq \bigcup E_n$ and $\sum |J_n| = \sum_{n=1}^{\infty} 2^{-n}\varepsilon = \varepsilon$.

ii. There also exist uncountable sets of measure zero (such as the Cantor set).

iii. If E has measure zero and $F \subseteq E$ then F has measure zero.

iv. Let J be an interval with $|J| > 0$. (This condition just means that J is not empty or of the form $\{a\}$.) Then J does *not* have measure zero. This is not as obvious as it sounds. If you do think it's obvious, try proving it!

Definition A3.12 Let $S \subseteq \mathbb{R}$. A property P of points $x \in S$ holds **almost everywhere (a.e.)**, or **for almost all** $x \in S$, if there is some $E \subseteq S$ of measure zero such that P holds for all $x \in S \setminus E$.

For example, if $f, g \colon I \to \mathbb{C}$ then

'$f = g$ a.e.'

and

'$f(x) = g(x)$ for almost all x'

both mean: there exists a subset $E \subseteq I$ of measure zero such that $f(x) = g(x)$ for all $x \in I \setminus E$.

What's this got to do with integration?

Fact A3.13 *If $f, g\colon I \to \mathbb{C}$ are integrable and $f = g$ a.e. then $\int_I f(x)\,dx = \int_I g(x)\,dx$.*

Fact A3.14 *If $h\colon I \to \mathbb{R}$ is integrable, $h(x) \geq 0$ for all $x \in I$, and $\int_I h(x)\,dx = 0$, then $h = 0$ a.e..*

Proposition A3.15 *Let $f, g\colon I \to \mathbb{C}$ be integrable functions with $f = g$ a.e.. Assume that $|I| > 0$.*

i. Let $x \in I$. If f and g are both continuous at x, then $f(x) = g(x)$.

ii. If f and g are continuous then $f = g$.

Proof For (i), suppose for a contradiction that $f(x) \neq g(x)$. By continuity, we can find some $\delta > 0$ such that $f(t) \neq g(t)$ for all $t \in (x - \delta, x + \delta) \cap I$. Also, since $f = g$ a.e., we can choose $E \subseteq I$ of measure zero such that $f(t) = g(t)$ for all $t \in I \setminus E$. Then $(x - \delta, x + \delta) \cap I \subseteq E$, so $(x - \delta, x + \delta) \cap I$ has measure zero by Example A3.11(iii). But $|I| > 0$, so $(x - \delta, x + \delta) \cap I$ is an interval of length > 0, contradicting Example A3.11(iv).

Part (ii) follows immediately. □

The next result makes no mention of the concept of measure zero.

Proposition A3.16 *Let $h\colon I \to \mathbb{R}$ be an integrable function with $h(t) \geq 0$ for all $t \in I$ and $\int_I h(t)\,dt = 0$. Assume that $|I| > 0$.*

i. Let $x \in I$. If h is continuous at x, then $h(x) = 0$.

ii. If h is continuous then $h = 0$.

Proof We *could* deduce this from Fact A3.14 and Proposition A3.15. This is unnecessarily complicated, though. Sheet 1, q.4 asks you to find a direct proof. □

* * *

The last two results of this lecture are non-examinable, and are included just for background.

Fact A3.17 *A function $f\colon I \to \mathbb{C}$ is integrable if and only if it is bounded and continuous a.e. (that is, the set $\{x \in I : f \text{ is not continuous at } x\}$ has measure zero).*

For example, $\chi_{\mathbb{Q}\cap[0,1]}\colon [0, 1] \to \mathbb{C}$ is not integrable, since the set of discontinuities is $[0, 1]$, which does not have measure zero.

At the end of the last lecture, I said that in a certain sense, Carleson's theorem (Theorem A2.6) cannot be improved. Here is the precise statement.

Theorem A3.18 (Kahane and Katznelson, 1960s) *Let $E \subseteq [0, 1]$ be a set of measure zero. Then there is a 1-periodic, continuous function $f\colon \mathbb{R} \to \mathbb{C}$ such that for all $x \in E$, the sequence $\big((S_n f)(x)\big)_{n=0}^{\infty}$ fails to converge.*

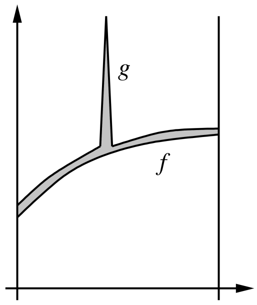


Figure A.3: The area between f and g is small, but the largest difference between them is large.

A4 What does it mean for a sequence of functions to converge?

For the lecture of 23 January 2014

There is no single right answer to this question. Compare the question 'how big is a person?' I might interpret that as a question about height, but you might interpret it as a question about weight. Neither of us would be right or wrong. Height and weight are correlated, but not logically related in an absolute sense.

Related to the question of the title is another question: what does it mean for two functions to be 'close'?

We might say that two functions f and g are close if the area between their graphs is small (that is, $\int |f(x) - g(x)|\, dx$ is small). Or, we might say that they are close if their values are never too different (that is, $\sup_x |f(x) - g(x)|$ is small). Both ideas are sensible, but they are not the same, as Figure A.3 demonstrates.

For the rest of this section, let $I \subseteq \mathbb{R}$ be a bounded interval of length > 0.

Definition A4.1 Let $f\colon I \to \mathbb{C}$ be an integrable function. Define:

- $\|f\|_1 = \int_I |f(x)|\, dx$, the **1-norm** of f;
- $\|f\|_2 = \sqrt{\int_I |f(x)|^2\, dx}$, the **2-norm** of f;
- $\|f\|_\infty = \sup\limits_{x \in I} |f(x)|$, the **$\infty$-norm** or **sup-norm** of f.

Remarks A4.2 i. In the definition of $\|f\|_2$, it really makes a difference that we write $|f(x)|^2$, not $f(x)^2$, since f is a *complex*-valued function.

ii. The definition of $\|f\|_\infty$ is only really appropriate for *continuous* functions f. We will mostly stick to continuous functions when we speak of $\|f\|_\infty$.

iii. The word 'norm' is actually not quite right, for reasons that will soon be explained.

iv. For $1 \leq p < \infty$, put $\|f\|_p = \left(\int_I |f(x)|^p\, dx\right)^{1/p}$. It can be shown that $\lim\limits_{p \to \infty} \|f\|_p = \|f\|_\infty$ for continuous f, which explains why it's called $\|\cdot\|_\infty$.

If we want to refer to the 1-norm in the abstract, we sometimes write it as $\|\cdot\|_1$. The dot is a blank or placeholder, into which arguments can be inserted. The same goes for $\|\cdot\|_2$ and $\|\cdot\|_\infty$.

Lemma A4.3 *Let $\|\cdot\|$ stand for any of $\|\cdot\|_1$, $\|\cdot\|_2$ or $\|\cdot\|_\infty$. Then for all integrable functions $f, g\colon I \to \mathbb{C}$ and all $c \in \mathbb{C}$:*

i. $\|f\| \geq 0$;

ii. $\|cf\| = |c| \cdot \|f\|$;

iii. $\|f + g\| \leq \|f\| + \|g\|$.

Proof All easy except (iii) for $\|\cdot\|_2$, which is Sheet 1, q.6(iv). □

A **norm** on the set {integrable functions $I \to \mathbb{C}$} is an operation satisfying conditions (i)–(iii) and one further condition: that if $\|f\| = 0$ then $f = 0$. This further condition fails for $\|\cdot\|_1$ and $\|\cdot\|_2$ (as the next example shows), so strictly speaking we should call them 'seminorms'. But I will abuse terminology and go on calling them the 1-norm and 2-norm. (The ∞-norm really is a norm.)

Example A4.4 Define $f\colon [0, 1) \to \mathbb{C}$ by

$$f(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\|f\|_1 = \|f\|_2 = 0$ but $f \neq 0$.

Lemma A4.5 *i. For an integrable function $f\colon I \to \mathbb{C}$,*

$$\|f\|_1 = 0 \iff f = 0 \textit{ a.e.} \iff \|f\|_2 = 0.$$

ii. For an integrable function $f\colon I \to \mathbb{C}$, if $\|f\|_1 = 0$ or $\|f\|_2 = 0$ then $f(x) = 0$ for all $x \in I$ such that f is continuous at x.

iii. For a continuous function $f\colon I \to \mathbb{C}$,

$$\|f\|_1 = 0 \iff f = 0 \iff \|f\|_2 = 0.$$

iv. For an integrable (or just bounded) function $f\colon I \to \mathbb{C}$,

$$\|f\|_\infty = 0 \iff f = 0.$$

Proof i. Follows from Facts A3.13 and A3.14.

ii. Follows from Proposition A3.16(i) (taking $h = |f|$ or $h = |f|^2$).

iii. Follows from (ii).

iv. Follows from the definition. □

We've seen that $\|f\|_1$, $\|f\|_2$ and $\|f\|_\infty$ are three different measures of the *size* of a function f.

So, $\|f - g\|_1$, $\|f - g\|_2$ and $\|f - g\|_\infty$ are three different measures of the *distance* between functions f and g.

They genuinely measure different things! Look back at the opening paragraphs of this lecture and Figure A.3.

We now consider the three resulting notions of convergence, plus two more.

Definition A4.6 Let (f_n) be a sequence of integrable functions from I to $\mathbb{C}$, and let f be an integrable function from I to $\mathbb{C}$. We say:

- $f_n \to f$ **in** $\|\cdot\|_1$ if $\|f_n - f\|_1 \to 0$ as $n \to \infty$;
- $f_n \to f$ **in** $\|\cdot\|_2$ (or **in mean square**) if $\|f_n - f\|_2 \to 0$ as $n \to \infty$;
- $f_n \to f$ **in** $\|\cdot\|_\infty$ (or **uniformly**) if $\|f_n - f\|_\infty \to 0$ as $n \to \infty$;
- $f_n \to f$ **pointwise** if for all $x \in I$, $f_n(x) \to f(x)$ as $n \to \infty$;
- $f_n \to f$ **a.e.** if for almost all $x \in I$, $f_n(x) \to f(x)$ as $n \to \infty$.

All five types of convergence are useful and meaningful.

We now work out the logical relationships between them. First, here's something in common between the first three.

Lemma A4.7 *Let $\|\cdot\|$ stand for $\|\cdot\|_1$, $\|\cdot\|_2$ or $\|\cdot\|_\infty$. If $f_n \to f$ in $\|\cdot\|$ then $\|f_n\| \to \|f\|$ as $n \to \infty$.*

Proof For all n, we have

$$\|f_n\| - \|f\| \leq \|f_n - f\|$$

by Lemma A4.3(iii), and similarly

$$\|f\| - \|f_n\| \leq \|f - f_n\| = \|f_n - f\|,$$

so

$$\big| \|f_n\| - \|f\| \big| \leq \|f_n - f\|.$$

But $\|f_n - f\| \to 0$ as $n \to \infty$, so $\|f_n\| - \|f\| \to 0$ as $n \to \infty$, as required. □

Here's a classic fact from PAA:

Fact A4.8 *Let $f, f_1, f_2, \ldots : I \to \mathbb{C}$ be functions such that $f_n \to f$ uniformly. Suppose that each f_n is continuous. Then f is continuous, and $\int_I f_n(x)\,dx \to \int_I f(x)\,dx$ as $n \to \infty$.*

In words: a uniform limit of continuous functions is continuous, and the integral of a uniform limit is the limit of the integrals.

Now let's begin to record the implications between the different types of convergence.

Lemma A4.9 *Let $f, f_1, f_2, \ldots : I \to \mathbb{C}$ be integrable functions. Then*

$$\begin{aligned} & f_n \to f \text{ uniformly} \\ \Rightarrow\ & f_n \to f \text{ pointwise} \\ \Rightarrow\ & f_n \to f \text{ a.e..} \end{aligned}$$

Proof Immediate from the definitions. □

We now restrict ourselves to the case $I = [0, 1)$, since that's what we'll need once we start to consider 1-periodic functions.

Lemma A4.10 *i. Let $f\colon [0,1) \to \mathbb{C}$ be an integrable function. Then*

$$\|f\|_1 \leq \|f\|_2 \leq \|f\|_\infty.$$

ii. Let $f, f_1, f_2, \ldots : [0,1) \to \mathbb{C}$ be integrable functions. Then

$$\begin{aligned} & f_n \to f \textit{ in } \|\cdot\|_\infty \\ \Rightarrow\ & f_n \to f \textit{ in } \|\cdot\|_2 \\ \Rightarrow\ & f_n \to f \textit{ in } \|\cdot\|_1. \end{aligned}$$

Proof i. We prove $\|f\|_1 \leq \|f\|_2$ later. (Reader: fill in the reference once we've got to it!) For the other inequality:

$$\|f\|_2 = \sqrt{\int_0^1 |f(x)|^2\,dx} \leq \sqrt{\int_0^1 \|f\|_\infty^2\,dx} = \|f\|_\infty.$$

ii. Follows from (i) (with $f_n - f$ in place of f). □

Remark A4.11 Here's a summary of the implications:

$$\begin{tikzcd}
 & & \|\cdot\|_\infty \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & & \\
 & \|\cdot\|_2 \arrow[dl, Rightarrow] & & \text{pointwise} \arrow[dr, Rightarrow] & \\
\|\cdot\|_1 & & & & \text{a.e.}
\end{tikzcd}$$

That is, $\|\cdot\|_\infty$-convergence implies $\|\cdot\|_2$-convergence, and so on. Uniform convergence is the strongest type there is.

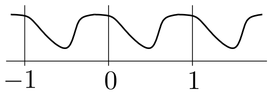
No further implications hold: there are examples of sequences of functions that converge in $\|\cdot\|_1$ but not in $\|\cdot\|_2$, etc.

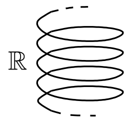
A5 Periodic functions

For the lecture of 27 January 2014; part one of two

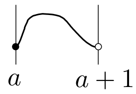
You know what it means for a function f on $\mathbb{R}$ to be 1-periodic: $f(x+1) = f(x)$ for all x. But in fact, there are at least three ways to think about a periodic function: as a...

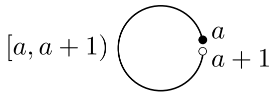
i. 1-periodic function $\mathbb{R} \to \mathbb{C}$:



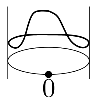


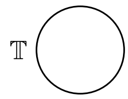
ii. function $[a, a+1) \to \mathbb{C}$
(for any $a \in \mathbb{R}$):





iii. function $\mathbb{T} \to \mathbb{C}$,
where $\mathbb{T}$ is the circle:





In (i), it's best to think of the real line $\mathbb{R}$ as coiled up into a spiral of period 1, so that $x, x \pm 1, x \pm 2, \ldots$ all lie on the same vertical line.

In (ii), it's best to think of the interval $[a, a+1)$ as bent round into a circle, as shown. (The values of a we most commonly use are 0 and $-1/2$.)

The viewpoint in (iii) is ultimately the most satisfactory. To get from (i) to (iii), think of pushing down on the coil to squash it into a circle. To get from (ii) to (iii), join the two ends of the arc.

Formally, $\mathbb{T}$ is the quotient group $\mathbb{R}/\mathbb{Z}$. (In case you've forgotten what this means, there is an equivalence relation $\sim$ on $\mathbb{R}$ given by $x \sim y \iff x - y \in \mathbb{Z}$; then $\mathbb{R}/\mathbb{Z}$ is the set of equivalence classes.) The elements of $\mathbb{T}$ are the elements of $\mathbb{R}$ *but* with x regarded as the same as $x + n$ in $\mathbb{T}$ whenever $x \in \mathbb{R}$ and $n \in \mathbb{Z}$. So,

$$\ldots, -1.9, -0.9, 0.1, 1.1, 2.1, \ldots$$

are all names for the same element of $\mathbb{T}$. (Compare the fact that

$$\ldots, -19, -9, 1, 11, 21, \ldots$$

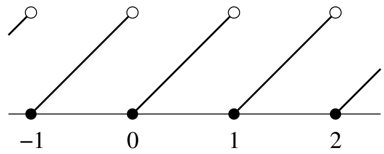
are all names for the same element of $\mathbb{Z}/10\mathbb{Z}$, i.e. the same integer mod 10.)

There are one-to-one correspondences between functions as in (i), functions as in (ii), and functions as in (iii). We will switch freely between the three ways of thinking about periodic functions, but most often, we will adopt the viewpoint of (iii).

We will also take for granted some easy lemmas about periodic functions, e.g. that any linear combination or product of 1-periodic functions is again 1-periodic.

Definition A5.1 A function $\mathbb{T} \to \mathbb{C}$ is **continuous** if the corresponding 1-periodic function $\mathbb{R} \to \mathbb{C}$ is continuous.

Note that a 1-periodic function $f\colon \mathbb{R} \to \mathbb{C}$ is continuous if and only if its restriction $\widetilde{f}$ to $[0,1)$ is continuous *and* $\lim_{x\to 1-} \widetilde{f}(x) = \widetilde{f}(0)$. (Here $\lim_{x\to 1-}$ means the limit as x tends to 1 from below.) To see why the second condition is needed, consider the function $f\colon \mathbb{R} \to \mathbb{C}$ defined by $f(x) = x - \lfloor x \rfloor$, where $\lfloor x \rfloor$ is the integer part of x:



The restriction of f to $[0,1)$ is continuous, but f itself is not continuous, because $\lim_{x\to 1-} f(x) \neq f(0)$.

Definition A5.2 A function $f\colon \mathbb{T} \to \mathbb{C}$ is **integrable** if the corresponding function $\widetilde{f}\colon [0,1) \to \mathbb{C}$ is integrable. We then define

$$\int_{\mathbb{T}} f(x)\,dx = \int_0^1 \widetilde{f}(x)\,dx, \qquad \|f\|_1 = \|\widetilde{f}\|_1, \qquad \text{etc.}$$

Remarks A5.3 i. Definition A5.2 is unchanged if we replace 0 by a and 1 by $a+1$, for any $a \in \mathbb{R}$.

ii. A little common sense is called for. When we say that a 1-*periodic* function $f\colon \mathbb{R} \to \mathbb{C}$ is integrable, it does not mean that f itself is integrable! For example, if $f\colon \mathbb{R} \to \mathbb{C}$ has constant value 6, then f is not integrable *as an ordinary function*, but it is integrable *as a* 1-*periodic function* (and its integral is 6). In practice, confusion shouldn't arise.

Functions on a half-open interval can be continuous but not integrable, since they might fail to be bounded by shooting off to $\pm\infty$ or oscillating wildly near the open end. But things are easier in the world of *periodic* functions:

Lemma A5.4 *Every continuous,* 1-*periodic function is integrable.*

Proof Let $f\colon \mathbb{R} \to \mathbb{C}$ be a continuous 1-periodic function. We have to prove that its restriction $\widetilde{f}\colon [0,1) \to \mathbb{C}$ is integrable. We know that $\widetilde{f}$ is continuous, so it is enough to prove that $\widetilde{f}$ is bounded. Indeed, $f|_{[0,1]}$ is a continuous function on a closed bounded interval, and therefore bounded; hence $f|_{[0,1)} = \widetilde{f}$ is certainly bounded. $\square$

A6 The inner product

For the lecture of 27 January 2014; part two of two

You're familiar with the scalar product $v.w$ of two vectors $v, w \in \mathbb{R}^n$, defined by $v.w = \sum_i v_i w_i$. Sometimes this is written as $\langle v, w \rangle$. Perhaps you're familiar with the complex version: given $v, w \in \mathbb{C}^n$, we put $\langle v, w \rangle = \sum_i v_i \overline{w_i}$. There is also a version for complex-valued *functions*:

Definition A6.1 Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions. We define

$$\langle f, g \rangle = \int_{\mathbb{T}} f(x)\overline{g(x)}\, dx \in \mathbb{C}.$$

In order for this definition to make sense, we need to know that the function $f \cdot \overline{g}$ is integrable. Lemmas A3.4(iii) and A3.7 guarantee this.

Lemma A6.2 *Let $f, g, h\colon \mathbb{T} \to \mathbb{C}$ be integrable functions, and let $a, b \in \mathbb{C}$. Then:*

i. $\langle g, f \rangle = \overline{\langle f, g \rangle}$*;*

ii. $\langle af + bg, h \rangle = a\langle f, h \rangle + b\langle g, h \rangle$ *and* $\langle f, ag + bh \rangle = \overline{a}\langle f, g \rangle + \overline{b}\langle f, h \rangle$*;*

iii. $\langle f, f \rangle = \|f\|_2^2 \geq 0$.

Proof Sheet 1, q.6. □

The properties of $\langle \cdot, \cdot \rangle$ stated in this lemma *nearly* say that it is an inner product. All that prevents it from being one is that $\langle f, f \rangle = 0$ does not quite imply $f = 0$; it only implies that $f = 0$ almost everywhere (by Lemma A4.5(i)). However, I will abuse terminology slightly by referring to $\langle f, g \rangle$ as the inner product of f and g anyway.

Here are some further properties of $\langle \cdot, \cdot \rangle$.

Lemma A6.3 *Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions. Then:*

i. $\|f + g\|_2^2 = \|f\|_2^2 + \|g\|_2^2 + 2\mathrm{Re}\langle f, g \rangle$ *('cosine rule');*

ii. if $\langle f, g \rangle = 0$ *then* $\|f + g\|_2^2 = \|f\|_2^2 + \|g\|_2^2$ *('Pythagoras').*

Proof For (i), use $\|h\|_2^2 = \langle h, h \rangle$. Part (ii) follows immediately. □

Here is the most fundamental result about the inner product.

Theorem A6.4 (Cauchy–Schwarz inequality) *Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions. Then*

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2.$$

Proof Sheet 1, q.6. □

Note the modulus sign on the left-hand side. Without it, the inequality would not even make sense, since $\langle f, g \rangle$ is not usually a real number.

From this, we deduce that $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$ (Sheet 1, q.6 again; see also Lemma A4.3). We also deduce the following, which completes the proof of Lemma A4.10:

Lemma A6.5 *Let* $f\colon [0,1) \to \mathbb{C}$ *be an integrable function. Then* $\|f\|_1 \le \|f\|_2$.

Proof Let us apply the Cauchy–Schwarz inequality to $|f|$ and the constant function 1. We have

$$\langle |f|, 1\rangle = \|f\|_1, \qquad \|\,|f|\,\|_2 = \|f\|_2, \qquad \|1\|_2 = 1,$$

giving $\|f\|_1 \le \|f\|_2 \cdot 1 = \|f\|_2$, as required. $\square$

There is a cousin of the Cauchy–Schwarz inequality that is also useful:

Lemma A6.6 *Let* $f, g\colon \mathbb{T} \to \mathbb{C}$ *be integrable functions. Then*

$$|\langle f, g\rangle| \le \|f\|_1 \|g\|_\infty.$$

Proof We have

$$\begin{aligned} |\langle f, g\rangle| &= \left| \int_{\mathbb{T}} f(x)\overline{g(x)}\,dx \right| \le \int_{\mathbb{T}} \left| f(x)\overline{g(x)} \right| dx \\ &= \int_{\mathbb{T}} |f(x)| \cdot |g(x)|\,dx \le \int_{\mathbb{T}} |f(x)| \cdot \|g\|_\infty\,dx = \|f\|_1 \|g\|_\infty, \end{aligned}$$

using Lemma A3.9 in the first inequality. $\square$

Remark A6.7 (Non-examinable.) Both the Cauchy–Schwarz inequality and Lemma A6.6 are special cases of a result known as Hölder's inequality, which states that $|\langle f, g\rangle| \le \|f\|_p \|g\|_q$ whenever $1/p + 1/q = 1$ (with $1 \le p \le \infty$, $1 \le q \le \infty$). Here $\|f\|_p$ and $\|g\|_q$ are defined as in Remark A4.2(iv).

Lemma A6.6 easily implies two more small results, both of which will be useful later:

Lemma A6.8 *Let* $f\colon \mathbb{T} \to \mathbb{C}$ *be an integrable function. Then* $\|f\|_2 \le \sqrt{\|f\|_1 \|f\|_\infty}$.

Proof Put $g = f$ in Lemma A6.6. $\square$

Lemma A6.9 *Let* $f_1, f_2, \ldots, f, g\colon \mathbb{T} \to \mathbb{C}$ *be integrable functions. Suppose that* $f_n \to f$ *as* $n \to \infty$ *in* $\|\cdot\|_1$. *Then* $\langle f_n, g\rangle \to \langle f, g\rangle$ *as* $n \to \infty$.

Proof For each n, we have

$$|\langle f_n, g\rangle - \langle f, g\rangle| = |\langle f_n - f, g\rangle| \le \|f_n - f\|_1 \|g\|_\infty,$$

and $\|f_n - f\|_1 \to 0$ as $n \to \infty$, so the result follows. $\square$

A7 Characters and Fourier series

For the lecture of 30 January 2014; part one of two

Among all periodic functions, certain ones are special. These are the so-called 'characters'. Fourier theory can be seen as an attempt to build all periodic functions out of characters.

Let $k \in \mathbb{Z}$. We define $e_k : \mathbb{R} \to \mathbb{C}$ by $e_k(x) = e^{2\pi ikx}$. This function is 1-periodic, so can be seen as a function $e_k : \mathbb{T} \to \mathbb{C}$, the kth **character** of $\mathbb{T}$.

Remarks A7.1 i. The notation e_k is not standard. No one outside this class will know what you mean by 'e_k' unless you define it.

ii. The apparently strange terminology 'character of $\mathbb{T}$' will be put into context in the very last part of this course.

Here are some elementary properties of the characters.

Lemma A7.2 *Let $k \in \mathbb{Z}$ and $x, y \in \mathbb{T}$. Then:*

i. e_k *is continuous;*

ii. $|e_k(x)| = 1$*;*

iii. $e_k(x + y) = e_k(x)e_k(y)$, $e_k(-x) = 1/e_k(x)$, *and* $e_k(0) = 1$.

Note that in (iii), it does make sense to add and subtract elements of $\mathbb{T}$, because $\mathbb{T}$ is by definition a group (the quotient $\mathbb{R}/\mathbb{Z}$).

Proof Straightforward. □

Some further elementary properties:

Lemma A7.3 *Let $k, \ell \in \mathbb{Z}$. Then:*

i. $e_{k+\ell} = e_k \cdot e_\ell$, $e_{-k} = 1/e_k$, *and* $e_0 = 1$*;*

ii. $e_{-k} = \overline{e_k}$*;*

iii. $e_k = e_1^k$.

Proof Straightforward. □

We now come to a crucial property of the characters.

Lemma A7.4 *The characters $(e_k)_{k\in\mathbb{Z}}$ are orthonormal. That is, for $k, \ell \in \mathbb{Z}$,*

$$\langle e_k, e_\ell \rangle = \begin{cases} 1 & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell. \end{cases}$$

Proof We have

$$\langle e_k, e_\ell \rangle = \int_{\mathbb{T}} e_k(x)\overline{e_\ell(x)}\, dx = \int_0^1 e_k(x)e_{-\ell}(x)\, dx = \int_0^1 e_{k-\ell}(x)\, dx,$$

using Lemma A7.3. If $k = \ell$ then the integrand is $e_0 = 1$, so $\langle e_k, e_\ell \rangle = 1$. If $k \neq \ell$ then

$$\langle e_k, e_\ell \rangle = \left[\frac{1}{2\pi i(k - \ell)} e^{2\pi i(k-\ell)x} \right]_0^1 = 0,$$

as required. □

You can think of the characters as analogous to the standard basis vectors in $\mathbb{R}^n$ (which are also orthonormal). When we express a point of $\mathbb{R}^n$ in terms of its coordinates, we are viewing it as a linear combination of the standard basis vectors. Similarly, in Fourier theory, we seek to view any periodic function as a linear combination of the characters. The analogy is not exact, because there are *infinitely* many characters, so we have to take *infinite* linear combinations of characters. This is what gives the subject its subtlety.

Here are the central definitions of this course.

Definition A7.5 Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function.

i. For $k \in \mathbb{Z}$, the kth **Fourier coefficient** of f is

$$\hat{f}(k) = \langle f, e_k \rangle.$$

ii. For $n \geq 0$, the nth **Fourier partial sum** of f is the function

$$S_n f = \sum_{k=-n}^{n} \hat{f}(k) e_k \colon \mathbb{T} \to \mathbb{C}.$$

iii. The **Fourier series** of f is the expression

$$Sf = \sum_{k=-\infty}^{\infty} \hat{f}(k) e_k.$$

Remarks A7.6 i. Explicitly,

$$\begin{aligned}
\hat{f}(k) &= \int_0^1 f(x) e^{-2\pi i k x}\, dx \qquad (k \in \mathbb{Z}), \\
(S_n f)(x) &= \sum_{k=-n}^{n} \hat{f}(k) e^{2\pi i k x} \qquad (n \geq 0,\ x \in \mathbb{T}), \\
(Sf)(x) &= \sum_{k=-\infty}^{\infty} \hat{f}(k) e^{2\pi i k x}.
\end{aligned}$$

ii. The Fourier series of f is $\sum_{k=-\infty}^{\infty} \langle f, e_k \rangle e_k$. Compare: if $u_1, \ldots, u_n$ denote the standard basis vectors of $\mathbb{R}^n$, then $v = \sum_{k=1}^{n} (v.u_k) u_k$ for all $v \in \mathbb{R}^n$. So, we might guess that f is 'equal' to its Fourier series (whatever that means). The central question of this subject is whether, and in what sense, this is actually true.

We finish by recording two basic properties of Fourier coefficients.

Lemma A7.7 *Let $k \in \mathbb{Z}$. Then:*

i. $\widehat{af + bg}(k) = a\hat{f}(k) + b\hat{g}(k)$ for all $a, b \in \mathbb{C}$ and integrable $f, g\colon \mathbb{T} \to \mathbb{C}$.

ii. Let $f_1, f_2, \ldots, f\colon \mathbb{T} \to \mathbb{C}$ be integrable functions such that $f_n \to f$ as $n \to \infty$ in $\|\cdot\|_1$. Then $\widehat{f_n}(k) \to \hat{f}(k)$ as $n \to \infty$.

Proof Part (i) follows from Lemma A6.2, and part (ii) from Lemma A6.9. $\square$

A8 Trigonometric polynomials

For the lecture of 30 January 2014; part two of two

As we saw in the last section, Fourier theory asks whether a periodic function f can be expressed as a linear combination $Sf = \sum_{k=-\infty}^{\infty} \hat{f}(k)e_k$ of characters e_k. This is, in general, an *infinite* linear combination (whatever that means). But to get started, it is useful to consider the *finite* linear combinations of characters. These are called trigonometric polynomials.

Definition A8.1 A function $g\colon \mathbb{T} \to \mathbb{C}$ is a **trigonometric polynomial** if there exist $n \geq 0$ and $c_{-n}, \ldots, c_0, \ldots, c_n \in \mathbb{C}$ such that $g = \sum_{k=-n}^{n} c_k e_k$. The **degree** of g is the least n for which this is possible.

Example A8.2 For any integrable $f\colon \mathbb{T} \to \mathbb{C}$ and $n \geq 0$, the nth Fourier partial sum $S_n f = \sum_{k=-n}^{n} \hat{f}(k)e_k$ is a trigonometric polynomial of degree $\leq n$.

It is perhaps not obvious that the coefficients of a trigonometric polynomial are unique. The next two results show that, in fact, they are.

Lemma A8.3 *Let $g = \sum_{k=-n}^{n} c_k e_k$ be a trigonometric polynomial. Then*

$$\hat{g}(k) = \begin{cases} c_k & \text{if } |k| \leq n, \\ 0 & \text{if } |k| > n \end{cases}$$

($k \in \mathbb{Z}$).

Proof We have

$$\begin{aligned} \hat{g}(k) = \langle g, e_k \rangle &= \left\langle \sum_{\ell=-n}^{n} c_\ell e_\ell, e_k \right\rangle = \sum_{\ell=-n}^{n} c_\ell \langle e_\ell, e_k \rangle \\ &= \begin{cases} c_k & \text{if } -n \leq k \leq n \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

where in the last step we used the fact that the characters are orthonormal. □

Corollary A8.4 *If $\sum_{k=-n}^{n} c_k e_k = \sum_{k=-n}^{n} d_k e_k$ then $c_k = d_k$ for all $k \in \{-n, \ldots, 0, \ldots, n\}$.* □

In other words, the characters are linearly independent.

Example A8.5 Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then for all $n \geq 0$ and $k \in \mathbb{Z}$,

$$\widehat{S_n f}(k) = \begin{cases} \hat{f}(k) & \text{if } |k| \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $S_n f$ and f have the same kth Fourier coefficients for $|k| \leq n$.

Here is a baby version of the whole of Fourier theory.

Proposition A8.6 *Let $n \geq 0$. Then the functions*

$$\{\textit{trigonometric polynomials of degree } \leq n\} \rightleftarrows \mathbb{C}^{2n+1}$$

given by

$$\begin{array}{rcl} g & \mapsto & \big(\hat{g}(-n), \ldots, \hat{g}(0), \ldots, \hat{g}(n)\big) \\ \sum_{k=-n}^{n} c_k e_k & \leftarrow\!\shortmid & (c_{-n}, \ldots, c_0, \ldots, c_n) \end{array}$$

are mutually inverse.

Proof

- Let $g = \sum_{k=-n}^{n} c_k e_k$ be a trigonometric polynomial of degree $\leq n$. We must show that $g = \sum_{k=-n}^{n} \hat{g}(k) e_k$. This follows from Lemma A8.3.
- Let $(c_{-n}, \ldots, c_n) \in \mathbb{C}^{2n+1}$ and put $g = \sum_{k=-n}^{n} c_k e_k$. We must show that $\hat{g}(k) = c_k$ for all $k \in \{-n, \ldots, n\}$. This also follows from Lemma A8.3. $\square$

Fantasy A8.7 We can fantasize about extending Proposition A8.6 to a pair of mutually inverse functions

$$\{\text{nice functions } \mathbb{T} \to \mathbb{C}\} \rightleftarrows \{\text{nice double sequences in } \mathbb{C}\}$$

given by

$$\begin{array}{rcl} f & \mapsto & \big(\hat{f}(k)\big)_{k=-\infty}^{\infty} \\ \sum_{k=-\infty}^{\infty} c_k e_k & \leftarrow\!\shortmid & (c_k)_{k=-\infty}^{\infty}. \end{array}$$

The rest of the course explores this fantasy.

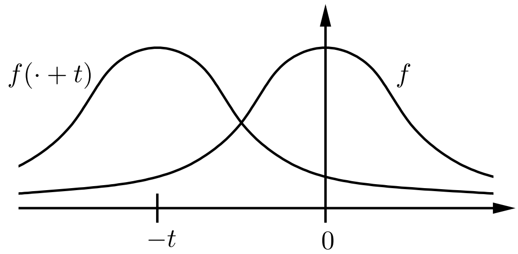


Figure A.4: Translating a function.

A9 Integrable functions are sort of continuous

For the lecture of 3 February 2014

Integrable functions are not necessarily continuous. However, integrable functions *do* satisfy a continuity-like condition of a not-so-obvious kind, as follows.

Let $f\colon \mathbb{T} \to \mathbb{C}$ be a function. Given $t \in \mathbb{T}$, we obtain a new function

$$f(\cdot + t)\colon \mathbb{T} \to \mathbb{C}$$

defined by

$$x \mapsto f(x+t).$$

Geometrically, this means shifting the graph of the function by t units to the left (Fig. A.4). (Or really, since $\mathbb{T}$ is a circle, it means rotating the graph by a fraction t of a revolution.) It's not so hard to see (try it!) that

$$f \text{ is continuous} \iff f(\cdot + t) \to f \text{ pointwise as } t \to 0$$

—in other words, for each $x \in \mathbb{T}$, $f(x+t) \to f(x)$ as $t \to 0$. Similarly, it's not hard to see that

$$f \text{ is uniformly continuous} \iff f(\cdot + t) \to f \text{ in } \|\cdot\|_\infty \text{ as } t \to 0$$

—in other words, $\|f(\cdot + t) - f\|_\infty \to 0$ as $t \to 0$. Obviously, neither condition is satisfied by an arbitrary integrable function. However, it *is* true that for any integrable function,

$$f(\cdot + t) \to f \text{ in } \|\cdot\|_1 \text{ as } t \to 0. \tag{A:4}$$

Proving this will take the rest of this lecture, and will require us to go right back to the definition of integrability.

Here's the plan. First we'll introduce a class of particularly simple functions, the so-called 'step functions'. We'll show that every integrable function can be closely approximated by a step function. The proof of (A:4) for arbitrary integrable functions f proceeds in two steps: (i) prove it for step functions (which is relatively easy, as step functions are so simple); (ii) extend the result to an arbitrary integrable f by approximating it with step functions.

Definition A9.1 Let $I \subseteq \mathbb{R}$ be a bounded interval. A **step function** is a function $f\colon I \to \mathbb{C}$ such that $f = \sum_{k=1}^{n} c_k \chi_{J_k}$ for some $n \geq 0$, $c_1, \ldots, c_n \in \mathbb{C}$, and bounded intervals $J_1, \ldots, J_n \subseteq I$.

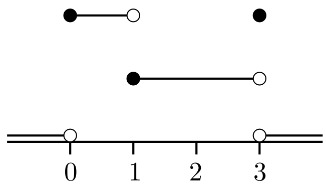


Figure A.5: A step function.

In other words, a step function is a finite linear combination of characteristic functions of intervals.

Example A9.2 $2\chi_{[0,3]} - \chi_{[1,3)}$ is a step function (Figure A.5).

Step functions are integrable, by Example A3.2(ii) and Lemma A3.4. Indeed,

$$\int_I \sum_k c_k \chi_{J_k}(x)\, dx = \sum_k c_k \left|J_k\right|.$$

Of course, not every integrable function is a step function. However, there is a sense in which every integrable function can be well approximated by a step function. This is somewhat similar to the fact that arbitrary images can be displayed on a computer screen, which is a grid of discrete pixels; the finer the grid is, the better the quality of the display.

To make 'well approximated' precise, we introduce some more terminology.

Definition A9.3 Let I be a bounded interval, and let $\|\cdot\|$ stand for $\|\cdot\|_1$, $\|\cdot\|_2$ or $\|\cdot\|_\infty$. Let $\mathcal{F}$ be a set of functions $I \to \mathbb{C}$, and let $\mathcal{G} \subseteq \mathcal{F}$. Then $\mathcal{G}$ is **dense** in $\mathcal{F}$ with respect to $\|\cdot\|$ if:

for all $f \in \mathcal{F}$, for all $\varepsilon > 0$, there exists $g \in \mathcal{G}$ such that $\|f - g\| < \varepsilon$.

(You may be familiar with something like this definition from the theory of metric spaces; for example, $\mathbb{Q}$ is dense in $\mathbb{R}$.)

Proposition A9.4 *Let I be a bounded interval. Then $\{$step functions $I \to \mathbb{C}\}$ is dense in $\{$integrable functions $I \to \mathbb{C}\}$ with respect to $\|\cdot\|_1$.*

Proof We prove it just for $I = [0, 1)$, since this is the case that will matter most to us and the proof for other bounded intervals is very similar.

First take a *real*-valued integrable function $f\colon [0, 1) \to \mathbb{R}$, and let $\varepsilon > 0$. By definition of integration, we can choose a partition P of $[0, 1)$ such that

$$L(f, P) > \int_0^1 f(x)\, dx - \varepsilon.$$

Here $L(f, P)$ is the lower Darboux sum: if P is the partition

$$0 = x_0 < x_1 < \cdots < x_n = 1$$

and we put $m_k = \inf\{f(x) : x_{k-1} \leq x \leq x_k\}$ (for $1 \leq k \leq n$), then

$$L(f, P) = \sum_{k=1}^{n} m_k (x_k - x_{k-1}).$$

Put $J_k = [x_{k-1}, x_k)$ and put

$$g = \sum_{k=1}^{n} m_k \chi_{J_k}.$$

(Draw a picture!) Then g is a step function, with $L(f,P) = \int_0^1 g(x)\,dx$. We have $g(x) \leq f(x)$ for all $x \in [0,1)$, so

$$\|f - g\|_1 = \int_0^1 (f(x) - g(x))\,dx = \int_0^1 f(x)\,dx - L(f,P) < \varepsilon,$$

as required.

Now take an arbitrary integrable function $f\colon [0,1) \to \mathbb{C}$, say $f = f_1 + if_2$ with $f_1, f_2\colon [0,1) \to \mathbb{R}$. Let $\varepsilon > 0$. By definition of integrability of complex-valued functions (Definition A3.1(i)), f_1 and f_2 are integrable. So by the previous paragraph, we can choose step functions $g_1, g_2\colon [0,1) \to \mathbb{R}$ such that

$$\|f_1 - g_1\|_1 < \varepsilon/2, \qquad \|f_2 - g_2\|_1 < \varepsilon/2.$$

Put $g = g_1 + ig_2\colon [0,1) \to \mathbb{C}$, which is also a step function. Then

$$\|f - g\|_1 = \|(f_1 - g_1) + i(f_2 - g_2)\|_1 \leq \|f_1 - g_1\|_1 + \|f_2 - g_2\|_1 < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

(using Lemma A4.3), as required. □

We are now ready to prove that integrable functions are 'sort of continuous', in the sense of (A:4), following the plan described above.

Theorem A9.5 *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then $f(\cdot + t) \to f$ in $\|\cdot\|_1$ as $t \to 0$ (that is, $\|f(\cdot + t) - f\|_1 \to 0$ as $t \to 0$).*

Proof For the duration of this proof only, let us say that an integrable function $g\colon \mathbb{T} \to \mathbb{C}$ is 'good' if $g(\cdot + t) \to g$ in $\|\cdot\|_1$ as $t \to 0$. We will prove that every integrable function is good.

First, for any interval $J \subseteq [0,1)$, the characteristic function χ_J is good. Indeed, write $a = \inf J$ and $b = \sup J$. If $a = b$, or if $a = 0$ and $b = 1$, then $\|\chi_J(\cdot + t) - \chi_J\|_1 = 0$ for all t. Otherwise, we can show that

$$\|\chi_J(\cdot + t) - \chi_J\|_1 = 2|t|$$

whenever $|t|$ is sufficiently small (exercise; see Figure A.6). Hence $\|\chi_J(\cdot + t) - \chi_J\|_1 \to 0$ as $t \to 0$.

Second, every step function $g\colon [0,1) \to \mathbb{C}$ is good. Indeed, if $g = \sum_{k=1}^{n} c_k \chi_{J_k}$ (in the usual notation) then

$$\begin{aligned}
\|g(\cdot + t) - g\|_1 &= \left\| \sum_{k=1}^{n} c_k \left\{ \chi_{J_k}(\cdot + t) - \chi_{J_k} \right\} \right\|_1 \\
&\leq \sum_{k=1}^{n} |c_k| \left\| \chi_{J_k}(\cdot + t) - \chi_{J_k} \right\|_1 \\
&\to 0
\end{aligned}$$

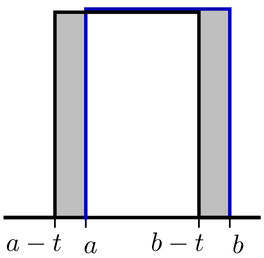


Figure A.6: The area between a characteristic function and a translation of it by a small distance t is $2|t|$.

as $t \to 0$, by the first part.

Finally, every integrable function $f\colon [0,1) \to \mathbb{C}$ is good. Indeed, let $\varepsilon > 0$. By Proposition A9.4, we can choose a step function $g\colon [0,1) \to \mathbb{C}$ such that $\|f-g\|_1 < \varepsilon/3$. This implies that for all $t \in \mathbb{R}$,

$$\begin{aligned}
\|f(\cdot+t) - g(\cdot+t)\|_1 &= \int_{\mathbb{T}} |f(x+t) - g(x+t)|\,dx \\
&= \int_{\mathbb{T}} |f(y) - g(y)|\,dy = \|f-g\|_1 < \varepsilon/3,
\end{aligned}$$

where the second equality is by substitution. (Intuitively, the area between the graphs f and g is unchanged if we shift everything horizontally by t.) Also, by the second part, we can choose $\delta > 0$ such that for all $t \in (-\delta, \delta)$, we have $\|g(\cdot+t) - g\|_1 < \varepsilon/3$. Now for all $t \in (-\delta, \delta)$, we have

$$\begin{aligned}
\|f(\cdot+t) - f\|_1 &\le \|f(\cdot+t) - g(\cdot+t)\|_1 + \|g(\cdot+t) - g\|_1 + \|g-f\|_1 \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon,
\end{aligned}$$

as required. □

Chapter B

Convergence of Fourier series in the 2- and 1-norms

In this chapter, we'll prove:

> for *any* integrable function $f\colon \mathbb{T} \to \mathbb{C}$, the Fourier series of f converges to f in both $\|\cdot\|_2$ and $\|\cdot\|_1$.

It's as simple as that. Compare the long and complicated saga of pointwise convergence recounted in Section A2. In contrast, this result is clean, easily stated, and not too difficult to prove.

The 2-norm will play a much greater role than the 1-norm in this chapter. Because $\|\cdot\|_2$ goes hand in hand with inner products (recalling that $\|f\|_2^2 = \langle f, f\rangle$), working in the 2-norm has a great deal in common with ordinary Euclidean geometry. It's the most easily visualized context to work in.

B1 Nothing is better than a Fourier partial sum

For the lecture of 6 February 2014; part one of two

The title of this section is made precise by part (i) of the following lemma, illustrated in Figure B.1. Recall that given an integrable function f, its nth Fourier partial sum $S_n f$ is a trigonometric polynomial of degree $\leq n$.

Lemma B1.1 *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function, and let $n \geq 0$. Then:*

i. $\|f - S_n f\|_2 \leq \|f - g\|_2$ *whenever g is a trigonometric polynomial of degree $\leq n$.*

ii. $\|f\|_2^2 = \|S_n f\|_2^2 + \|f - S_n f\|_2^2$.

iii. $\|S_n f\|_2^2 = \sum_{k=-n}^{n} |\hat{f}(k)|^2$.

Proof Before we prove any of the three parts, we show that $\langle f - S_n f, h\rangle = 0$ for all trigonometric polynomials h of degree $\leq n$. (Figure B.1 makes this look

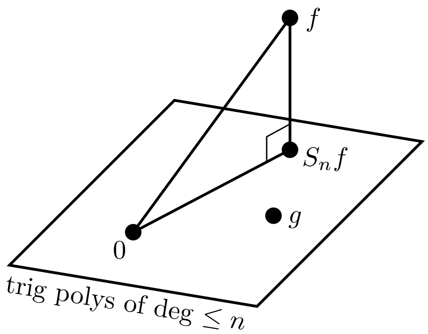


Figure B.1: Approximating a function f by a trigonometric polynomial of degree $\leq n$.

plausible.) Indeed, by linearity, it is enough to prove this when $h = e_k$ for some $k \in \{-n, \dots, 0, \dots, n\}$, and

$$\langle f - S_n f, e_k \rangle = \langle f, e_k \rangle - \langle S_n f, e_k \rangle = \hat{f}(k) - \widehat{S_n f}(k) = 0$$

by Example A8.5, as required.

Now let g be a trigonometric polynomial of degree $\leq n$. By what we have just shown, $\langle f - S_n f, S_n f - g \rangle = 0$, so

$$\begin{aligned} \|f - g\|_2^2 &= \|(f - S_n f) + (S_n f - g)\|_2^2 \\ &= \|f - S_n f\|_2^2 + \|S_n f - g\|_2^2 \end{aligned} \tag{B:1}$$

by Pythagoras (Lemma A6.3(ii)). Part (i) follows, then part (ii) by putting $g = 0$ in (B:1).

For part (iii), we have

$$\begin{aligned} \|S_n f\|_2^2 &= \left\langle \sum_{k=-n}^{n} \hat{f}(k) e_k, \sum_{\ell=-n}^{n} \hat{f}(\ell) e_l \right\rangle \\ &= \sum_{k,\ell=-n}^{n} \hat{f}(k) \overline{\hat{f}(\ell)} \langle e_k, e_\ell \rangle \\ &= \sum_{k=-n}^{n} \left| \hat{f}(k) \right|^2, \end{aligned}$$

where the first two equalities use Lemma A6.2 and the third uses the orthonormality of the characters (Lemma A7.4). □

To elaborate a little on the title of this section: among all trigonometric polynomials of degree $\leq n$, nothing approximates f better than the Fourier partial sum $S_n f$.

Proposition B1.2 *Let $f \colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then:*

i. $\sum_{k=-\infty}^{\infty} \left| \hat{f}(k) \right|^2$ *converges; indeed,* $\sum_{k=-\infty}^{\infty} \left| \hat{f}(k) \right|^2 \leq \|f\|_2^2$.

ii. ***(Riemann–Lebesgue Lemma)*** $\hat{f}(k) \to 0$ *as* $k \to \pm\infty$.

Proof By parts (ii) and (iii) of Lemma B1.1, we have $\sum_{k=-n}^{n} |\hat{f}(k)|^2 \leq \|f\|_2^2$ for all $n \geq 0$. The results follow. □

The Riemann–Lebesgue lemma tells us that not every double sequence $(c_k)_{k=-\infty}^{\infty}$ of complex numbers arises as the sequence of Fourier coefficients of some function. This is already a substantial result. Looking back again at Fantasy A8.7, we begin to get a sense of what a 'nice' double sequence might be.

The next lemma will help us to prove that $(S_n f)$ always converges to f in the 2-norm.

Lemma B1.3 *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. The following are equivalent:*

i. $S_n f \to f$ in $\|\cdot\|_2$;

ii. $g_n \to f$ in $\|\cdot\|_2$ for some sequence (g_n) of trigonometric polynomials.

Proof (i) $\Rightarrow$ (ii) is trivial, since each $S_n f$ is a trigonometric polynomial.

For (ii) $\Rightarrow$ (i), let $\varepsilon > 0$. Choose m such that $\|f - g_m\|_2 < \varepsilon$. Put $N = \deg(g_m)$. Then

$$\varepsilon > \|f - g_m\|_2 \geq \|f - S_N f\|_2,$$

using Lemma B1.1(i) in the second inequality. But then

$$\|f - S_N f\|_2 \geq \|f - S_{N+1} f\|_2$$

by Lemma B1.1(i) again, noting that $S_N f$ has degree $\leq N+1$. Continuing like this, we find that

$$\varepsilon > \|f - S_N f\|_2 \geq \|f - S_{N+1} f\|_2 \geq \|f - S_{N+2} f\|_2 \geq \cdots,$$

and in particular, $\|f - S_n f\|_2 < \varepsilon$ for all $n \geq N$. □

Our strategy for proving that $S_n f \to f$ in $\|\cdot\|_2$ will be to prove that $g_n \to f$ in $\|\cdot\|_2$ for some *other* sequence (g_n) of trigonometric polynomials—a sequence that is easier to work with than $(S_n f)$ itself.

In order to carry out this strategy, we need to come up with some convenient sequence (g_n) of trigonometric polynomials. How can we cook up new trigonometric polynomials? By convolution, as we'll see in the next two sections.

B2 Convolution: definition and examples

For the lecture of 6 February 2014; part two of two

You've already met convolution of functions on $\mathbb{R}$. Convolution of functions on $\mathbb{T}$, which is what we'll mostly be concerned with, is very similar.

Definition B2.1 Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions. The **convolution** $f * g\colon \mathbb{T} \to \mathbb{C}$ is the function defined by

$$(f * g)(x) = \int_{\mathbb{T}} f(t)g(x-t)\,dt$$

$(x \in \mathbb{T})$.

Remark B2.2 Recall from Section A5 that the circle $\mathbb{T}$ is a group (the quotient $\mathbb{R}/\mathbb{Z}$). The group operation is addition mod 1. So, it does make sense to add and subtract elements of $\mathbb{T}$, as we did in the definition of convolution.

Our first example of convolution is important enough to be stated as a lemma.

Lemma B2.3 *For any $k \in \mathbb{Z}$ and integrable function $f\colon \mathbb{T} \to \mathbb{C}$,*

$$f * e_k = \langle f, e_k \rangle e_k = \hat{f}(k) e_k.$$

Proof For all $x \in \mathbb{T}$,

$$\begin{aligned}
(f * e_k)(x) &= \int_{\mathbb{T}} f(t) e_k(x-t)\,dt \\
&= \int_{\mathbb{T}} f(t) e_k(x) e_k(-t)\,dt \\
&= e_k(x) \int_{\mathbb{T}} f(t) \overline{e_k(t)}\,dt \\
&= \langle f, e_k \rangle e_k(x) \\
&= \hat{f}(k) e_k(x)
\end{aligned}$$

by Lemmas A7.2 and A7.3. □

This is remarkable. It tells us that when we convolve a character e_k with anything (*anything!*), the result is a scalar multiple of e_k. This is a very special property of the characters.

Convolution has a smoothing effect. For the purposes of the following example, let us consider functions on $\mathbb{R}$ rather than $\mathbb{T}$.

Example B2.4 Given an integrable function $f\colon \mathbb{R} \to \mathbb{C}$, we have

$$\big(f * \chi_{[-1/2,1/2]}\big)(x) = \int_{x-1/2}^{x+1/2} f(t)\,dt.$$

This is a 'moving average' of f: the value of $f * \chi_{[-1/2,1/2]}$ at x is the mean value of f over the interval $[x - 1/2, x + 1/2]$ (Figure B.2).

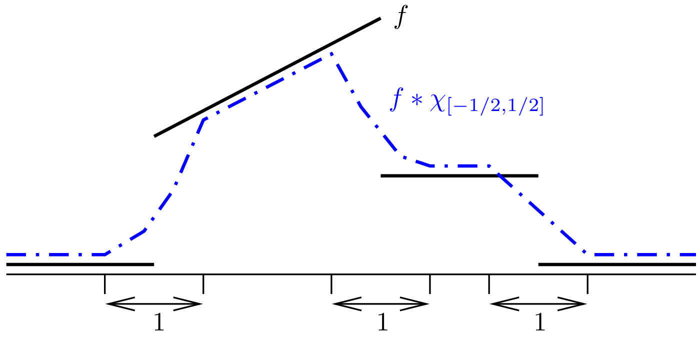


Figure B.2: Convolution has a smoothing effect.

B3 Convolution: properties

For the lecture of 10 February 2014; part one of two

Example B2.4 showed the smoothing effect of convolution. In that particular example, the convolution of two discontinuous functions was continuous. This surprising behaviour is, in fact, a completely general phenomenon:

Lemma B3.1 *Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions. Then their convolution $f * g\colon \mathbb{T} \to \mathbb{C}$ is continuous.*

This is far stronger than we might have guessed in advance. It's not merely true that the convolution of two integrable functions is integrable, or that the convolution of two continuous functions is continuous. In fact, for *any* two integrable functions f and g, no matter how discontinuous they may be, $f * g$ is continuous.

Proof Let $x, h \in \mathbb{T}$. Then

$$\begin{aligned}
|(f * g)(x+h) - (f * g)(x)| &= \left| \int_{\mathbb{T}} f(t)[g(x+h-t) - g(x-t)]\,dt \right| \\
&\leq \int_{\mathbb{T}} |f(t)|\,|g(x+h-t) - g(x-t)|\,dt \qquad \text{(B:2)} \\
&\leq \|f\|_\infty \int_{\mathbb{T}} |g(x+h-t) - g(x-t)|\,dt \\
&= \|f\|_\infty \int_{\mathbb{T}} |g(u+h) - g(u)|\,du \qquad \text{(B:3)} \\
&= \|f\|_\infty\, \|g(\cdot + h) - g\|_1 \\
&\to 0 \text{ as } h \to 0. \qquad \text{(B:4)}
\end{aligned}$$

Here (B:2) is by the triangle inequality for integration (Lemma A3.9), equation (B:3) comes from substituting $u = x - t$, and (B:4) is by Theorem A9.5 ('integrable functions are sort of continuous'). □

Here are some basic properties of convolution. In future, we will use them without explicitly referring back to this lemma.

Lemma B3.2 *Let $f, g, h\colon \mathbb{T} \to \mathbb{C}$ be integrable functions, and let $c \in \mathbb{C}$. Then:*

i. $f * g = g * f$;

ii. $(f * g) * h = f * (g * h)$;

iii. $f * (g + h) = (f * g) + (f * h)$;

iv. $f * (cg) = c(f * g)$.

Proof For (i), let $x \in \mathbb{T}$. Then

$$\begin{aligned}
(f * g)(x) &= \int_{\mathbb{T}} f(t)g(x-t)\,dt \\
&= \int_{\mathbb{T}} f(x-u)g(u)\,du \\
&= (g * f)(x),
\end{aligned}$$

by substituting $u = x - t$. The other parts are similarly straightforward. □

Remark B3.3 This lemma tells us that $+$ and $*$ give the set

$$\{\text{integrable functions } \mathbb{T} \to \mathbb{C}\}$$

the structure of a 'commutative algebra' over $\mathbb{C}$, that is, both a vector space over $\mathbb{C}$ and a ring.

Or nearly. The only missing part is that it has no multiplicative identity. Nowadays, the definition of 'ring' is usually taken to include the existence of a multiplicative identity; but in analysis especially, there are important rings that do not have one.

We'll see very soon that if the identity *did* exist, it would be the mythical 'delta function'.

Lemma B3.4 *Let $f: \mathbb{T} \to \mathbb{C}$ be an integrable function, and let $g: \mathbb{T} \to \mathbb{C}$ be a trigonometric polynomial. Then $f * g$ is a trigonometric polynomial.*

Again, this is stronger than might be expected. It's not merely true that the convolution of two trigonometric polynomials is a trigonometric polynomial. In fact, the convolution of *anything* with a trigonometric polynomial is a trigonometric polynomial.

Proof We may write $g = \sum_{k=-n}^{n} c_k e_k$ for some $n \geq 0$ and $c_k \in \mathbb{C}$. Then

$$f * g = \sum_{k=-n}^{n} c_k (f * e_k) = \sum_{k=-n}^{n} \big(c_k \hat{f}(k)\big) e_k,$$

which is a trigonometric polynomial. Here the second equality uses Lemma B2.3. $\square$

Remark B3.5 This lemma says that $\{\text{trigonometric polynomials}\}$ is an ideal in the ring $\{\text{integrable functions } \mathbb{T} \to \mathbb{C}\}$.

Example B3.6 Every Fourier partial sum of a function f is a convolution of f with a trigonometric polynomial. More exactly, using Lemma B2.3 again,

$$S_n f = \sum_{k=-n}^{n} \hat{f}(k) e_k = \sum_{k=-n}^{n} f * e_k = f * \sum_{k=-n}^{n} e_k.$$

Since $\sum_{k=-n}^{n} e_k$ is a trigonometric polynomial, Lemma B3.4 tells us in this case that $S_n f$ is also a trigonometric polynomial—which of course we already knew.

Since Fourier partial sums are important in Fourier theory, this example suggests that $\sum_{k=-n}^{n} e_k$ is also important. It is, and it has its own name:

Definition B3.7 Let $n \geq 0$. The **Dirichlet kernel** of order n is $D_n = \sum_{k=-n}^{n} e_k$.

So, Example B3.6 can be restated as:

Lemma B3.8 *$S_n f = f * D_n$, for all integrable $f: \mathbb{T} \to \mathbb{C}$ and $n \geq 0$.* $\square$

Remark B3.9 We can dream of taking $n \to \infty$ in Lemma B3.8, so that $Sf = f * D_\infty$ where $D_\infty = \sum_{k=-\infty}^{\infty} e_k$. However, the sum $\sum_{k=-\infty}^{\infty} e_k$ does not converge *anywhere*, so D_∞ does not really exist.

Here are two further properties of convolution.

Lemma B3.10 *For integrable functions $f, g\colon \mathbb{T} \to \mathbb{C}$,*

$$\|f * g\|_\infty \le \|f\|_\infty \|g\|_1.$$

Proof For all $x \in \mathbb{T}$,

$$\begin{aligned}
|(f * g)(x)| &= \left| \int_{\mathbb{T}} f(x-t)g(t)\,dt \right| \\
&\le \int_{\mathbb{T}} |f(x-t)|\,|g(t)|\,dt \\
&\le \|f\|_\infty \int_{\mathbb{T}} |g(t)|\,dt = \|f\|_\infty \|g\|_1.
\end{aligned}$$

□

The final property should remind you of an important fact about Fourier *transforms.*

Lemma B3.11 *Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions, and let $k \in \mathbb{Z}$. Then*

$$\widehat{f * g}(k) = \hat{f}(k)\hat{g}(k).$$

Proof Sheet 2, q.5. □

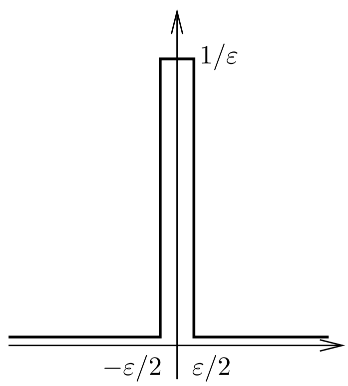


Figure B.3: Approximation to the mythical delta function.

B4 The mythical delta function

For the lecture of 10 February 2014; part two of two

This short section is largely intended as motivation for the rest of Part B. It contains no definitions or theorems, but the ideas are important for what follows.

Fact: there is no integrable function $\delta\colon \mathbb{T} \to \mathbb{C}$ such that

$$\text{for all continuous } f\colon \mathbb{T} \to \mathbb{C}, \qquad \int_{\mathbb{T}} f(x)\delta(x)\,dx = f(0). \tag{B:5}$$

(Compare Sheet 1, q.5.) But we can get close. Non-rigorously, for a 'small' $\varepsilon > 0$, put

$$\Delta_\varepsilon = \frac{1}{\varepsilon}\chi_{[-\varepsilon/2,\varepsilon/2]}\colon [-1/2, 1/2) \to \mathbb{C}$$

(Figure B.3). Then for any continuous function $f\colon \mathbb{T} \to \mathbb{C}$,

$$\int_{\mathbb{T}} f(x)\Delta_\varepsilon(x)\,dx = \frac{1}{\varepsilon}\int_{-\varepsilon/2}^{\varepsilon/2} f(x)\,dx \approx \frac{1}{\varepsilon}\int_{-\varepsilon/2}^{\varepsilon/2} f(0)\,dx = f(0),$$

using continuity in the approximate equality.

Imagine now that there *is* a function δ satisfying (B:5).

Let $f\colon \mathbb{T} \to \mathbb{C}$ be a continuous function. Then

$$f * \delta = f,$$

since

$$(f * \delta)(x) = \int_{\mathbb{T}} f(x-t)\delta(t)\,dt = f(x-0) = f(x)$$

for all $x \in \mathbb{T}$. (So δ is an identity for convolution with continuous functions; compare Remark B3.3.)

If δ was actually a trigonometric polynomial, then the equation $f * \delta = f$ together with Lemma B3.4 would imply that every continuous function f was a trigonometric polynomial. This is obviously false.

If δ is merely a *limit* of trigonometric polynomials K_n, say $K_n \to \delta$ in $\|\cdot\|_2$, then perhaps it follows that $f * K_n \to f * \delta$ in $\|\cdot\|_2$. In that case, we would have $f * K_n \to f$ in $\|\cdot\|_2$. Lemma B1.3 would then imply that $S_n f \to f$ in $\|\cdot\|_2$. This is the result we're aiming for.

We know that no δ satisfying (B:5) exists. However, the previous paragraph suggests a strategy:

> Look for a sequence (K_n) of trigonometric polynomials such that for all continuous (or even integrable) functions $f\colon \mathbb{T} \to \mathbb{C}$, $f * K_n \to f$ in $\|\cdot\|_2$.

Our fantasies about the delta function suggest that the sequence (K_n) should somehow 'converge to δ'. So, the plan now is to look for such a sequence (K_n), and show that it does what this informal chain of reasoning leads us to hope it will.

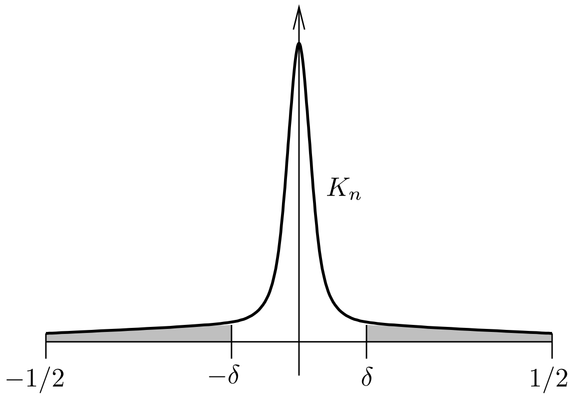


Figure B.4: Axiom PAD3 states that for large n, the shaded area is small.

B5 Positive approximations to delta

For the lecture of 13 February 2014

The last section culminated in a plan: to show that $S_n f \to f$ in $\|\cdot\|_2$ for all f,

- look for a sequence (K_n) of trigonometric polynomials that in some sense 'converges to the (non-existent) delta function'; then
- show that $f * K_n \to f$ in $\|\cdot\|_2$ for all f.

We'll carry out this plan. In this section, we give a precise meaning to the phrase 'converges to the delta function'. We also show that if (K_n) has this property, then $f * K_n \to f$ in $\|\cdot\|_2$ for all f. Actually *finding* such a sequence (K_n) is left until later.

Definition B5.1 A **positive approximation to delta (PAD)** is a sequence $(K_n)_{n=0}^{\infty}$ of integrable functions $\mathbb{T} \to \mathbb{R}$ such that:

PAD1 for all $n \geq 0$ and $t \in \mathbb{T}$, we have $K_n(t) \geq 0$;

PAD2 for all $n \geq 0$, we have $\int_{\mathbb{T}} K_n(t)\,dt = 1$;

PAD3 for all $\delta \in (0, 1/2)$, we have $\lim_{n\to\infty} \int_{\delta<|t|\leq 1/2} K_n(t)\,dt = 0$.

Remarks B5.2 i. PAD2 is inspired by the thought that if $\int_{\mathbb{T}} \delta(t) f(t)\,dt = f(0)$ for all continuous f (as in the previous section), then in particular this is true when f is the constant function 1, giving $\int_{\mathbb{T}} \delta(t)\,dt = 1$.

ii. In PAD3, the 'δ' mentioned is a real number, not the delta function!

iii. PAD2 tells us that the area under the graph of K_n is always 1. PAD3 says that as n gets larger, that area gets concentrated into an ever-narrower strip around the y-axis (Figure B.4).

iv. For this part of the theory, it's convenient to view functions on $\mathbb{T}$ as functions on $[-1/2, 1/2)$.

v. The name 'approximation to delta' is non-standard. The usual name is 'approximation to the identity', since delta is the (non-existent) identity for convolution.

vi. Almost every definition in this course is conceptually well-motivated. This one, however, is not. The conditions PAD1–3 are just what is needed in order to make the arguments work. (Other variants are possible.) However, it's only a stepping stone, which we'll use to reach theorems that are clean and free of arbitrary conditions.

Examples B5.3 i. $\left(n\chi_{[-1/2n,1/2n)}\right)_{n=1}^{\infty}$ is a PAD. (Check!) A typical element of this sequence is shown in Figure B.3, taking $\varepsilon = 1/n$.

ii. $(D_n)_{n=0}^{\infty}$ is *not* a PAD. First, it fails PAD1, since (for instance) $D_1(1/2) = -1 < 0$. More seriously, it can be shown that $\|D_n\|_1 \to \infty$ as $n \to \infty$, whereas if (K_n) is a PAD then

$$\|K_n\|_1 = \int_{\mathbb{T}} K_n(t)\, dt = 1 \tag{B:6}$$

for all n (by PAD1 and PAD2).

The rest of this section is devoted to showing that when (K_n) is a PAD, $f * K_n \to f$ in $\|\cdot\|_2$ for any integrable f. The proof is quite delicate.

The first step is to prove the weaker result that $f * K_n \to f$ in $\|\cdot\|_1$. (To see why it's 'weaker', recall Lemma A4.10.)

Proposition B5.4 *Let $(K_n)_{n=0}^{\infty}$ be a PAD and let $f \colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then $f * K_n \to f$ in $\|\cdot\|_1$.*

Proof We begin by finding an upper bound for $\|f * K_n - f\|_1$.

For all $x \in \mathbb{T}$ and $n \geq 0$,

$$\begin{aligned}
|(f * K_n)(x) - f(x)| &= \left| \int_{-1/2}^{1/2} (f(x-t) - f(x)) K_n(t)\, dt \right| \\
&\leq \int_{-1/2}^{1/2} |f(x-t) - f(x)| K_n(t)\, dt,
\end{aligned}$$

using PAD2 in the first line and PAD1 in the second.

(The idea now is roughly as follows. Let $\delta > 0$ be small. Then $|f(x-t) - f(x)|$ is small for $|t| < \delta$, at least if f is sort of continuous. Also, $K_n(t)$ is small for $|t| > \delta$, since K_n is something like the delta function.)

It follows that for all $x \in \mathbb{T}$, $n \geq 0$ and $\delta \in (0, 1/2)$,

$$\begin{aligned}
&|(f * K_n)(x) - f(x)| \\
&\leq \int_{|t|<\delta} |f(x-t) - f(x)| K_n(t)\, dt + \int_{\delta < |t| \leq 1/2} |f(x-t) - f(x)| K_n(t)\, dt,
\end{aligned}$$

so by the triangle inequality,

$$\begin{aligned}
&|(f * K_n)(x) - f(x)| \\
&\leq \int_{|t|<\delta} |f(x-t) - f(x)| K_n(t)\, dt + 2\|f\|_\infty \int_{\delta < |t| \leq 1/2} K_n(t)\, dt.
\end{aligned} \tag{B:7}$$

Integrating (B:7) with respect to x shows that for all $n \geq 0$ and $\delta \in (0, 1/2)$,

$$
\begin{aligned}
&\|f * K_n - f\|_1 \\
&\leq \int_{-1/2}^{1/2} \int_{|t|<\delta} |f(x-t) - f(x)| K_n(t)\, dt\, dx + 2\|f\|_\infty \int_{\delta<|t|\leq 1/2} K_n(t)\, dt \quad \text{(B:8)} \\
&= \int_{|t|<\delta} \|f(\cdot - t) - f\|_1 K_n(t)\, dt + 2\|f\|_\infty \int_{\delta<|t|\leq 1/2} K_n(t)\, dt. \quad \text{(B:9)}
\end{aligned}
$$

(To get (B:8), we used the fact that $2\|f\|_\infty \int_{\delta<|t|\leq 1/2} K_n(t)\, dt$ is independent of x, so that integrating it with respect to x over an interval of length 1 leaves it unchanged.)

We now prove convergence. Let $\varepsilon > 0$. By Theorem A9.5 ('integrable functions are sort of continuous'), we can choose $\delta \in (0, 1/2)$ such that for all $t \in (-\delta, \delta)$, $\|f(\cdot - t) - f\|_1 < \varepsilon/2$. By PAD3, we can then choose $N \geq 0$ such that for all $n \geq N$,

$$\int_{\delta<|t|\leq 1/2} K_n(t)\, dt < \frac{\varepsilon}{4\|f\|_\infty}.$$

So by (B:9), for all $n \geq N$,

$$
\begin{aligned}
\|f * K_n - f\|_1 &\leq \frac{\varepsilon}{2} \int_{|t|<\delta} K_n(t)\, dt + 2\|f\|_\infty \frac{\varepsilon}{4\|f\|_\infty} \\
&\leq \frac{\varepsilon}{2} \int_{-1/2}^{1/2} K_n(t)\, dt + \frac{\varepsilon}{2} \quad \text{(B:10)} \\
&= \varepsilon, \quad \text{(B:11)}
\end{aligned}
$$

using PAD1 in (B:10) and PAD2 in (B:11). □

We now know that $f * K_n \to f$ in $\|\cdot\|_1$ for any integrable f. It is relatively easy to deduce the stronger result that $f * K_n \to f$ in $\|\cdot\|_2$. The key is Lemma A6.8, which gives an upper bound on the 2-norm in terms of the 1- and ∞-norms.

Proposition B5.5 *Let $(K_n)_{n=0}^\infty$ be a PAD and let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then $f * K_n \to f$ in $\|\cdot\|_2$.*

Proof For all $n \geq 0$, we have

$$\|f * K_n\|_\infty \leq \|f\|_\infty \|K_n\|_1 = \|f\|_\infty$$

by Lemma B3.10 and (B:6). So for all $n \geq 0$,

$$\|f * K_n - f\|_\infty \leq \|f * K_n\|_\infty + \|f\|_\infty \leq 2\|f\|_\infty.$$

(Idea: we now have control over the ∞-norm of $(f * K_n - f)$. The previous proposition gives us control over its 1-norm. Putting them together will give us control over its 2-norm.)

Hence for all $n \geq 0$, using Lemma A6.8 and Proposition B5.4,

$$
\begin{aligned}
\|f * K_n - f\|_2 &\leq \sqrt{\|f * K_n - f\|_1 \, \|f * K_n - f\|_\infty} \\
&\leq \sqrt{\|f * K_n - f\|_1}\, \sqrt{2\|f\|_\infty} \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}
$$

□

B6 Summing the unsummable

For the lecture of Monday 24 February 2014

Before I explain the title, recall: we're looking for a sequence $K_0, K_1, \ldots$ of trigonometric polynomials that approximate the mythical δ function increasingly well.

How can we find one?

Idea If δ existed, its Fourier coefficients would be given by

$$\hat{\delta}(k) = \int_{\mathbb{T}} \delta(t) e^{-2\pi i k t} \, dt = e^{-2\pi i k 0} = 1$$

for all $k \in \mathbb{Z}$, and so $S_n \delta = \sum_{k=-n}^{n} e_k = D_n$. If also $S_n \delta \to \delta$ as $n \to \infty$, then $D_n \to \delta$. Since D_n is a trigonometric polynomial, we might try $K_n = D_n$.

Problem (D_n) is not a PAD, as noted in Example B5.3(ii). Also, the sequence $(D_n)_{n=0}^{\infty}$ does not converge in *any* of our usual five senses. In other words, $\sum_{k=-\infty}^{\infty} e_k$ is thoroughly unsummable.

So, we want to be able to sum unsummable series. This sounds impossible, but it's not: you just need to be more generous about what 'summable' means. To explain this, it's useful to step back from our particular problem—indeed, step back from Fourier analysis entirely—and think about the general question: how can we sum a divergent series?

Example B6.1 The series

$$\sum_{n=0}^{\infty} (-1)^n = 1 - 1 + 1 - 1 + \cdots$$

does not converge, since the partial sums $S_N = \sum_{n=0}^{N} (-1)^n$ are

$$S_0 = 1, \quad S_1 = 0, \quad S_2 = 1, \quad S_3 = 0, \quad \ldots,$$

and the *sequence* (S_N) does not converge.

Nevertheless, there are non-rigorous ways of evaluating $S = \sum_{n=0}^{\infty} (-1)^n$. For instance:

i. Alice thinks that

$$S = (1 + -1) + (1 + -1) + (1 + -1) + \cdots = 0 + 0 + 0 + \cdots = 0.$$

Bob thinks that

$$S = 1 + (-1 + 1) + (-1 + 1) + \cdots = 1 + 0 + 0 + \cdots = 1.$$

Alice and Bob agree to split the difference, and so conclude that $S = 1/2$.

ii. Consider two copies of S lined up in columns:

$$\begin{aligned}2S &= (1 - 1 + 1 - 1 + \cdots) \\ &\qquad + (1 - 1 + 1 - \cdots) \\ &= 1 + 0 + 0 + 0 + \cdots \\ &= 1.\end{aligned}$$

So $S = 1/2$, the same answer as obtained by Alice and Bob.

iii. Apply the formula for the sum of a geometric series (even though it's invalid, as that formula requires the ratio to have modulus strictly less than 1):

$$S = \frac{1}{1 - (-1)} = 1/2.$$

Again, this is the same answer.

Those three methods were just for fun. Each of them can actually be made respectable, but I won't show you how. Instead, here's a fourth way, which is the one that will matter to us the most.

iv. The sequence of partial sums $(1, 0, 1, 0, \ldots)$ doesn't converge, but it oscillates evenly around $1/2$. The centre of mass is $1/2$, if you like. Since a centre of mass (or centroid) is a mean, this suggests considering the mean of all the partial sums so far.

Put

$$A_N = \frac{1}{N+1}(S_0 + \cdots + S_N).$$

A short calculation shows that

$$A_N = \begin{cases} \frac{1}{2} + \frac{1}{2(N+1)} & \text{if } N \text{ is even,} \\ \frac{1}{2} & \text{if } N \text{ is odd.} \end{cases}$$

So $\lim_{N\to\infty} A_N = 1/2$, as expected.

There is some general terminology for this fourth method.

Definition B6.2 i. Let $(s_n)_{n=0}^\infty$ be a sequence in $\mathbb{C}$. Its **Nth Cesàro mean** ($N \geq 0$) is

$$a_N = \frac{1}{N+1}(s_0 + \cdots + s_N) \in \mathbb{C}.$$

If $a_N \to s$ as $N \to \infty$, we say that s is the **Cesàro limit** of (s_n).

ii. The **Cesàro sum** of a series $\sum_{n=0}^\infty x_n$ is the Cesàro limit of the partial sums $S_N = \sum_{n=0}^N x_n$, if it exists.

We have just seen an example of a series that is Cesàro-summable but not summable in the usual sense. Now we show that the method of Cesàro summation *extends* the method of ordinary summation, in the sense that if the ordinary sum exists then so does the Cesàro sum, and the two sums agree.

Proposition B6.3 *i. Let $(s_n)_{n=0}^\infty$ be a sequence in $\mathbb{C}$, and let $s \in \mathbb{C}$. If s is the limit of (s_n) then s is also the Cesàro limit of (s_n).*

ii. Let $\sum_{n=0}^{\infty} x_n$ be a series in $\mathbb{C}$. If the sum exists, then the Cesàro sum exists and is equal to $\sum x_n$.

Proof We prove (i); then (ii) follows immediately.

Let a_N be the Nth Cesàro mean of (s_n). For all $N \geq 0$,

$$|a_N - s| = \left| \frac{(s_0 - s) + \cdots + (s_N - s)}{N+1} \right| \leq \frac{|s_0 - s| + \cdots + |s_N - s|}{N+1}.$$

(The idea now: we want $|a_N - s|$ to be small. Let L be a large integer. Then $|s_n - s|$ is small when $n \geq L$, and $\frac{|s_0-s|+\cdots+|s_{L-1}-s|}{N+1}$ is small when N is much greater than L, since the numerator does not depend on N.)

Let $\varepsilon > 0$. Since $s_n \to s$, we can choose L such that $|s_n - s| < \varepsilon/2$ for all $n \geq L$. We can then choose an integer

$$M \geq \max\Big\{L,\ \frac{2}{\varepsilon}\big(|s_0 - s| + \cdots + |s_{L-1} - s|\big) - 1\Big\}.$$

For all $N \geq M$,

$$\begin{aligned} |a_N - s| &\leq \frac{|s_0 - s| + \cdots + |s_{L-1} - s|}{N+1} + \frac{|s_L - s| + \cdots + |s_N - s|}{N+1} \\ &< \frac{(\varepsilon/2)(M+1)}{N+1} + \frac{(N-L+1)(\varepsilon/2)}{N+1} \qquad \text{(B:12)} \\ &\leq (\varepsilon/2) + (\varepsilon/2) = \varepsilon, \end{aligned}$$

where in (B:12), we used the definition of M in the first summand and the definition of L in the second. □

Remark B6.4 The partial sums $D_n = \sum_{k=-n}^{n} e_k$ are too wild to form a PAD. The proposition we have just proved suggests that the Cesàro means $\frac{1}{N+1}(D_0 + \cdots + D_N)$ might be tamer. It turns out that they are; indeed, they form a PAD. This will allow us to complete the plan described at the beginning of Section B5, thus proving that in the 2-norm, Fourier series always converge.

B7 The Fejér kernel

For the lecture of Thursday 27 February 2014; part one of two

At the end of the last section, we expressed the hope that although the Dirichlet kernels D_n do not form a PAD, perhaps their Cesàro means

$$\frac{1}{n+1}(D_0 + \cdots + D_n)$$

do. Here we show that this is indeed the case.

Definition B7.1 Let $n \geq 0$. The **Fejér kernel** of order n is

$$F_n = \frac{1}{n+1}(D_0 + \cdots + D_n) \colon \mathbb{T} \to \mathbb{C}.$$

Note that the Fejér kernel, like the Dirichlet kernel, is a trigonometric polynomial.

Way back in Section A1, we abandoned the sine-and-cosine formulation of Fourier series, choosing to work with the more elegant exponential formulation. But it will be useful to have expressions for D_n and F_n in traditional trigonometric form.

Lemma B7.2 *Let $n \geq 0$ and $0 \neq t \in \mathbb{T}$. Then*

$$D_n(t) = \frac{\sin((2n+1)\pi t)}{\sin \pi t}, \qquad F_n(t) = \frac{1}{n+1} \frac{\sin^2((n+1)\pi t)}{\sin^2 \pi t}.$$

Proof We have

$$D_n(t) = \sum_{k=-n}^{n} e_1(t)^k.$$

Since $t \neq 0$, we have $e_1(t) \neq 1$, and we may therefore apply the formula for summing a geometric series. After some routine algebra, we get

$$D_n(t) = \frac{e_1(t)^{n+1/2} - e_1(t)^{-(n+1/2)}}{e_1(t)^{1/2} - e_1(t)^{-1/2}}. \tag{B:13}$$

Noting that $e_1(t)^\alpha = e^{2\pi i \alpha t}$ and applying the formula $e^{i\theta} = \cos\theta + i \sin\theta$, we obtain the result on $D_n(t)$.

We now use a trick: multiply the top and bottom of (B:13) by the bottom. This gives

$$\begin{aligned} D_n(t) &= \frac{\left[e_1(t)^{n+1/2} - e_1(t)^{-(n+1/2)}\right]\left[e_1(t)^{1/2} - e_1(t)^{-1/2}\right]}{\left[e_1(t)^{1/2} - e_1(t)^{-1/2}\right]^2} \\ &= \frac{\left[e_1(t)^{n+1} + e_1(t)^{-(n+1)}\right] - \left[e_1(t)^n + e_1(t)^{-n}\right]}{\left[e_1(t)^{1/2} - e_1(t)^{-1/2}\right]^2}. \end{aligned}$$

So for any $N \geq 0$, the sum $\sum_{n=0}^{N} D_n(t)$ telescopes, giving

$$\begin{aligned}
\sum_{n=0}^{N} D_n(t) &= \frac{\left[e_1(t)^{N+1} + e_1(t)^{-(N+1)}\right] - \left[e_1(t)^0 + e_1(t)^{-0}\right]}{\left[e_1(t)^{1/2} - e_1(t)^{-1/2}\right]^2} \\
&= \frac{\left[e_1(t)^{(N+1)/2} - e_1(t)^{-(N+1)/2}\right]^2}{\left[e_1(t)^{1/2} - e_1(t)^{-1/2}\right]^2} \\
&= \frac{[2i\sin((N+1)\pi t)]^2}{[2i\sin(\pi t)]^2},
\end{aligned}$$

hence the result on $F_N(t)$. □

This explicit formula helps us to prove our main result on the Fejér kernel.

Proposition B7.3 *$(F_n)_{n=0}^{\infty}$ is a PAD. In particular, there is a PAD consisting of trigonometric polynomials.*

Proof PAD1: By Lemma B7.2, $F_N(t) \geq 0$ for all N and t.

PAD2: First, $\int_{\mathbb{T}} D_n(t)\,dt = 1$ for all n, since

$$\int_{\mathbb{T}} D_n(t)\,dt = \langle D_n, e_0 \rangle = \sum_{k=-n}^{n} \langle e_k, e_0 \rangle = 1.$$

Hence

$$\begin{aligned}
\int_{\mathbb{T}} F_N(t)\,dt &= \frac{1}{N+1}\left(\int_T D_0(t)\,dt + \cdots + \int_{\mathbb{T}} D_N(t)\,dt\right) \\
&= \frac{1}{N+1}(1 + \cdots + 1) = 1.
\end{aligned}$$

PAD3: Let $\delta \in (0, 1/2)$. We must prove that

$$\lim_{n\to\infty} \int_{\delta<|t|<1/2} F_N(t)\,dt = 0.$$

We use the fact that $\sin\theta \geq \frac{2}{\pi}\theta$ for all $\theta \in [0, \pi/2]$. (This can be proved using convexity; see Figure B.5.) Now

$$\begin{aligned}
\int_{\delta<|t|\leq 1/2} F_N(t)\,dt &= \frac{2}{N+1}\int_{\delta}^{1/2} \frac{\sin^2((N+1)\pi t)}{\sin^2 \pi t}\,dt && \text{(B:14)} \\
&\leq \frac{2}{N+1}\int_{\delta}^{1/2} \frac{1}{(2t)^2}\,dt && \text{(B:15)} \\
&= \frac{1}{2(N+1)}\left(\frac{1}{\delta} - 2\right) && \text{(B:16)} \\
&\to 0 \text{ as } N \to \infty. && \text{(B:17)}
\end{aligned}$$

Here (B:14) follows from Lemma B7.2 and F_N being even, (B:15) is because $\sin^2((n+1)\pi t) \leq 1$ and $|\sin \pi t| \geq \frac{2}{\pi} \cdot \pi t$, and (B:16) is a routine calculation. □

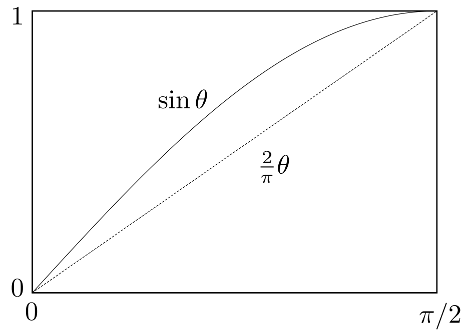


Figure B.5: $\sin\theta \geq \frac{2}{\pi}\theta$ for all $\theta \in [0, \pi/2]$.

B8 The main theorem

For the lecture of Thursday 27 February 2014; part two of two

We can now prove the main theorem of Part B. It states that Fourier's idea works perfectly when we use the 2- or 1-norm.

Theorem B8.1 *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then $S_n f \to f$ in both $\|\cdot\|_2$ and $\|\cdot\|_1$.*

Proof For all n, the Fejér kernel F_n is a trigonometric polynomial, so $f * F_n$ is a trigonometric polynomial (Lemma B3.4). Also, (F_n) is a PAD (Proposition B7.3), so $f * F_n \to f$ in $\|\cdot\|_2$ (Proposition B5.5). Hence f is the limit in $\|\cdot\|_2$ of a sequence of trigonometric polynomials. So by Lemma B1.3, $S_n f \to f$ in $\|\cdot\|_2$. Finally, by Lemma A4.10, $S_n f \to f$ in $\|\cdot\|_1$. $\square$

Corollary B8.2 *The set {trigonometric polynomials} is dense in {integrable functions $\mathbb{T} \to \mathbb{C}$} with respect to both $\|\cdot\|_2$ and $\|\cdot\|_1$.*

Proof This follows from Theorem B8.1, noting that the Fourier partial sums $S_n f$ are trigonometric polynomials. $\square$

In other words, if f is a point in the space of all integrable functions on $\mathbb{T}$, then there are trigonometric polynomials arbitrarily close to f. (Compare Proposition A9.4 on step functions.)

Theorem B8.3 (Parseval) *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then*

$$\|f\|_2 = \sqrt{\sum_{k=-\infty}^{\infty} \left|\hat{f}(k)\right|^2}.$$

Proof $S_n f \to f$ in $\|\cdot\|_2$, so $\|S_n f\|_2 \to \|f\|_2$ by Lemma A4.7, so $\|S_n f\|_2^2 \to \|f\|_2^2$. But

$$\|S_n f\|_2^2 = \sum_{k=-n}^{n} \left|\hat{f}(k)\right|^2$$

by Lemma B1.1, so the result follows. $\square$

In a suitable sense, Parseval's theorem says that the map $f \mapsto \hat{f}$ is an isometry (distance-preserving). It is also angle-preserving:

Corollary B8.4 *Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions. Then*

$$\int_{\mathbb{T}} f(x)\overline{g(x)}\,dx = \sum_{k=-\infty}^{\infty} \hat{f}(k)\overline{\hat{g}(k)}.$$

Proof Sheet 3, q.5. $\square$

Finally, the map $f \mapsto \hat{f}$ is, essentially, injective:

Corollary B8.5 *Let $f, g\colon \mathbb{T} \to \mathbb{C}$ be integrable functions such that $\hat{f}(k) = \hat{g}(k)$ for all $k \in \mathbb{Z}$. Then:*

i. $f = g$ a.e.;

ii. $f(x) = g(x)$ for all $x \in \mathbb{T}$ such that f and g are both continuous at x;

iii. $f = g$ if f and g are both continuous.

Proof We have $\widehat{f-g}(k) = 0$ for all $k \in \mathbb{Z}$ (using Lemma A7.7(i)), so $\|f - g\|_2 = 0$ by Parseval's theorem. All three parts now follow from Lemma A4.5. $\square$

This result does not mention the 2-norm (or indeed any norm) in its statement, although it does use the 2-norm in its proof. It answers a very fundamental question about Fourier series, telling us:

Different functions have different Fourier series.

Here 'different functions' must be understood as 'functions that are not almost everywhere equal', since it's a basic fact that if $f = g$ a.e. then f and g have the same Fourier coefficients; see A3.13.

This encourages us to believe in Fantasy A8.7.

Chapter C

Uniform and pointwise convergence of Fourier series

In Remark A4.11, we met this diagram:

$$\begin{tikzcd}[column sep=small]
 & & \|\cdot\|_\infty \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & & \\
 & \|\cdot\|_2 \arrow[dl, Rightarrow] & & \text{pointwise} \arrow[dr, Rightarrow] & \\
\|\cdot\|_1 & & & & \text{a.e.}
\end{tikzcd}$$

Part B covered the lower-left region of the diagram: convergence of Fourier series in the 2- and 1-norms. In Part C, we look at the upper-right region: uniform and pointwise convergence. (We will never look seriously at almost everywhere convergence.)

C1 Warm-up

For the lecture of Monday 3 March 2014; part one of two

In Part B, we showed:

for all integrable $f\colon \mathbb{T} \to \mathbb{C}$, $S_n f \to f$ in $\|\cdot\|_2$ and $\|\cdot\|_1$.

Is it also true in $\|\cdot\|_\infty$?

No, since if $S_n f \to f$ in $\|\cdot\|_\infty$ then f is continuous (by Fact A4.8), whereas not every integrable function is continuous.

But this leaves open the possibility that:

for all *continuous* $f\colon \mathbb{T} \to \mathbb{C}$, $S_n f \to f$ in $\|\cdot\|_\infty$.

However, this is not true either. By du Bois–Reymond's example (Theorem A2.2), there is some continuous $f\colon \mathbb{T} \to \mathbb{C}$ such that $S_n f$ does not even converge *pointwise* to f, let alone uniformly.

So if we're hoping for all Fourier series to converge in either the uniform or pointwise sense, we'll be disappointed. But perhaps it's true if we dilute our hopes, asking only for convergence almost everywhere. In other words, we might hope that:

$$\text{for all continuous } f\colon \mathbb{T} \to \mathbb{C},\ \ S_n f \to f \text{ almost everywhere.}$$

This *is* true, by Carleson's theorem (Theorem A2.6). But as previously mentioned, all known proofs are much too hard for this course.

However, there are some easier results available. For example, we will be able to prove the theorem of Dirichlet that if f is continuously differentiable, then $S_n f \to f$ uniformly (hence in all five senses). We will also look at an application that seems to have nothing to do with Fourier theory: the so-called equidistribution theorem of Weyl, concerning the statistical behaviour of multiples of irrational numbers.

We begin by collecting a few basic facts.

As described in Fantasy A8.7, the theory of Fourier series is about the back-and-forth between functions on $\mathbb{T}$ and doubly infinite sequences. So far, we have concentrated on the passage in one direction, starting with a function f on $\mathbb{T}$ and deriving the double sequence $(\hat{f}(k))_{k\in\mathbb{Z}}$. Let us now consider the opposite direction. Write

$$\mathbb{C}^{\mathbb{Z}} = \big\{\text{double sequences } c = (c_k)_{k=-\infty}^{\infty} \text{ in } \mathbb{C}\big\}.$$

Lemma C1.1 *Let $c \in \mathbb{C}^{\mathbb{Z}}$. Then*

$$\sum_{k=-\infty}^{\infty} |c_k| < \infty \ \Rightarrow \ \sum_{k=-\infty}^{\infty} |c_k|^2 < \infty \ \Rightarrow \ \sup_{k\in\mathbb{Z}} |c_k| < \infty.$$

(This should remind you of the situation with convergence in the 1-, 2- and ∞-norms—but it's the other way round. Notice that squaring a small positive number *decreases* it.)

Proof For the first implication, suppose that $\sum_{k=-\infty}^{\infty} |c_k| < \infty$. Then certainly the set $\{|c_k| : k \in \mathbb{Z}\}$ is bounded; put $\|c\|_\infty = \sup_{k\in\mathbb{Z}} |c_k|$. For all $n \geq 0$, we have

$$\sum_{k=-n}^{n} |c_k|^2 \leq \|c\|_\infty \sum_{k=-n}^{n} |c_k| \leq \|c\|_\infty \sum_{k=-\infty}^{\infty} |c_k|,$$

so $\sum_{k=-\infty}^{\infty} |c_k|^2 < \infty$.

For the second implication, if $\sum_{k=-\infty}^{\infty} |c_k|^2 < \infty$ then $\{|c_k|^2 : k \in \mathbb{Z}\}$ is bounded, so $\{|c_k| : k \in \mathbb{Z}\}$ is bounded too. □

Remark C1.2 No further implications hold. To see that the converse of the first implication fails, consider

$$c_k = \begin{cases} 1/k & \text{if } k \geq 1 \\ 0 & \text{if } k \leq 0. \end{cases}$$

To see that the converse of the second implication fails, take $c_k = 1$ for all k.

Here is another basic fact. It is about convergence in the 1-norm, but will be useful for arguments about uniform and pointwise convergence.

Lemma C1.3 *Let $c \in \mathbb{C}^{\mathbb{Z}}$ and let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Suppose that $\sum_{k=-n}^{n} c_k e_k \to f$ in $\| \cdot \|_1$ as $n \to \infty$. Then $c_k = \hat{f}(k)$ for all $k \in \mathbb{Z}$.*

Proof For each $n \geq 0$, write $g_n = \sum_{k=-n}^{n} c_k e_k$. Let $\ell \in \mathbb{Z}$. By Lemma A7.7(ii), $\widehat{g_n}(\ell) \to \hat{f}(\ell)$ as $n \to \infty$. But by Lemma A8.3, $\widehat{g_n}(\ell) = c_\ell$ whenever $n \geq |\ell|$, so $c_\ell = \hat{f}(\ell)$, as required. $\square$

We can now answer a subtle question that has been present from the beginning. Can it happen that the Fourier series of f converges, *but not to f*? For convergence in the uniform sense, no.

Lemma C1.4 *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Suppose that the sequence $(S_n f)_{n=0}^{\infty}$ converges uniformly (not necessarily to f). Then:*

i. $(S_n f)(x) \to f(x)$ for almost all $x \in \mathbb{T}$;

ii. $(S_n f)(x) \to f(x)$ for all $x \in \mathbb{T}$ such that f is continuous at x;

iii. $S_n f \to f$ uniformly if f is continuous.

Part (iii) says that for a continuous function f, if $S_n f$ converges uniformly to *something*, that something must be f.

Proof Let g be the uniform limit of the sequence $(S_n f)$. By Fact A4.8, g is continuous. Certainly $S_n f \to g$ in $\| \cdot \|_1$, so $\hat{f}(k) = \hat{g}(k)$ for all $k \in \mathbb{Z}$ by Lemma C1.3 (taking 'c_k' to be $\hat{f}(k)$ and 'f' to be g). The result now follows from Corollary B8.5. $\square$

C2 What if the Fourier coefficients are absolutely summable?

For the lecture of Monday 3 March 2014; part two of two

We saw in Proposition B1.2 that $\sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2 < \infty$ for any integrable function $f\colon \mathbb{T} \to \mathbb{C}$. However, f may or may not have the stronger property that $\sum_{k=-\infty}^{\infty} |\hat{f}(k)| < \infty$. (To see why this is stronger, recall Lemma C1.1.)

In this section, we will see that *if* $\sum_{k=-\infty}^{\infty} |\hat{f}(k)| < \infty$ then the Fourier series of f behaves well, in a sense that will be made precise. In the next section, we will see that many common functions do have this property.

We begin with a result on double sequences.

Proposition C2.1 *Let $c \in \mathbb{C}^{\mathbb{Z}}$. Suppose that $\sum_{k=-\infty}^{\infty} |c_k| < \infty$. Then the sequence $\left(\sum_{k=-n}^{n} c_k e_k\right)_{n=0}^{\infty}$ converges uniformly to a continuous function g. Moreover, $\hat{g}(k) = c_k$ for all $k \in \mathbb{Z}$.*

Proof For $n \geq 0$, write $s_n = \sum_{k=-n}^{n} c_k e_k$.

First we show that the sequence (s_n) converges pointwise. Indeed, let $x \in \mathbb{T}$. Given $\varepsilon > 0$, we can choose N such that $\sum_{k\colon |k| \geq N} |c_k| < \varepsilon$; then for all $m \geq n \geq N$,

$$|s_m(x) - s_n(x)| = \left| \sum_{k\colon n<|k|\leq m} c_k e_k(x) \right| \leq \sum_{k\colon n<|k|\leq m} |c_k| < \varepsilon.$$

So the sequence $(s_n(x))_{n=0}^{\infty}$ is Cauchy, and therefore convergent, as claimed.

Next, for each $x \in \mathbb{T}$, put

$$g(x) = \lim_{n\to\infty} s_n(x) = \sum_{k=-\infty}^{\infty} c_k e_k(x).$$

This defines a function $g\colon \mathbb{T} \to \mathbb{C}$. Now (s_n) converges *uniformly* to g, since

$$\|s_n - g\|_\infty = \sup_{x\in\mathbb{T}} \left| \sum_{k\colon |k|>n} c_k e_k(x) \right| \leq \sum_{k\colon |k|>n} |c_k| \to 0$$

as $n \to \infty$. By Fact A4.8, g is therefore continuous, and by Lemma C1.3, $\hat{g}(k) = c_k$ for all $k \in \mathbb{Z}$. $\square$

We now come to the most important result of Part C so far.

Theorem C2.2 *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Suppose that $\sum_{k=-\infty}^{\infty} |\hat{f}(k)| < \infty$. Then:*

i. $(S_n f)(x) \to f(x)$ as $n \to \infty$ for all $x \in \mathbb{T}$ such that f is continuous at x;

ii. $S_n f \to f$ uniformly if f is continuous.

Proof By Proposition C2.1 (taking c_k to be $\hat{f}(k)$), the Fourier partial sums $S_n f$ converge uniformly to some continuous function $g\colon \mathbb{T} \to \mathbb{C}$, and $\hat{g}(k) = \hat{f}(k)$ for all $k \in \mathbb{Z}$. The result follows from Corollary B8.5. $\square$

C3 Continuously differentiable functions

For the lecture of Thursday 6 March 2014; part one of two

We have just shown that if $\sum_{k=-\infty}^{\infty} |\hat{f}(k)| < \infty$ then the Fourier series of f behaves well (Theorem C2.2). If you had a *particular* function f in front of you, you might be able to calculate its Fourier coefficients, you might be able to calculate the sum of their absolute values, and that sum might be finite. In that case, you could conclude that the Fourier series of that particular f was well-behaved. But are there *general* conditions on f guaranteeing that $\sum_{k=-\infty}^{\infty} |\hat{f}(k)| < \infty$?

The answer turns out to be yes. But beware that not *all* integrable functions $f\colon \mathbb{T} \to \mathbb{C}$ satisfy $\sum |\hat{f}(k)| < \infty$. In fact, there exist quite 'nice' functions f with $\sum |\hat{f}(k)| = \infty$:

Example C3.1 Define $f\colon \mathbb{T} \to \mathbb{C}$ by

$$f(x) = \sum_{n \geq 2} \frac{\sin(2\pi n x)}{n \log n}.$$

One can show that $S_n f \to f$ uniformly, and so f is continuous. One can also show that $\sum |\hat{f}(k)| = \infty$. (The proof of both these statements is omitted and non-examinable.) So a continuous function f need not satisfy $\sum |\hat{f}(k)| < \infty$.

However, we will see that if f is smooth enough, the Fourier coefficients decay fast enough that $\sum |\hat{f}(k)|$ converges.

Definition C3.2 Let $n \geq 0$. We write $C^n(\mathbb{T})$ for the set of functions $\mathbb{T} \to \mathbb{C}$ whose corresponding 1-periodic function $g\colon \mathbb{R} \to \mathbb{C}$ is n times continuously differentiable (that is, $g^{(n)}$ exists and is continuous). We also write

$$C(\mathbb{T}) = C^0(\mathbb{T}) = \{\text{continuous functions } \mathbb{T} \to \mathbb{C}\}.$$

For example, a function $\mathbb{T} \to \mathbb{C}$ belongs to the set $C^1(\mathbb{T})$ if and only if the corresponding 1-periodic function $\mathbb{R} \to \mathbb{C}$ is continuously differentiable.

Lemma C3.3 *Let $n \geq 0$, $f \in C^n(\mathbb{T})$, and $k \in \mathbb{Z}$. Then*

$$\widehat{f^{(n)}}(k) = (2\pi i k)^n \hat{f}(k).$$

Proof Sheet 3, q.6(ii). □

In other words, differentiating f amounts to multiplying its kth Fourier coefficient by $2\pi i k$.

We can now prove the 1829 theorem of Dirichlet, one of the early landmarks of Fourier analysis. In fact, what we prove is stronger than the version stated earlier (Theorem A2.1): there, we only asserted pointwise convergence, but here we prove uniform convergence.

Theorem C3.4 (Dirichlet) *Let $f \in C^1(\mathbb{T})$. Then $\sum_{k=-\infty}^{\infty} |\hat{f}(k)| < \infty$ and $S_n f \to f$ uniformly.*

Since uniform convergence is the strongest kind of convergence, this tells us that the Fourier series of a continuously differentiable function converges in all five senses. Thus, it behaves as well as any function could.

Proof For $k \neq 0$, we have

$$\begin{aligned} |\hat{f}(k)| &= |\widehat{f'}(k)| \cdot \frac{1}{2\pi|k|} && \text{(C:1)} \\ &\leq \frac{1}{2}\Big(|\widehat{f'}(k)|^2 + \frac{1}{(2\pi|k|)^2}\Big) && \text{(C:2)} \\ &= \frac{|\widehat{f'}(k)|^2}{2} + \frac{1}{8\pi^2 k^2}, && \text{(C:3)} \end{aligned}$$

where in (C:1) we used Lemma C3.3 and in (C:2) we used the fact that $uv \leq \frac{1}{2}(u^2 + v^2)$ for all $u, v \in \mathbb{R}$ (proof: exercise). So to show that $\sum_{k=-\infty}^{\infty} |\hat{f}(k)|$ is finite, it suffices to show that the sum of (C:3) over all $k \in \mathbb{Z}$ is finite. First, f' is continuous and so integrable; hence $\sum_{k=-\infty}^{\infty} |\widehat{f'}(k)|^2 < \infty$ by Proposition B1.2. Second,

$$\sum_{0 \neq k \in \mathbb{Z}} \frac{1}{k^2} = 2 \sum_{k \geq 1} \frac{1}{k^2} < \infty.$$

Hence $\sum_{k=-\infty}^{\infty} |\hat{f}(k)| < \infty$, giving $S_n f \to f$ uniformly by Theorem C2.2. $\square$

Remark C3.5 We have just shown that

$$f \in C^1(\mathbb{T}) \;\Rightarrow\; \sum |\hat{f}(k)| < \infty.$$

But the converse fails, even for continuous functions. Here is an example due to Weierstrass:

$$f(x) = \sum_{n=1}^{\infty} 2^{-n} \cos(2^n \pi x).$$

It can be shown (non-examinably) that $\sum |\hat{f}(k)| < \infty$ and that f is continuous. But it can also be shown that f is not differentiable *anywhere*, let alone *continuously* differentiable *everywhere*.

C4 Fejér's theorem

For the lecture of Thursday 6 March 2014; part two of two

Du Bois–Reymond's example (Theorem A2.2) tells us that for a continuous function $f\colon \mathbb{T} \to \mathbb{C}$ and a point $x \in \mathbb{T}$, the sequence $\big((S_n f)(x)\big)_{n=0}^{\infty}$ need not have a limit. However, Fejér showed that it always has a *Cesàro* limit, namely, $f(x)$. His result is the centrepiece of this section.

In Part B, we proved that $S_n f \to f$ in $\|\cdot\|_2$ by proving that $f * F_n \to f$ in $\|\cdot\|_2$. We did *that* by proving that (i) (F_n) is a PAD, and (ii) $f * K_n \to f$ in $\|\cdot\|_2$ for any PAD (K_n). We adopt a similar strategy here, but with pointwise and uniform convergence in place of mean square convergence.

For integrable $f\colon \mathbb{T} \to \mathbb{C}$, we have

$$\begin{aligned} f * F_n &= f * \frac{1}{n+1}(D_0 + \cdots + D_n) \\ &= \frac{1}{n+1}(f * D_0 + \cdots + f * D_n) \\ &= \frac{1}{n+1}(S_0 f + \cdots + S_n f). \end{aligned}$$

This is nothing but the nth Cesàro mean of the sequence $S_0 f, S_1 f, \ldots$. We will write

$$A_n f = \frac{1}{n+1}(S_0 f + \cdots + S_n f) = f * F_n.$$

To say that $f(x)$ is the Cesàro limit of the sequence $((S_n f)(x))$ is to say that $(A_n f)(x) \to f(x)$ as $n \to \infty$, or equivalently that $(f * F_n)(x) \to f(x)$ as $n \to \infty$.

Here is a summary of the notation and terminology:

	Ordinary convergence	Cesàro convergence
nth term in sequence	$S_n f$	$A_n f$
nth kernel	D_n (Dirichlet)	F_n (Fejér)

To prove Fejér's theorem, we will need to know a little more about PADs. We showed in Section B5 that for any PAD (K_n) and integrable function f, we have $f * K_n \to f$ in both $\|\cdot\|_2$ and $\|\cdot\|_1$. That result cannot be strengthened to convergence in $\|\cdot\|_\infty$ for *arbitrary* integrable functions f, since each $f * K_n$ is continuous and a uniform limit of continuous functions is continuous. So the best we could hope for is that *if f is continuous* then $f * K_n \to f$ in $\|\cdot\|_\infty$. We might also hope to be able to prove a result about convergence at particular points. Both these hopes are fulfilled:

Proposition C4.1 *Let (K_n) be a PAD and $f\colon \mathbb{T} \to \mathbb{C}$ an integrable function. Then:*

i. $(f * K_n)(x) \to f(x)$ as $n \to \infty$ for all $x \in \mathbb{T}$ such that f is continuous at x;

ii. $f * K_n \to f$ uniformly if f is continuous.

The proof is similar to the proof of Proposition B5.4, the corresponding result for the 1-norm.

Proof For (i), let $x \in \mathbb{T}$. In the proof of Proposition B5.4, equation (B:7) stated that for all $n \geq 0$ and $\delta \in (0, 1/2)$,

$$|(f*K_n)(x)-f(x)| \leq \int_{|t|<\delta} |f(x-t)-f(x)|K_n(t)\,dt+2\|f\|_\infty \int_{\delta<|t|\leq 1/2} K_n(t)\,dt.$$

Suppose that f is continuous at x. Let $\varepsilon > 0$. Choose $\delta \in (0, 1/2)$ such that $|f(x-t) - f(x)| < \varepsilon/2$ for all $t \in (-\delta, \delta)$. By PAD3, we can also choose N such that for all $n \geq N$,

$$\int_{\delta<|t|\leq 1/2} K_n(t)\,dt < \frac{\varepsilon}{4\|f\|_\infty}.$$

Then for all $n \geq N$,

$$\begin{aligned} |(f * K_n)(x) - f(x)| &\leq \frac{\varepsilon}{2} \int_{|t|<\delta} K_n(t)\,dt + 2\|f\|_\infty \cdot \frac{\varepsilon}{4\|f\|_\infty} \\ &\leq \frac{\varepsilon}{2} \int_{-1/2}^{1/2} K_n(t)\,dt + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

as required.

For (ii), suppose that f is continuous. Let $\varepsilon > 0$. Every continuous map from a compact metric space to a metric space is uniformly continuous, and $\mathbb{T}$ is compact, so f is uniformly continuous. Hence we can choose $\delta \in (0, 1/2)$ such that for all $x \in \mathbb{T}$ and $t \in (-\delta, \delta)$, we have $|f(x-t) - f(x)| < \varepsilon/2$.

(In case you have not previously encountered this theorem about metric spaces, an alternative argument is available. Here we use instead the theorem that any continuous function on a closed bounded interval is uniformly continuous. Since f is continuous on $[-1, 1]$, it is uniformly continuous there. So by interpreting $\mathbb{T}$ as $[-1/2, 1/2)$, we can choose δ as in the last paragraph.)

By PAD3, we can choose N as in the proof of (i). For all $x \in \mathbb{T}$, for all $n \geq N$, we calculate that $|(f * K_n)(x) - f(x)| < \varepsilon$, again as in the proof of (i). So $\|f * K_n - f\|_\infty \leq \varepsilon$ for all $n \geq N$, as required. □

Theorem C4.2 (Fejér) *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then:*

i. $(A_n f)(x) \to f(x)$ as $n \to \infty$ for all $x \in \mathbb{T}$ such that f is continuous at x;

ii. $A_n f \to f$ uniformly if f is continuous.

Proof (F_n) is a PAD and $A_n f = f * F_n$, so this follows from the last proposition. □

Looking back at du Bois–Reymond's example (Theorem A2.2), we see that the sequence $(A_n f)$ is more likely to converge than the sequence $(S_n f)$.

Remark C4.3 If $f, g\colon \mathbb{T} \to \mathbb{C}$ with $\hat{f}(k) = \hat{g}(k)$ for all k, then $A_n f = A_n g$ for all n, so by Fejér's theorem, $f(x) = g(x)$ for all $x \in \mathbb{T}$ such that f and g are both continuous at x. This gives an alternative proof of Corollary B8.5(ii), (iii). But unlike that corollary, Fejér's theorem gives an explicit way of *reconstructing* f from its Fourier coefficients (at least for continuous f): from knowing the Fourier coefficients, we can compute $A_n f$ for each n, and we then obtain $f(x)$ as $\lim_{n\to\infty}(A_n f)(x)$.

Corollary C4.4 *The set {trigonometric polynomials} is dense in $C(\mathbb{T})$ with respect to $\|\cdot\|_\infty$.*

Proof Let $f \in C(\mathbb{T})$. For each $n \geq 0$, the function $A_n f$ is a trigonometric polynomial, since $S_0 f, \ldots, S_n f$ are. The result follows from part (ii) of Fejér's theorem. $\square$

By Lemma A4.10(i), density with respect to $\|\cdot\|_\infty$ is the strongest type of density (that is, stronger than density with respect to $\|\cdot\|_2$ or $\|\cdot\|_1$). Corollary C4.4 can be compared to Corollary B8.2, which stated that the set of trigonometric polynomials is dense in the set of *integrable* functions with respect to $\|\cdot\|_2$ (hence $\|\cdot\|_1$).

(Faced with this, it is natural to ask whether Corollary B8.2 can be improved to density with respect to $\|\cdot\|_\infty$. The answer is no, again because a uniform limit of continuous functions is continuous.)

Corollary C4.5 (Weierstrass approximation theorem) *Let $I \subseteq \mathbb{R}$ be a closed bounded interval and $f\colon I \to \mathbb{C}$ a continuous function. Then for all $\varepsilon > 0$, there exists a polynomial p over $\mathbb{C}$ such that $\sup_{x \in I} |f(x) - p(x)| \leq \varepsilon$.*

Proof Sheet 4. $\square$

This result is important in many ways. For example, it is a basic fact in numerical analysis, where we might wish to approximate an arbitrary continuous function by a polynomial. However, we will not need it for anything in the rest of this course.

C5 Differentiable functions and the Riemann localization principle

For the lecture of Monday 10 March

This section contains two major results. First, there is a convergence theorem for differentiable functions, to add to our existing convergence theorem for continuously differentiable functions (Dirichlet's Theorem C3.4). Then, we meet the deep and shocking 'localization principle' of Riemann.

Dirichlet's theorem states that if f is *continuously* differentiable then $S_n f \to f$ *uniformly*. We will weaken both the hypothesis and the conclusion, proving that if f is differentiable (but not necessarily continuously so) then $S_n f \to f$ pointwise (but not necessarily uniformly).

To prove this, we need another fact about integration.

Fact C5.1 *Let $I \subseteq \mathbb{R}$ be a bounded interval and $x \in I$. Let $g\colon I \to \mathbb{C}$ be a bounded function such that for all $\delta > 0$, the restriction $g|_{I\setminus(x-\delta,x+\delta)}$ is integrable. Then g is integrable.*

The proof is omitted and non-examinable, but can be found in the appendix of the book *Fourier Analysis* by Stein and Shakarchi.

Theorem C5.2 *Let $f\colon \mathbb{T} \to \mathbb{C}$ be an integrable function. Then:*

i. $(S_n f)(x) \to f(x)$ for all $x \in \mathbb{T}$ such that f is differentiable at x;

ii. $S_n f \to f$ pointwise if f is differentiable.

Proof We just have to prove (i), since this immediately implies (ii). Take $x \in \mathbb{T}$ such that f is differentiable at x.

For all $n \geq 0$,

$$\begin{aligned} f(x) - (S_n f)(x) &= \int_{\mathbb{T}} D_n(t) f(x)\, dt - \int_{\mathbb{T}} D_n(t) f(x-t)\, dt && \text{(C:4)}\\ &= \int_{\mathbb{T}} \frac{e_{n+1}(t) - e_{-n}(t)}{e_1(t) - 1} (f(x) - f(x-t))\, dt && \text{(C:5)}\\ &= \int_{\mathbb{T}} (e_{n+1}(t) - e_{-n}(t)) g(t)\, dt, && \text{(C:6)} \end{aligned}$$

where in (C:4) we used the fact that $\int_{\mathbb{T}} D_n(t)\, dt = 1$, in (C:5) we used equation (B:13) (from the proof of Lemma B7.2), and in (C:6) we put

$$g(t) = \frac{f(x) - f(x-t)}{e_1(t) - 1}$$

for $0 < |t| \leq 1/2$. If g is integrable then

$$f(x) - (S_n f)(x) = \hat{g}(-(n+1)) - \hat{g}(n) \to 0 - 0 = 0$$

as $n \to \infty$, by the Riemann–Lebesgue lemma (Proposition B1.2), completing the proof. So it is enough to prove that g is integrable. To do this, we will need the assumption that f is differentiable at x, which we have not used yet.

We prove that g is integrable using Fact C5.1.

First we show that for each $\delta > 0$, the function $g|_{[-1/2,1/2)\setminus(-\delta,\delta)}$ is integrable. Let $\delta > 0$. By Lemma A3.7, it is enough to show that the restrictions of the functions

$$t \mapsto 1/(e_1(t) - 1), \qquad t \mapsto f(x) - f(x - t)$$

to $[-1/2, 1/2) \setminus (-\delta, \delta)$ are both integrable. And indeed, the first restricted function is integrable because it is continuous and bounded, and the second is integrable because f is.

Next we show that g is bounded. For $0 < |t| \leq 1/2$, we have

$$g(t) = \frac{f(x) - f(x - t)}{t} \cdot \frac{t}{e_1(t) - 1} \to f'(x) \cdot \frac{1}{2\pi i} \quad \text{as } t \to 0,$$

using l'Hôpital's rule in the second factor. Define $g(0) = f'(x)/2\pi i$; then g is continuous at 0, so there exists $\eta > 0$ such that $g|_{(-\eta,\eta)}$ is bounded. But also, $g|_{[-1/2,1/2)\setminus(-\eta,\eta)}$ is bounded (since it is integrable). Hence g itself is bounded.

It follows from Fact C5.1 that g is integrable, as required. □

It once seems to have been believed that although a continuous function need not be differentiable everywhere, it must be differentiable *somewhere*. If this were the case, Theorem C5.2 would tell us that the Fourier series of a continuous function f must converge to $f(x)$ for at least one value of x. However, Weierstrass's example of a continuous function that is nowhere differentiable (Remark C3.5) shows that this strategy will not work.

We now use Theorem C5.2 to deduce the second major result of this section, the Riemann localization principle. A good way to prepare for this is to look back at Section A1 (the algebraist's dream), and especially the two points labelled '7' there.

Remark C5.3 The Taylor series

$$(Tf)(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

of an infinitely differentiable function f is 'determined locally' at 0. That is, if $f, g \colon \mathbb{R} \to \mathbb{C}$ with $f|_{(-\delta,\delta)} = g|_{(-\delta,\delta)}$ for some $\delta > 0$, then $Tf = Tg$. In contrast, the Fourier series of $f \colon \mathbb{C} \to \mathbb{C}$ is *not* determined locally at 0 (or indeed at any other point), since each coefficient

$$\hat{f}(k) = \int_{\mathbb{T}} f(t) e^{-2\pi i k t} \, dt$$

involves all of f. Thus, given $f, g \colon \mathbb{T} \to \mathbb{C}$ with $f|_{(-\delta,\delta)} = g|_{(-\delta,\delta)}$ for some $\delta > 0$, we can *not* deduce that $Sf = Sg$.

The next result therefore comes as a surprise.

Corollary C5.4 (Riemann localization principle) *Let $f, g \colon \mathbb{T} \to \mathbb{C}$ be integrable functions and $x \in \mathbb{T}$. Suppose there exists $\delta > 0$ such that $f|_{(x-\delta,x+\delta)} = g|_{(x-\delta,x+\delta)}$. Then*

$$(S_n f)(x) \to f(x) \text{ as } n \to \infty \iff (S_n g)(x) \to g(x) \text{ as } n \to \infty.$$

In other words, whether or not the Fourier series of f at x converges to $f(x)$ depends only on the behaviour of f *near* x.

Proof The integrable function $f - g \colon \mathbb{T} \to \mathbb{C}$ has constant value 0 on some neighbourhood of x, so is differentiable at x. Hence by Theorem C5.2,

$$(S_n(f-g))(x) \to (f-g)(x) \text{ as } n \to \infty.$$

Equivalently,

$$(S_n f)(x) - (S_n g)(x) \to 0 \text{ as } n \to \infty.$$

The result follows. □

The theme of 'local versus global' becomes increasingly visible at this level of mathematics. You may have met it in differential geometry; for example, a surface is locally like $\mathbb{R}^2$, but globally may be quite unlike it. In number theory, there is the local-to-global principle of Hasse, which has to do with p-adic numbers and is helpful in determining the solvability of polynomial equations over the integers. The interplay between local and global is captured formally by the notion of sheaf, which is especially prominent in algebraic geometry.

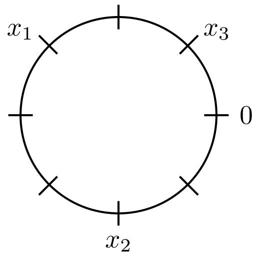


Figure C.1: The beginning of the sequence $(x_n)_{n=1}^\infty$, where $x_n = \left\langle \frac{3}{8}n \right\rangle$.

C6 Weyl's equidistribution theorem

For the lecture of Thursday 13 March

This section is an application of our theory. We will use what we have learned to solve a problem that appears to have nothing to do with Fourier analysis.

Question:

> Which sequences $x_1, x_2, \ldots$ in $[0,1)$ are 'evenly spread'?

More fundamentally:

> What does it *mean* for a sequence $x_1, x_2, \ldots$ in $[0,1)$ to be 'evenly spread'?

We answer the second question immediately. Write $\#S$ for the cardinality of a finite set S. (You might be more used to $|S|$, but that could be confused with the notation $|I|$ for the length of an interval I.)

Definition C6.1 A sequence $(x_n)_{n=1}^\infty$ in $[0,1)$ is **equidistributed** if for all intervals $I \subseteq [0,1)$,

$$\frac{1}{n} \cdot \#\big\{j \in \{1, \ldots, n\} : x_j \in I\big\} \to |I| \quad \text{as } n \to \infty.$$

So, roughly speaking, for (x_n) to be equidistributed means that for large n, about $1/2$ of $x_1, \ldots, x_n$ are in $(0, 1/2]$, about $1/5$ are in $[3/5, 4/5]$, and so on.

Given $x \in \mathbb{R}$, write $\langle x \rangle = x - \lfloor x \rfloor$. Thus, $\langle x \rangle$ is the unique element of $[0,1)$ such that $x - \langle x \rangle \in \mathbb{Z}$ (or equivalently, such that $\langle x \rangle$ represents the same element of $\mathbb{T}$ as x).

Examples C6.2

i. For $n \geq 1$, put $x_n = \left\langle \frac{3}{8}n \right\rangle$ (Figure C.1). Then $x_n \in \{0, 1/8, \ldots, 7/8\}$ for all n, so

$$\frac{1}{n} \cdot \#\big\{j \in \{1, \ldots, n\} : x_j \in (0, 1/8)\big\} = 0$$

for all n, whereas $|(0,1/8)| = 1/8$. So (x_n) is not equidistributed.

ii. Similarly, for any $\alpha \in \mathbb{Q}$, the sequence $\big(\langle n\alpha \rangle\big)_{n=1}^\infty$ is not equidistributed.

So far, we haven't seen any examples of equidistributed sequences, only non-examples. It's not obvious that there are any equidistributed sequences at all! But using Fourier analysis, we'll construct lots.

When we defined equidistributed sequence, we were choosing a precise meaning for the phrase 'evenly spread'. But here's a different, less obvious interpretation. We could say that a sequence (x_n) in $[0, 1)$ is 'evenly spread' if for all integrable functions $f\colon \mathbb{T} \to \mathbb{C}$,

$$\frac{1}{n}\sum_{j=1}^{n} f(x_j) \to \int_0^1 f(t)\,dt \quad \text{as } n \to \infty. \tag{C:7}$$

Think for a while about why this is a sensible idea—understanding the rest of this section depends on it.

There are variations on this idea in which we use different classes of functions. For instance, we could regard (x_n) as 'evenly spread' if (C:7) holds for all continuous (rather than integrable) functions f, or all step functions, or all characters, etc. Actually, one of these variations produces the notion of equidistributed sequence itself:

Lemma C6.3 *Let $(x_n)_{n=1}^{\infty}$ be a sequence in $[0, 1)$. Then (x_n) is equidistributed if and only if* (C:7) *holds whenever f is the characteristic function of an interval in* $[0, 1)$.

Proof For an interval $I \subseteq [0, 1)$, and for $n \geq 1$, we have

$$\frac{1}{n}\sum_{j=1}^{n} \chi_I(x_j) = \frac{1}{n}\#\bigl\{j \in \{1, \ldots, n\} : x_j \in I\bigr\}$$

and

$$\int_0^1 \chi_I(t)\,dt = |I|\,.$$

The result follows. □

We now show that all these notions of 'evenly spread' sequence produced by choosing different classes of function are, in fact, the same.

Theorem C6.4 *Let $(x_n)_{n=1}^{\infty}$ be a sequence in $[0, 1)$. The following are equivalent:*

i. (C:7) *holds for all integrable $f\colon \mathbb{T} \to \mathbb{C}$;*

ii. (C:7) *holds for all characters $f = e_k\colon \mathbb{T} \to \mathbb{C}$, where $k \in \mathbb{Z}$;*

iii. (C:7) *holds for all trigonometric polynomials $f\colon \mathbb{T} \to \mathbb{C}$;*

iv. (C:7) *holds for all continuous $f\colon \mathbb{T} \to \mathbb{C}$;*

v. (C:7) *holds for all characteristic functions $f = \chi_I\colon \mathbb{T} \to \mathbb{C}$, where $I \subseteq$* $[0, 1)$ *is an interval;*

vi. (C:7) *holds for all step functions $f\colon \mathbb{T} \to \mathbb{C}$.*

Proof We show that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (vi) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii): trivial.

(ii) $\Rightarrow$ (iii): assume (ii). Let $f = \sum_{k=-m}^{m} c_k e_k$ be a trigonometric polynomial. Then

$$\frac{1}{n}\sum_{j=1}^{m} f(x_j) = \sum_{k=-m}^{m} c_k \cdot \frac{1}{n}\sum_{j=1}^{n} e_k(x_j) \to \sum_{k=-m}^{m} c_k \int_0^1 e_k(t)\,dt = \int_0^1 f(t)\,dt$$

as $n \to \infty$.

(iii) $\Rightarrow$ (iv): assume (iii). Let $f\colon \mathbb{T} \to \mathbb{C}$ be a continuous function. Let $\varepsilon > 0$. By Corollary C4.4, we can find a trigonometric polynomial g such that $\|f - g\|_\infty < \varepsilon/3$. By assumption, we can choose N such that for all $n \geq N$,

$$\left|\frac{1}{n}\sum_{j=1}^{n} g(x_j) - \int_0^1 g(t)\,dt\right| < \varepsilon/3.$$

Then for all $n \geq N$, using the triangle inequality,

$$\begin{aligned}
\left|\frac{1}{n}\sum_{j=1}^{n} f(x_j) - \int_0^1 f(t)\,dt\right| &\leq \left|\frac{1}{n}\sum_{j=1}^{n} f(x_j) - \frac{1}{n}\sum_{j=1}^{n} g(x_j)\right| + \left|\frac{1}{n}\sum_{j=1}^{n} g(x_j) - \int_0^1 g(t)\,dt\right| \\
&\quad + \left|\int_0^1 g(t)\,dt - \int_0^1 f(t)\,dt\right| \\
&< \frac{1}{n}\sum_{j=1}^{n} \|f - g\|_\infty + \varepsilon/3 + \|f - g\|_1 \\
&\leq \|f - g\|_\infty + \varepsilon/3 + \|f - g\|_\infty < \varepsilon.
\end{aligned}$$

(iv) $\Rightarrow$ (v): assume (iv). Let $I \subseteq [0,1)$ be an interval, and let $\varepsilon > 0$. We can find a continuous function $g_U\colon [0,1) \to \mathbb{R}$ such that $g_U(t) \geq \chi_I(t)$ for all $t \in [0,1)$ and

$$\int_0^1 g_U(t)\,dt - \int_0^1 \chi_I(t)\,dt < \varepsilon/2$$

(Figure C.2). By assumption, we can choose $N \geq 1$ such that for all $n \geq N$,

$$\left|\frac{1}{n}\sum_{j=1}^{n} g_U(x_j) - \int_0^1 g_U(t)\,dt\right| < \varepsilon/2.$$

So for all $n \geq N$,

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{n} \chi_I(x_j) - \int_0^1 \chi_I(t)\,dt && \text{(C:8)} \\
&\leq \frac{1}{n}\sum_{j=1}^{n} g_U(x_j) - \int_0^1 \chi_I(t)\,dt \\
&= \left(\frac{1}{n}\sum_{j=1}^{n} g_U(x_j) - \int_0^1 g_U(t)\,dt\right) + \left(\int_0^1 g_U(t)\,dt - \int_0^1 \chi_I(t)\,dt\right) \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

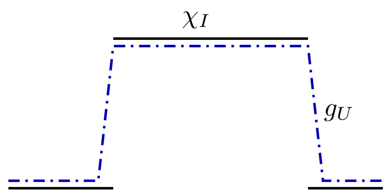


Figure C.2: The characteristic function χ_I of an interval, and a continuous function g_U approximating it. (It is intended here that $g_U(x) = \chi_I(x)$ for most values of x; the graphs are drawn a small distance apart for visual clarity only.)

Similarly, by approximating χ_I from below instead of above, (C:8) is greater than $-\varepsilon$.

(v) $\Rightarrow$ (vi): this is the same argument as in (ii) $\Rightarrow$ (iii). (Every step function is a finite linear combination of characteristic functions of intervals, just as every trigonometric polynomial is a finite linear combination of characters.)

(vi) $\Rightarrow$ (i): assume (vi). First take a *real*-valued integrable function $f\colon \mathbb{T} \to \mathbb{R}$. Let $\varepsilon > 0$. In the proof of Proposition A9.4, we showed that there is a step function g_L (called g there) such that $g_L(t) \leq f(t)$ for all $t \in [0, 1)$ and $\int_0^1 f(t)\,dt - \int_0^1 g_L(t)\,dt < \varepsilon/2$. Similarly, there is a step function g_U such that $g_U(t) \geq f(t)$ for all $t \in [0, 1)$ and $\int_0^1 g_U(t)\,dt - \int_0^1 f(t)\,dt < \varepsilon/2$. Arguing as in (iv) $\Rightarrow$ (v), we get $\frac{1}{n}\sum_{j=1}^n f(x_j) \to \int_0^1 f(t)\,dt$ as $n \to \infty$.

Now take *any* integrable function $f\colon \mathbb{T} \to \mathbb{C}$. Then $f = f_1 + if_2$ where $f_1, f_2\colon \mathbb{T} \to \mathbb{R}$ are integrable functions. The result follows by linearity, using the argument of (ii) $\Rightarrow$ (iii) again. $\square$

Corollary C6.5 *Let $(x_n)_{n=1}^\infty$ be a sequence in $[0, 1)$. Then (x_n) is equidistributed if and only if it satisfies* **Weyl's criterion***:*

$$\text{for each integer } k \neq 0, \quad \frac{1}{n}\sum_{j=1}^n e^{2\pi i k x_j} \to 0 \text{ as } n \to \infty.$$

Proof This is essentially (v) $\iff$ (ii) of Theorem C6.4. Indeed, condition (v) is equivalent to (x_n) being equidistributed, by Lemma C6.3. Condition (ii) holds if and only if for each $k \in \mathbb{Z}$,

$$\frac{1}{n}\sum_{j=1}^n e^{2\pi i k x_j} \to \int_0^1 e_k(t)\,dt \text{ as } n \to \infty. \tag{C:9}$$

When $k = 0$, (C:9) holds for any sequence (x_n) (check!). On the other hand, $\int_0^1 e_k(t)\,dt = \langle e_k, e_0\rangle = 0$ whenever $k \neq 0$. The result follows. $\square$

This enables us to find, at last, some examples of equidistributed sequences.

Corollary C6.6 (Weyl's equidistribution theorem) *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then the sequence $\big(\langle n\alpha\rangle\big)_{n=1}^\infty$ in $[0, 1)$ is equidistributed.*

Proof We verify Weyl's criterion. Let $0 \neq k \in \mathbb{Z}$. For all $n \geq 1$,

$$\frac{1}{n}\sum_{j=1}^n e^{2\pi i k \langle j\alpha\rangle} = \frac{1}{n}\sum_{j=1}^n e^{2\pi i k j \alpha}$$

since $\langle j\alpha\rangle - j\alpha \in \mathbb{Z}$. We sum this geometric series using the standard formula, which is valid as long as the ratio is not equal to 1. Here, this means $e^{2\pi ik\alpha} \neq 1$, or equivalently $k\alpha \notin \mathbb{Z}$, which is true as $k \neq 0$ and $\alpha \notin \mathbb{Q}$. Thus, the formula is valid and the sum of the geometric series is

$$\frac{1}{n} e^{2\pi ik\alpha} \frac{e^{2\pi ikn\alpha} - 1}{e^{2\pi ik\alpha} - 1}.$$

So for all $n \geq 1$,

$$\left| \frac{1}{n} \sum_{j=1}^{n} e^{2\pi ik\langle j\alpha\rangle} \right| \leq \frac{1}{n} \cdot \frac{\left|e^{2\pi ikn\alpha}\right| + |1|}{\left|e^{2\pi ik\alpha} - 1\right|} = \frac{1}{n} \cdot \frac{2}{\left|e^{2\pi ik\alpha} - 1\right|} \to 0$$

as $n \to \infty$. □

Example C6.7 Consider the sequence

$$e = 2.718\ldots,\ 2e = 5.436\ldots,\ 3e = 8.154\ldots,\ \ldots.$$

The sequence $\big(\langle ne\rangle\big)_{n=1}^{\infty}$ of fractional parts is

$$0.718\ldots,\ 0.436\ldots,\ 0.154\ldots,\ldots.$$

Weyl's equidistribution theorem tells us that 'in the long run', about 1/100 of these fractional parts lie between 0.12 and 0.13 (for instance). In particular, there are infinitely many natural numbers n such that

$$0.12 \leq \langle ne\rangle < 0.13.$$

Could you prove this without Weyl's theorem? Could you prove that there is even *one* number n with this property?

Chapter D

Duality

This course is called Fourier Analysis, but we have focused throughout on Fourier *series*. You're aware that Fourier *transforms* are also important. What is the relationship between the two?

Section D1 gives the answer. We'll see there that by looking at Fourier analysis from sufficiently high altitude—by taking a bird's eye view—it becomes apparent that Fourier series and Fourier transforms are two special cases of the same general construction.

But it also becomes apparent that there are other interesting special cases. One case that is both interesting and easy is Fourier analysis on finite abelian groups. That's what we'll study for the rest of Part D.

D1 An abstract view of Fourier analysis

For the lecture of Monday 17 March

This section (D1) is not directly *examinable, but reading it is recommended: it should help you to understand the ideas behind Part D and how it relates to the rest of the course.*

The right general context for Fourier analysis is that of topological groups. To understand what a topological group is, you first need to know roughly what a topological space is. So here is a short, informal introduction to topological spaces.

Roughly speaking, a set X is called a **topological space** if we know what it means to 'move gradually' within X. For instance, $\mathbb{R}^n$ is a topological space because we know what it means for one point to be close to another, and we know what a 'continuous path' in $\mathbb{R}^n$ is (namely, a continuous map $[0,1] \to \mathbb{R}^n$). Similarly, the sphere is a topological space: you know what it means to move gradually on the surface of the earth. The circle $\mathbb{T}$ is a topological space for the same reason.

Some non-mathematical examples to help your intuition: the set of all colours is a topological space, because you know what it means for a colour to change gradually. The set of all possible human faces is a topological space, because you know what it means for a face to change gradually (e.g. as you age) or for one face to be similar to another. Your own changing face defines a continuous map f from $[0,1]$ to the space of all faces, with $f(0)$ being your face at birth and $f(1)$ your face at death.

Formally, a topological space is a set X equipped with some extra data (specifying what it means to 'move gradually'). That extra data is called a 'topology' on X. We will not need the formal definition.

Some topological spaces are rather trivial. For instance, you can't move gradually from one integer to another without passing through non-integers, so the set $\mathbb{Z}$ is a topological space in a trivial way: the only way of moving gradually is to stay still. Formally, this is called the **discrete** topology on $\mathbb{Z}$. (If you know the definition of topological space, the discrete topology is the topology in which all subsets are open.) Any set can be given the discrete topology.

Topological spaces are the right context for the definition of continuous map. You know what continuity means for a function $\mathbb{R} \to \mathbb{R}$, and more generally for a function $\mathbb{R}^n \to \mathbb{R}^m$, and also for a function $\mathbb{T} \to \mathbb{C}$. Given topological spaces X and Y, there is a definition of what it means for a function $X \to Y$ to be continuous. The definitions for $\mathbb{R}^n$, $\mathbb{T}$, $\mathbb{C}$, etc. are all special cases of this general definition.

Roughly speaking, a **topological group** is a topological space that is also a group. We are most interested in those topological groups that are abelian (that is, the multiplication is commutative) and 'locally compact' (a condition that I will not explain; it is satisfied by all the examples I will mention). So, what we are interested in is locally compact, abelian topological groups, which are usually just called 'locally compact abelian groups' or LCAGs for short.

Examples D1.1 i. $\mathbb{R}^n$ is a LCAG (with addition as the group operation), for any $n \geq 0$.

ii. The circle $\mathbb{T}$ is a LCAG too. Recall that $\mathbb{T}$ is the quotient group $\mathbb{R}/\mathbb{Z}$, so that its group operation is addition too.

iii. $\mathbb{Z}$ is a LCAG, with $+$ as the group operation and the discrete topology.

iv. Any finite abelian group is a LCAG, with the discrete topology.

v. $\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\}$ is a LCAG, with $\cdot$ as the group operation. In fact, $\mathbb{T} \cong \mathbb{S}$, via the map $t \mapsto e^{2\pi i t}$.

Definition D1.2 Let G be a LCAG. A **character** of G is a continuous group homomorphism $G \to \mathbb{S}$.

Examples D1.3 i. For each $\xi \in \mathbb{R}^n$, there is a character e_ξ of $\mathbb{R}^n$ defined by

$$e_\xi(x) = e^{2\pi i \xi . x}$$

($x \in \mathbb{R}^n$), where $\xi.x$ is the dot product of ξ and x. (This is the ordinary product in the case $n = 1$.) You know that e_ξ is continuous, and you can easily check that e_ξ is a group homomorphism (try it!). Fact: these are the *only* characters of $\mathbb{R}^n$.

ii. For each $k \in \mathbb{Z}$, there is a character e_k of $\mathbb{T}$ defined in the usual way. Lemma A7.2 states exactly that each e_k is a character in the sense of Definition D1.2. Fact: these are the *only* characters of $\mathbb{T}$. So, the characters of $\mathbb{T}$ in our new sense are precisely the characters of $\mathbb{T}$ in our old sense.

iii. For each $t \in \mathbb{T}$, there is a character ε_t of $\mathbb{Z}$ defined by

$$\varepsilon_t(k) = e^{2\pi ikt}$$

($t \in \mathbb{T}$). Again, you can easily check that each ε_t really is a character, that is, a continuous homomorphism. Fact: these are the *only* characters of $\mathbb{Z}$.

The set
$$\widehat{G} = \{\text{characters of } G\}$$
forms a LCAG in a natural way. I won't describe the topology, but the group structure is given 'pointwise': if e_1 and e_2 are characters of G then the character $e_1 \cdot e_2$ is defined by $(e_1 \cdot e_2)(x) = e_1(x) \cdot e_2(x)$ $(x \in G)$. The hat notation is related to, but different from, the hat notation $\hat{f}$ for the Fourier coefficients or Fourier transform of a function f.

Examples D1.3 tell us what $\widehat{G}$ is for various LCAGs G:

G	$\widehat{G}$
$\mathbb{R}^n$	$\mathbb{R}^n$
$\mathbb{T}$	$\mathbb{Z}$
$\mathbb{Z}$	$\mathbb{T}$

For instance, $\widehat{\mathbb{T}} \cong \mathbb{Z}$ because we have an isomorphism

$$\begin{array}{ccc} \mathbb{Z} & \to & \widehat{\mathbb{T}} \\ k & \mapsto & e_k. \end{array}$$

We saw in Lemma A7.3 that this map is a homomorphism, and in Example D1.3(ii) that this map is surjective. It is not hard to see that it is also injective (that is, e_k and e_ℓ are different if $k \neq \ell$). So, it is an isomorphism of groups, and in fact an isomorphism of *topological* groups (that is, a homeomorphism too).

In each of the three cases in the table, $\widehat{\widehat{G}} \cong G$. In fact, this is a general phenomenon:

Theorem D1.4 (Pontryagin duality) *$\widehat{\widehat{G}} \cong G$ for all LCAGs G.*

The group $\widehat{G}$ can be thought of as the 'mirror image' of the group G (Figure D.1). Pontryagin duality states that the mirror image of the mirror image of G is G itself.

We now look briefly at Fourier analysis on locally compact abelian groups.

Let G be a LCAG. 'Nice' functions on G can be integrated. For example, you're used to integrating nice functions on $\mathbb{R}^n$ or $\mathbb{T}$. In the case of $\mathbb{Z}$, a function $\mathbb{Z} \to \mathbb{C}$ is just a double sequence $(c_k)_{k=-\infty}^{\infty}$ in $\mathbb{C}$, and integrating it means taking its sum (which is possible if it's 'nice' enough).

We can define the Fourier transform of a nice function on G, and also the inverse Fourier transform of a nice function on $\widehat{G}$:

$$\{\text{nice functions on } G\} \underset{(\check{\ })}{\overset{(\hat{\ })}{\rightleftarrows}} \{\text{nice functions on } \widehat{G}\},$$

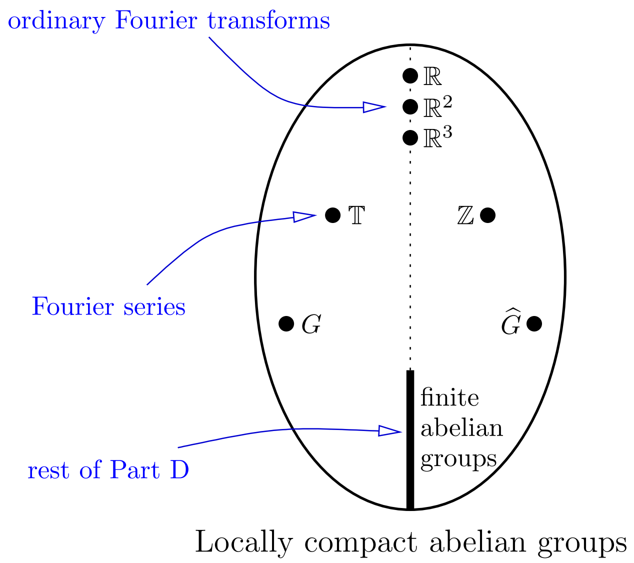


Figure D.1: The world of locally compact abelian groups. Think of the dotted line down the middle as a mirror. Reflecting in the mirror swaps G with $\widehat{G}$.

as follows: for a 'nice' function $f\colon G \to \mathbb{C}$, we define its Fourier transform $\hat{f}\colon \widehat{G} \to \mathbb{C}$ by

$$\hat{f}(e) = \int_G f(x)\overline{e(x)}\,dx$$

($e \in \widehat{G}$), and for a 'nice' function $\phi\colon \widehat{G} \to \mathbb{C}$, we define its inverse Fourier transform $\check{\phi}\colon G \to \mathbb{C}$ by

$$\check{\phi}(x) = \int_{\widehat{G}} \phi(e)e(x)\,de$$

($x \in G$). The terminology suggests that the two processes should be mutually inverse, which they are if the functions are 'nice' enough.

When $G = \mathbb{R}^n$ (and so $\widehat{G} = \mathbb{R}^n$ too), $\hat{f}$ is the usual Fourier transform of a function $f\colon \mathbb{R}^n \to \mathbb{C}$ (see Example D1.3(i)), and $\check{\phi}$ is the inverse Fourier transform of a function $\phi\colon \mathbb{R}^n \to \mathbb{C}$.

When $G = \mathbb{T}$ (and so $\widehat{G} = \mathbb{Z}$), what is here called $\hat{f}(e_k)$ is what we usually call $\hat{f}(k)$, the kth Fourier coefficient of f. Given a double sequence $\phi\colon \mathbb{Z} \to \mathbb{C}$, what is here called $\check{\phi}(x)$ is $\sum_{k\in\mathbb{Z}} \phi(k)e_k(x)$. So, this is nothing but Fantasy A8.7.

Both Fourier transforms and Fourier series therefore arise as special cases of the general theory of Fourier transforms on locally compact abelian groups. Developing this theory is a major undertaking, requiring some of the theory of topological groups and also some measure theory. But there is a special case in which all the complications disappear, and that is what we will study for the rest of the course. It is the theory of Fourier transforms on finite abelian groups.

D2 The dual of a finite abelian group

For the lecture of Monday 24 March. (Lecture of Thu 20 March is cancelled)

Maybe you're fed up with fussy analytic conditions: this function is *continuously* differentiable, that series is *absolutely* summable, and so on. If so, the last few lectures are for you. They will show you a world in which Fourier analysis works beautifully without any analytic conditions at all.

Everything will be developed from the ground up, assuming only some basic group theory. I will not assume any definitions, notation or results from Section D1.

Definition D2.1 We denote by $\mathbb{S}$ the multiplicative group $\{z \in \mathbb{C} : |z| = 1\}$.

Note that $\overline{z} = z^{-1}$ for $z \in \mathbb{S}$.

When we do Fourier analysis on finite abelian groups, we will always use multiplicative notation for our groups (even though the locally compact abelian groups in Section D1 were written additively). It's only notation.

Lemma D2.2 *Let G be a group. Then the set*

$$\widehat{G} = \{\textit{homomorphisms } G \to \mathbb{S}\}$$

is a group under the following operations:

- *the product of $e_1, e_2 \in \widehat{G}$ is the product function $e_1 \cdot e_2$ (defined as in Notation A3.3);*
- *the inverse of $e \in \widehat{G}$ is $\overline{e}$ (again, defined as in Notation A3.3);*
- *the identity is the constant function* 1*.*

Proof The set of *all* functions $G \to \mathbb{S}$ is certainly a group under these operations. We show that $\widehat{G}$ is a subgroup of this group, that is:

- if $e_1, e_2 \colon G \to \mathbb{S}$ are homomorphisms then so is $e_1 \cdot e_2$;
- if $e \colon G \to \mathbb{S}$ is a homomorphism then so is $\overline{e}$;
- the constant function $1 \colon G \to \mathbb{S}$ is a homomorphism.

For the first one, if $e_1, e_2 \in \widehat{G}$ then for all $x, y \in G$,

$$\begin{aligned}
(e_1 \cdot e_2)(xy) &= e_1(xy)e_2(xy) && \text{(by definition of } e_1 \cdot e_2\text{)} \\
&= e_1(x)e_1(y)e_2(x)e_2(y) && \text{(since } e_1 \text{ and } e_2 \text{ are homomorphisms)} \\
&= e_1(x)e_2(x)e_1(y)e_2(y) && \text{(since } \mathbb{S} \text{ is abelian)} \\
&= (e_1 \cdot e_2)(x) \cdot (e_1 \cdot e_2)(y) && \text{(by definition of } e_1 \cdot e_2\text{),}
\end{aligned}$$

so $e_1 \cdot e_2$ is a homomorphism. The arguments for $\overline{e}$ and 1 are similar (but simpler). □

Definition D2.3 Let G be a finite abelian group. A **character** of G is a homomorphism $G \to \mathbb{S}$. The **character group** or **Pontryagin dual** of G is the group $\widehat{G}$.

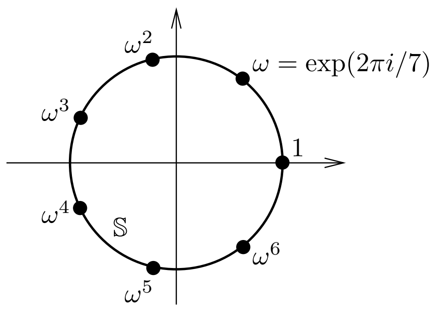


Figure D.2: The complex nth roots of unity are exactly the powers of $\exp(2\pi i/n)$. Shown: $n = 7$.

Let us pause to compare this with the definitions in Section D1 (although this is not logically necessary for anything that follows). In Definition D1.2, a character of a topological group G was defined to be a *continuous* homomorphism $G \to \mathbb{S}$. In Definition D2.3, continuity isn't mentioned. What's going on?

Implicitly, we are putting the discrete topology on our finite abelian groups (as in Example D1.1(iv)). The discrete topology has the special property that all maps out of a discrete space are automatically continuous. Thus, when G is given the discrete topology, a continuous homomorphism $G \to \mathbb{S}$ is the same thing as an ordinary homomorphism $G \to \mathbb{S}$. One of the nice things about Fourier theory for finite abelian groups is that the topological and analytic conditions vanish entirely.

Coming back to the main story, let C_n denote the cyclic group of order n.

Lemma D2.4 *$\widehat{C_n} \cong C_n$ for all $n \geq 1$.*

Proof Let $n \geq 1$. We prove that $\widehat{C_n}$ is a cyclic group of order n.

Choose a generator x of C_n. Let $\omega = \exp(2\pi i/n)$, and note that the complex numbers z satisfying $z^n = 1$ are exactly the integer powers of ω (Fig. D.2).

Define $d\colon C_n \to \mathbb{S}$ by $d(x^r) = \omega^r$ ($r \in \mathbb{Z}$). This is well-defined, as if $x^{r_1} = x^{r_2}$ then $r_1 \equiv r_2 \pmod{n}$, so $\omega^{r_1} = \omega^{r_2}$ (using the fact that $\omega^n = 1$). It is also a homomorphism, since for $r, s \in \mathbb{Z}$,

$$d(x^r \cdot x^s) = d(x^{r+s}) = \omega^{r+s} = \omega^r \cdot \omega^s = d(x^r) \cdot d(x^s).$$

So d is a character of C_n.

Now let e be any character of C_n. We have $e(x)^n = e(x^n) = e(1) = 1$, so $e(x) = \omega^r$ for some $r \in \mathbb{Z}$. I claim that $e = d^r$. Indeed, given $y \in C_n$, we have $y = x^s$ for some $s \in \mathbb{Z}$, and then

$$e(y) = e(x^s) = e(x)^s = \omega^{rs} = d(x)^{rs} = d(x^s)^r = d(y)^r = d^r(y),$$

proving the claim. So $\widehat{C_n}$ is cyclic, generated by d.

It remains to find the order of d. Certainly $d^n = 1$, since every element of C_n is of the form x^r for some $r \in \mathbb{Z}$, and $d^n(x^r) = d(x)^{nr} = \omega^{nr} = 1$. On the other hand, if $1 \leq m < n$ then $d^m(x) = \omega^m \neq 1$, so $d^m \neq 1$. So d has order n.

We have shown that $\widehat{C_n}$ is a cyclic group generated by an element of order n. Thus, $\widehat{C_n} \cong C_n$. □

Next recall that any two groups G_1 and G_2 have a product (also called a 'direct product') $G_1 \times G_2$.

Lemma D2.5 $\widehat{G_1 \times G_2} \cong \widehat{G_1} \times \widehat{G_2}$ *for all finite abelian groups* G_1 *and* G_2*.*

Proof Sheet 5. □

Fact D2.6 *Every finite abelian group is isomorphic to* $C_{n_1} \times C_{n_2} \times \cdots \times C_{n_k}$ *for some* $k, n_1, n_2, \ldots, n_k \geq 1$*.*

This is part of the classification theorem for finite abelian groups, which you may have met in other courses.

Putting together the last three results gives:

Proposition D2.7 $\widehat{G} \cong G$ *for every finite abelian group* G*.*

Proof This follows from Lemma D2.4, Lemma D2.5 and Fact D2.6. □

This proposition is the reason why the finite abelian groups were drawn on the central dotted line (the 'mirror') of Figure D.1.

Remark D2.8 For a given G, there is usually no *canonical* (i.e. God-given) isomorphism $G \to \widehat{G}$. In order to construct an isomorphism, you have to make some arbitrary choice, of the same kind you make when tossing a coin or choosing a basis for a vector space.

D3 Fourier transforms on a finite abelian group

For the lecture of Thursday 27 March

First we'll make some definitions analogous to the definitions for Fourier series. Then we'll prove some results analogous to results on Fourier series. We'll discover that life is much easier on a finite abelian group than on the circle.

For the rest of this section, let G be a finite abelian group.

Definition D3.1 i. Given a set X, write $\mathrm{Fn}(X) = \{\text{functions } X \to \mathbb{C}\}$.

ii. For $f \in \mathrm{Fn}(G)$, write $\int_G f(x)\,dx = \frac{1}{\#G}\sum_{x\in G} f(x)$.

iii. For $f, g \in \mathrm{Fn}(G)$, write $\langle f, g\rangle = \int_G f(x)\overline{g(x)}\,dx$.

In (i), $\mathrm{Fn}(X)$ is a vector space over $\mathbb{C}$, via the usual addition and scalar multiplication of functions (as in Notation A3.3).

In (ii), the factor $\frac{1}{\#G}$ ensures that $\int_G 1\,dx = 1$, just as $\int_{\mathbb{T}} 1\,dx = 1$. So the integral of a function on G can be thought of as its mean value, just as for functions on $\mathbb{T}$.

We now establish some elementary properties of integration, directly analogous to those for ordinary integration stated in Lemma A3.4.

Lemma D3.2 *i. For any functions $f, g\colon G \to \mathbb{C}$,*

$$\int_G (f+g)(x)\,dx = \int_G f(x)\,dx + \int_G g(x)\,dx.$$

ii. For any function $f\colon G \to \mathbb{C}$ and $c \in \mathbb{C}$,

$$\int_G (cf)(x)\,dx = c\int_G f(x)\,dx.$$

iii. For any function $f\colon G \to \mathbb{C}$,

$$\int_G \bar{f}(x)\,dx = \overline{\int_G f(x)\,dx}.$$

Proof For (i), let $f, g \in \mathrm{Fn}(G)$. Then

$$\begin{aligned}
\int_G (f+g)(x)\,dx &= \int_G (f(x)+g(x))\,dx && \text{(by definition of } f+g)\\
&= \frac{1}{\#G}\sum_{x\in G}(f(x)+g(x)) && \text{(by definition of } \textstyle\int_G)\\
&= \frac{1}{\#G}\sum_{x\in G} f(x) + \frac{1}{\#G}\sum_{x\in G} g(x)\\
&= \int_G f(x)\,dx + \int_G g(x)\,dx && \text{(by definition of } \textstyle\int_G).
\end{aligned}$$

Similar arguments prove parts (ii) and (iii). $\square$

Lemma D3.3 *$\langle \cdot, \cdot \rangle$ is an inner product on* $\mathrm{Fn}(G)$.

Proof This is straightforward, using Lemma D3.2. (Compare your solution to Sheet 1, q.6(i).) $\square$

Note that $\langle \cdot, \cdot \rangle$ is a *genuine* inner product; that is, if $\langle f, f \rangle = 0$ then $f = 0$. (Contrast the comments after Lemma A6.2.) This is because for $f \in \mathrm{Fn}(G)$,

$$\langle f, f \rangle = \frac{1}{\#G} \sum_{x \in G} |f(x)|^2,$$

and if this is 0 then $f(x) = 0$ for all $x \in G$.

In the proof of Lemma A7.4 (which stated that the characters of $\mathbb{T}$ are orthonormal), it was shown in passing that $\int_{\mathbb{T}} e_k(x)\, dx$ is 1 if $k = 0$ and 0 otherwise. Something very similar holds for finite abelian groups.

Lemma D3.4 *Let $e \in \widehat{G}$. Then*

$$\int_G e(x)\, dx = \begin{cases} 1 & \textit{if } e = 1 \\ 0 & \textit{otherwise.} \end{cases}$$

Remember that the elements of $\widehat{G}$ are homomorphisms, so the '1' in '$e = 1$' means the constant function 1 from G to $\mathbb{C}$.

Proof If $e = 1$ then $\int_G e(x)\, dx = 1$ (as noted after Definition D3.1).

If $e \neq 1$, choose $y \in G$ such that $e(y) \neq 1$. Then

$$\int_G e(x)\, dx = \frac{1}{\#G} \sum_{x \in G} e(x) = \frac{1}{\#G} \sum_{z \in G} e(yz)$$

since as x runs over G, $z = y^{-1}x$ runs over G too. (Morally, this is integration by substitution; we're substituting $x = yz$.) So

$$\int_G e(x)\, dx = \frac{1}{\#G} \sum_{z \in G} e(y)e(z) = e(y) \int_G e(x)\, dx.$$

But $e(y) \neq 1$, so $\int_G e(x)\, dx = 0$. $\square$

Proposition D3.5 *The characters of G are orthonormal: for $e_1, e_2 \in \widehat{G}$,*

$$\langle e_1, e_2 \rangle = \begin{cases} 1 & \textit{if } e_1 = e_2 \\ 0 & \textit{otherwise.} \end{cases}$$

(Compare Lemma A7.4, the analogous result on $\mathbb{T}$.)

Proof Let $e_1, e_2 \in \widehat{G}$. Since $\overline{e_2}$ is the inverse of e_2 in $\widehat{G}$, we have $e_1\overline{e_2} \in \widehat{G}$. By Lemma D3.4,

$$\langle e_1, e_2 \rangle = \int_G (e_1\overline{e_2})(x)\, dx = \begin{cases} 1 & \text{if } e_1\overline{e_2} = 1 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} 1 & \text{if } e_1 = e_2 \\ 0 & \text{otherwise.} \end{cases}$$ $\square$

Definition D3.6 Let $f \in \mathrm{Fn}(G)$. The **Fourier transform** of f is the function $\hat{f} \in \mathrm{Fn}(\widehat{G})$ defined by $\hat{f}(e) = \langle f, e \rangle \quad (e \in \widehat{G})$.

(Compare: for Fourier series, $\hat{f}(k) = \langle f, e_k \rangle$.)

Note that we've defined the Fourier transform of *any* function on G! In our simple world of finite groups, there's no integrability to worry about.

Proposition D3.7 *For all $\phi \in \mathrm{Fn}(\widehat{G})$,*

$$\phi = \left(\sum_{e \in \widehat{G}} \phi(e) \cdot e \right)^{\wedge}$$

(The right-hand side means $\hat{f}$ where $f = \sum_{e \in \widehat{G}} \phi(e) \cdot e$. This alternative notation is simply for typesetting convenience, to avoid very wide hats.)

The sum on the right-hand side is a *finite* sum, since $\widehat{G} \cong G$ (by Proposition D2.7) and G is finite. Unlike the circle, G has only finitely many characters.

Proof Both sides of this equation are functions on $\widehat{G}$, so to prove equality, we need to take an arbitrary element e' of $\widehat{G}$ and show that evaluating each side at e' gives the same result.

Let $e' \in \widehat{G}$. Then

$$\begin{aligned} \left(\sum_{e \in \widehat{G}} \phi(e) \cdot e \right)^{\wedge} (e') &= \left\langle \sum_{e} \phi(e) \cdot e, e' \right\rangle \\ &= \sum_{e} \phi(e) \langle e, e' \rangle \qquad \text{(by Lemma D3.3)} \\ &= \phi(e') \qquad \text{(by orthonormality).} \end{aligned}$$

□

This result is similar to the 'Moreover' part of Proposition C2.1, which stated that for 'nice' $c \in \mathrm{Fn}(\mathbb{Z})$, we have $c = \left(\sum_{k \in \mathbb{Z}} c_k e_k \right)^{\wedge}$.

Recall that $\mathrm{Fn}(G)$ is a vector space over $\mathbb{C}$.

Lemma D3.8 $\dim(\mathrm{Fn}(G)) = \#G$.

Proof For each $x \in G$, define $\delta_x \in \mathrm{Fn}(G)$ by

$$\delta_x(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

($y \in G$). The family $(\delta_x)_{x \in G}$ spans $\mathrm{Fn}(G)$, since if $f \in \mathrm{Fn}(G)$ then $f = \sum_{x \in G} f(x) \delta_x$ (check!). It is also linearly independent, since if $(c_x)_{x \in G}$ is a family of complex numbers such that $\sum_{x \in G} c_x \delta_x = 0$, then for each $y \in G$ we have

$$0 = \left(\sum_{x \in G} c_x \delta_x \right)(y) = \sum_{x \in G} c_x \delta_x(y) = c_y.$$

So it is a basis. □

Lemma D3.8 doesn't really have an analogue in the world of Fourier series (that is, Fourier theory on the circle group $\mathbb{T}$).

Theorem D3.9 *The characters of G form an orthonormal basis of* $\mathrm{Fn}(G)$.

Proof By Proposition D3.5, the characters are orthonormal. In particular, they are linearly independent, so the subspace of $\mathrm{Fn}(G)$ that they span has dimension $\#\widehat{G}$. But $\#\widehat{G} = \#G = \dim(\mathrm{Fn}(G))$ by Proposition D2.7 and Lemma D3.8, so their span is $\mathrm{Fn}(G)$. $\square$

Hence *every* function on G is a 'trigonometric polynomial' (a finite linear combination of characters). This makes Fourier analysis on finite groups considerably simpler than Fourier analysis on the circle.

For a function $f: G \to \mathbb{C}$, the analogue of the Fourier series of f is the sum $\sum_{e\in\widehat{G}} \hat{f}(e)e$. Since this sum is finite, there is no question of convergence; either it is equal to f, or it is not. The next result says that it is.

Corollary D3.10 *For all $f \in \mathrm{Fn}(G)$,* $\quad f = \sum_{e\in\widehat{G}} \hat{f}(e)e.$

Proof By Theorem D3.9, $f = \sum_{e\in\widehat{G}} \langle f, e\rangle e$; but $\hat{f}(e) = \langle f, e\rangle$ by definition. $\square$

So we can reconstruct a function from its Fourier transform. For Fourier series, we had to work hard to prove that!

Since we can reconstruct a function from its Fourier transform, if two functions have the same Fourier transform then they must in fact be the same:

Corollary D3.11 *Let $f, g \in \mathrm{Fn}(G)$. If $\hat{f} = \hat{g}$ then $f = g$.*

Proof Follows from Corollary D3.10. $\square$

This is the (simpler) analogue of Corollary B8.5; see also Remark C4.3.

D4 Fourier inversion on a finite abelian group

For the lecture of Monday 31 March

Throughout this section, let G be a finite abelian group.

Definition D4.1 For $\phi \in \mathrm{Fn}(\widehat{G})$, define $\check{\phi} \in \mathrm{Fn}(G)$ by $\check{\phi} = \sum_{e \in \widehat{G}} \phi(e)e$.

The upside-down hat is meant to suggest the opposite of the Fourier transform. It is pronounced 'check' (as in 'phi-check').

Remark D4.2 Suppose we were dealing with the circle $\mathbb{T}$ rather than a *finite* abelian group G. We have $\widehat{\mathbb{T}} = \mathbb{Z}$ (as in Section D1), so ϕ would be an element of $\mathrm{Fn}(\mathbb{Z})$, that is, a double sequence $c = (c_k)_{k \in \mathbb{Z}}$. Then $\check{c}$ is notation for the familiar expression $\sum_{k \in \mathbb{Z}} c_k e_k$.

Fantasy A8.7 comes true in the setting of finite abelian groups. (No fussy analytical details!) This is the content of the following theorem.

Theorem D4.3 *The maps*

$$\mathrm{Fn}(G) \underset{(\check{\ })}{\overset{(\hat{\ })}{\rightleftarrows}} \mathrm{Fn}(\widehat{G})$$

are linear and mutually inverse. In particular, the vector spaces $\mathrm{Fn}(G)$ *and* $\mathrm{Fn}(\widehat{G})$ *are isomorphic.*

Proof Corollary D3.10 states that $\check{\hat{f}} = f$ for all $f \in \mathrm{Fn}(G)$, and Proposition D3.7 states that $\hat{\check{\phi}} = \phi$ for all $\phi \in \mathrm{Fn}(\widehat{G})$. So $(\hat{\ })$ and $(\check{\ })$ are mutually inverse. The map $(\hat{\ })$ is linear since $\langle \cdot, \cdot \rangle$ is linear in the first argument; so its inverse $(\check{\ })$ is linear too. Hence $(\hat{\ })$ and $(\check{\ })$ define an isomorphism of vector spaces. □

Definition D4.4 For $f \in \mathrm{Fn}(G)$, define $\|f\|_2 = \sqrt{\langle f, f \rangle}$.

This defines a norm $\| \cdot \|_2$ on $\mathrm{Fn}(G)$, since $\langle \cdot, \cdot \rangle$ is an inner product.

Theorem D4.5 (Parseval for finite abelian groups) *For all* $f \in \mathrm{Fn}(G)$,

$$\|f\|_2 = \sqrt{\sum_{e \in \widehat{G}} |\hat{f}(e)|^2}.$$

Proof By Corollary D3.10, $f = \sum_{e \in \widehat{G}} \hat{f}(e)e$. So by orthogonality of the characters,

$$\|f\|_2^2 = \langle f, f \rangle = \sum_{e \in \widehat{G}} |\hat{f}(e)|^2.$$

□

We know from Theorem D4.3 that the function $(\hat{\ })\colon \mathrm{Fn}(G) \to \mathrm{Fn}(\widehat{G})$ is an isomorphism of vector spaces. We'd like to say that it's also an isomorphism of *metric* spaces, that is, distance-preserving: $\|f\|_2 = \|\hat{f}\|_2$ for all $f \in \mathrm{Fn}(G)$. But the right-hand side is not yet defined. So we make some definitions:

Definition D4.6 i. For $\phi \in \mathrm{Fn}(\widehat{G})$, define $\int_{\widehat{G}} \phi(e)\, de = \sum_{e \in \widehat{G}} \phi(e) \in \mathbb{C}$.

ii. For $\phi, \psi \in \mathrm{Fn}(\widehat{G})$, define $\langle \phi, \psi \rangle = \int_{\widehat{G}} \phi(e)\overline{\psi(e)}\, de \in \mathbb{C}$.

iii. For $\phi \in \mathrm{Fn}(\widehat{G})$, define $\|\phi\|_2 = \sqrt{\langle \phi, \phi \rangle}$.

Compare and contrast (i) with Definition D3.1(ii): here, there is no factor of $\frac{1}{\#G}$. This is to make the following true:

Corollary D4.7 (Parseval, restated) *For all $f \in \mathrm{Fn}(G)$,*

$$\|f\|_2 = \|\hat{f}\|_2.$$

Proof This is a restatement of Theorem D4.5, since for $\phi \in \mathrm{Fn}(\widehat{G})$ we have $\|\phi\|_2 = \sqrt{\sum_{e \in \widehat{G}} |\phi(e)|^2}$. □

The characters of G are functions $G \to \mathbb{C}$ with certain properties. Are there enough such functions to tell different points of G apart? That is, if x and y are two different points in G, can we find a character e such that $e(x)$ and $e(y)$ are different?

The answer is yes, and it's now easy for us to prove. But it's not so easy to prove from scratch: try it!

Proposition D4.8 (Characters separate points) *Let $x, y \in G$ with $x \neq y$. Then there is some $e \in \widehat{G}$ such that $e(x) \neq e(y)$.*

Proof Suppose that $e(x) = e(y)$ for all $e \in \widehat{G}$. Since $x \neq y$, there is some $f \in \mathrm{Fn}(G)$ such that $f(x) \neq f(y)$. (E.g. define f by $f(x) = 1$ and $f(z) = 0$ for all $z \neq x$.) Then by Corollary D3.10,

$$f(x) = \sum_{e \in \widehat{G}} \hat{f}(e)e(x) = \sum_{e \in \widehat{G}} \hat{f}(e)e(y) = f(y),$$

a contradiction. □

Remark D4.9 We know from Proposition D2.7 that $G \cong \widehat{G} \cong \widehat{\widehat{G}}$, so $G \cong \widehat{\widehat{G}}$. But for all we know so far, there is no *canonical* choice of isomorphism $G \to \widehat{\widehat{G}}$. (Recall Remark D2.8.) However, the next result tells us that although there is no canonical isomorphism between a group and its single dual, there *is* a canonical isomorphism between a group and its *double* dual.

Theorem D4.10 *For $x \in G$, define $\mathrm{ev}_x \in \mathrm{Fn}(\widehat{G})$ by*

$$\mathrm{ev}_x(e) = e(x)$$

($e \in \widehat{G}$). Then $\mathrm{ev}_x \in \widehat{\widehat{G}}$ for all $x \in G$, and the map

$$\begin{array}{ccc} G & \longrightarrow & \widehat{\widehat{G}} \\ x & \longmapsto & \mathrm{ev}_x \end{array}$$

is an isomorphism of groups.

Proof For each $x \in G$, the function ev_x is a homomorphism $\widehat{G} \to \mathbb{S}$: indeed, for all $e_1, e_2 \in \widehat{G}$, we have

$$\begin{aligned} \mathrm{ev}_x(e_1 \cdot e_2) &= (e_1 \cdot e_2)(x) && \text{(by definition of } \mathrm{ev}_x\text{)} \\ &= e_1(x) e_2(x) && \text{(by definition of } e_1 \cdot e_2\text{)} \\ &= \mathrm{ev}_x(e_1)\, \mathrm{ev}_x(e_2) && \text{(by definition of } \mathrm{ev}_x\text{).} \end{aligned}$$

This means we can define a function $\mathrm{ev}\colon G \to \widehat{\widehat{G}}$ by $\mathrm{ev}(x) = \mathrm{ev}_x$ $(x \in G)$.

This function ev is a homomorphism. Indeed,

$$\begin{aligned} & \mathrm{ev} \text{ is a homomorphism} \\ \iff & \forall x, y \in G,\ \mathrm{ev}(xy) = \mathrm{ev}(x) \cdot \mathrm{ev}(y) \\ \iff & \forall x, y \in G,\ \mathrm{ev}_{xy} = \mathrm{ev}_x \cdot \mathrm{ev}_y \\ \iff & \forall x, y \in G, \forall e \in \widehat{G},\ \mathrm{ev}_{xy}(e) = (\mathrm{ev}_x \cdot \mathrm{ev}_y)(e) \\ \iff & \forall x, y \in G, \forall e \in \widehat{G},\ \mathrm{ev}_{xy}(e) = \mathrm{ev}_x(e) \cdot \mathrm{ev}_y(e) \\ \iff & \forall x, y \in G, \forall e \in \widehat{G},\ e(xy) = e(x) \cdot e(y), \end{aligned}$$

which is true. So reading the implications from bottom to top, ev is a homomorphism.

Next, $\mathrm{ev}\colon G \to \widehat{\widehat{G}}$ is injective. For let $x, y \in G$ and suppose that $\mathrm{ev}(x) = \mathrm{ev}(y)$. Then $\mathrm{ev}_x = \mathrm{ev}_y$, that is, $\mathrm{ev}_x(e) = \mathrm{ev}_y(e)$ for all $e \in \widehat{G}$. But characters separate points (Proposition D4.8), so $x = y$.

Finally, $\mathrm{ev}\colon G \to \widehat{\widehat{G}}$ is surjective, since ev is injective, $\#G = \#\widehat{G} = \#\widehat{\widehat{G}}$ by Proposition D2.7, and an injection from one finite set to another set of the same cardinality is always surjective, by the pigeonhole principle. $\square$

This is Pontryagin's duality theorem (Theorem D1.4) in the special case of finite abelian groups.

We have now proved analogues of all the main results on Fourier series in the setting of finite abelian groups. We have also proved the central theorem of abstract Fourier analysis (Pontryagin duality) in this simplified setting. And we have done all this in just a few pages, calling on nothing more than basic group theory.

* * *