A First Look at Fourier Analysis

T. W. Körner

August 2, 2003

These are the skeleton notes of an undergraduate course given at the PCMI conference in 2003. I should like to thank the organisers and my audience for an extremely enjoyable three weeks. The document is written in LaTeX2e and should be available in tex, ps, pdf and dvi format from my home page

http://www.dpmms.cam.ac.uk/~twk/

Corrections sent to **twk@dpmms.cam.ac.uk** would be extremely welcome. Mihai Stoiciu who was TA for the course has kindly written out solutions to some of the exercises. These are also accessible via my home page.

Contents

1 Waves in strings

It is said that Pythagoras was the first to realise that the notes emitted by struck strings of lengths l. $l/2$, $l/3$ and so on formed particularly attractive harmonies for the human ear. From this he concluded, it is said, that all is number and the universe is best understood in terms of mathematics — one of the most outrageous and most important leaps of faith in human history.

Two millennia later the new theory of mechanics and the new method of mechanics enabled mathematicians to write down a model for a vibrating string. Our discussion will be exploratory with no attempt at rigour. Suppose that the string is in tension T and has constant density ρ. If the graph of the position of the string at time t is given by $y = Y(x,t)$ where $Y(x,t)$ is always very small then, working to the first order in δx, the portion of the string between x and $x + \delta x$ experiences a force parallel to the y-axis of

$$T\left(\frac{\partial Y}{\partial x}(x+\delta x,t) - \frac{\partial Y}{\partial x}(x,t)\right) = T\delta x\frac{\partial^2 Y}{\partial x^2}.$$

Applying Newton's second law we obtain (still working to first order)

$$\rho\delta x\frac{\partial^2 Y}{\partial t^2} = T\delta x\frac{\partial^2 Y}{\partial x^2}.$$

Thus we have the exact equation

$$\rho\frac{\partial^2 Y}{\partial t^2} = T\frac{\partial^2 Y}{\partial x^2}.$$

For reasons which will become apparent later, it is usual to write c for the positive square root of T/ρ giving our equation in the form

$$\frac{\partial^2 Y}{\partial t^2} = c^2 \frac{\partial^2 Y}{\partial x^2}. \qquad \bigstar$$

Equation $\bigstar$ is often called 'the wave equation'.

Let us try and solve the wave equation for a string fixed at 0 and l (that is, with $Y(0,t) = Y(0,l) = 0$ for all t). Since it is rather ambitious to try and find *all* solutions let us try and find some solutions. A natural approach is to seek solutions of the particular form $Y(x,t) = X(x)T(t)$. Substitution in $\bigstar$ gives

$$X(x)T''(t) = c^2 X''(x)T(t) \text{ which we can rewrite as } \frac{T''(t)}{T(t)} = c^2 \frac{X''(x)}{X(x)}.$$

Since a quantity which depends only on x and only on t must be constant $X''(x)/X(x)$ must be constant on $(0,l)$. Thus $\frac{X''(x)}{X(x)}$ must take a constant value K.

If $K = -\omega^2$ with $\omega > 0$, then

$$X(x) = Ae^{\omega x} + Be^{-\omega x}$$

for appropriate constants A and B. If $K = 0$, then

$$X(x) = A + Bx$$

for appropriate constants A and B. If $K = -\omega^2$ with $\omega > 0$, then

$$X(x) = A\cos\omega x + B\sin\omega x$$

for appropriate constants A and B. However, since $Y(0,t) = Y(0,l) = 0$, we must have $X(0) = X(l) = 0$. (We ignore the uninteresting possibility $T(t) = 0$ for all t.) The only way to obtain a non-trivial solution for X is to take $K = -(n\pi/l)^2$ with n a strictly positive integer. This yields

$$X(x) = B\sin n\pi x/l \text{ and } T''(t) + (n\pi c/l)^2 T(t) = 0$$

and gives us the particular solutions

$$Y(x,t) = (a_n \cos(n\pi ct/l) + b_n \sin(n\pi ct/l)) \sin n\pi x/l.$$

The wave equation is *linear* in the sense that if Y_1 and Y_2 are solutions of $\bigstar$ then so is $\lambda_1 Y_1 + \lambda_2 Y_2$. Subject to appropriate conditions

$$Y(x,t) = \sum_{n=1}^{\infty} (a_n \cos(n\pi ct/l) + b_n \sin(n\pi ct/l)) \sin n\pi x/l$$

will be a solution of our problem. It is natural to ask if this is the *most general* solution. More specifically, it is natural to ask, whether if u, $v : [0, l] \to \mathbb{R}$ are well behaved functions with $u(0) = v(0) = u(l) = v(l)$ we can find a_n and b_n such that,

$$Y(x, 0) = u(x) \text{ and } \frac{\partial Y}{\partial t}(x, 0) = v(x).$$

Without worrying too much about rigour, our question reduces to asking whether we can find a_n and b_n such that

$$\sum_{n=1}^{\infty} a_n \sin n\pi x/l = u(x) \text{ and } \sum_{n=1}^{\infty} \frac{n\pi c}{l} b_n \sin n\pi x/l = v(x).$$

The next lemma does not answer our question but indicates the direction our answer might take.

Lemma 1.1. *(i)* $\sin\theta \sin\phi = \frac{1}{2}\big(\cos(\theta - \phi) - \cos(\theta + \phi)\big)$.

(ii) $\displaystyle\int_0^l \sin\frac{n\pi x}{l} \sin\frac{m\pi x}{l}\, dx = 0$ *if n and m are distinct integers.*

(iii) $\displaystyle\int_0^l \sin\frac{n\pi x}{l} \sin\frac{n\pi x}{l}\, dx = \frac{l}{2}$ *if n is an integer.*

(iv) The **formal** *solution of* $\sum_{n=1}^{\infty} a_n \sin\frac{n\pi x}{l} = u(x)$ *is given by*

$$a_n = \frac{2}{l}\int_0^l u(x) \sin\frac{n\pi x}{l}\, dx.$$

In this course we will investigate to what extent a general function can be written in the kind of way suggested **formally** by lemma 1.1

D'Alambert came up with another very elegant way of solving the wave equation. Here we deal with an infinite string.

Lemma 1.2. *(i) If we write $\sigma = x+ct$ and $\tau = x-ct$ then the wave equation*

$$\frac{\partial^2 Y}{\partial t^2} = c^2 \frac{\partial^2 Y}{\partial x^2} \textit{ becomes } \frac{\partial^2 Y}{\partial \tau \partial \sigma} = 0.$$

(ii) The general solution of the wave equation is

$$Y(x, t) = f(x - ct) + g(x + ct).$$

Note how the solution can be interpreted as two 'signals' traveling with velocity c in two directions.

2 Approximation by polynomials

If you have met Fourier Analysis before you have probably met it as as the study of 'decompositions into sines and cosines'. We shall treat it as 'decomposition into functions of the form $\exp i\lambda x = \cos \lambda x + i \sin \lambda x$'. Technically, this is a rather trivial change but you are entitled to ask 'why bring complex numbers into the study of real objects?' In the first two sections I will try to convince you that there can be genuine advantages in such a procedure. Neither section is directly related to Fourier Analysis but this course is a ramble through fine scenery rather than a forced march to some distant goal.

Our first topic will be approximation by polynomials. Complex numbers and trigonometric functions will only make a brief (but useful) appearance at the very end.

We start by recalling some useful results from algebra.

Lemma 2.1. *(i) If P is a polynomial of degree $n \geq 1$ and a is constant then there exists a polynomial Q of degree $n-1$ and a constant r such that*

$$P(x) = (x-a)Q(x) + r.$$

(ii) If P is a polynomial of degree $n \geq 1$ and a is a zero of P then there exists a polynomial Q of degree $n-1$ such that

$$P(x) = (x-a)Q(x).$$

(iii) If P is a polynomial of degree at most n which vanishes at $n+1$ distinct points then $P = 0$.

Suppose that $b > a$, that $x_0, x_1, \dots, x_n$ are distinct points of $[a,b]$ and that $f : [a,b] \to \mathbb{R}$ is a given function. We say that at real polynomial P interpolates f at the points $x_0, x_1, \dots, x_n$ if $f(x_j) = P(x_j)$ for $0 \leq j \leq n$.

Lemma 2.1 (iii) gives the following useful fact.

Lemma 2.2. *With the notation just introduced, there can exist at most one polynomial P of degree at most n which interpolates f at the points $x_0, x_1, \dots, x_n$.*

This uniqueness result is complemented by an existence result.

Lemma 2.3. *We use the notation introduced above.*

(i) If we set

$$e_j(x) = \prod_{k \neq j} \frac{x - x_k}{x_j - x_k},$$

the e_j is a polynomial of degree n such that $e_j(x_k) = 0$ for $k \neq j$ and $e_j(x_j) = 1$.

(ii) There exists a polynomial P of degree at most n which interpolates f at the points $x_0, x_1, \dots, x_n$.

We have thus shown that there is unique interpolating polynomial P of degree at most n which agrees with f at $n+1$ points. How good an approximation is P to f at other points? If f is reasonably smooth, an ingenious use of Rolle's theorem gives a gives a partial answer.

Theorem 2.4. (Rolle's Theorem.) *If $F : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and $F(a) = F(b)$, then there exists a $c \in (a, b)$ such that $F'(c) = 0$.*

Lemma 2.5. *Suppose that f is $n+1$ times differentiable. With the notation of this section, let t be a point distinct from the x_j. Set*

$$E(t) = f(t) - P(t)$$

(so $E(t)$ is the 'error at point t) and write

$$g(x) = f(x) - P(x) - E(t) \prod_{k=0}^{n} \frac{x - x_k}{t - x_k}.$$

(i) The function g is $n+1$ times differentiable and vanishes at $n+2$ distinct points on $[a, b]$.

(ii) The function g' is n times differentiable and vanishes at $n+1$ distinct points on (a, b).

(iii) If $1 \leq r \leq n+1$ then the function $g^{(r)}$ is $n+1-r$ times differentiable and vanishes at $n+2-r$ distinct points on (a, b).

(iv) There exists a $\zeta \in (a, b)$ such that $g^{(n+1)}(\zeta) = 0$.

(v) There exists a $\zeta \in (a, b)$ such that

$$f^{(n+1)}(\zeta) = (n+1)! E(t) \prod_{k=0}^{n} (t - x_k)^{-1}.$$

Part (v) of Lemma 2.5 gives us the required estimate.

Theorem 2.6. *Suppose $f : [a, b] \to \mathbb{R}$ is $n+1$ times differentiable and $|f^{(n+1)}(x)| \leq M$ for all $x \in [a, b]$. If $x_0, x_1, \dots, x_n$ are distinct points of $[a, b]$ and P is the unique polynomial of degree n or less such that $P(x_j) = f(x_j)$ $[0 \leq j \leq n+1]$ then*

$$|P(t) - f(t)| \leq \frac{M}{(n+1)!} \prod_{k=0}^{n} (t - x_k).$$

In order to exploit the inequality of Theorem 2.6 fully we need a polynomial $\prod_{k=0}^{n}(t-x_k)$ which is small for all $t \in [a,b]$. Such a polynomial was found by Tchebychev. We start by recalling De Moivre's theorem.

Theorem 2.7. (De Moivre's Theorem.) *If θ is real and n is a positive integer*

$$\cos n\theta + i \sin n\theta = (\cos\theta + i\sin\theta)^n.$$

Taking real parts in the De Moivre formula we obtain the following result.

Lemma 2.8. *There is a real polynomial of degree n such that*

$$T_n(\cos\theta) = \cos n\theta \text{ for all real } \theta.$$

If $n \geq 1$,

(i) $T_{n+1}(t) = 2tT_n(t) - T_{n-1}(t)$ for all t.
(ii) The coefficient of t^n in $T_n(t)$ is 2^{n-1}.
(iii) $|T_n(t)| \leq 1$ for all $|t| \leq 1$.
(iv) T_n has n distinct roots all in $(-1,1)$.

We call T_n the nth Tchebychev polynomial.

Theorem 2.9. *Let $x_0, x_1, \dots, x_n$ be the $n+1$ roots of the $n+1$st Tchebychev polynomial. If $f : [-1,1] \to \mathbb{R}$ is $n+1$ times differentiable, $|f^{(n+1)}(x)| \leq M$ for all $x \in [-1,1]$ and P is the unique polynomial of degree n or less such that $P(x_j) = f(x_j)$ $[0 \leq j \leq n+1]$ then*

$$|P(t) - f(t)| \leq \frac{M}{2^n (n+1)!}$$

for all $t \in [-1,1]$.

The practical use of this result is restricted by the fact that the size of the nth derivative of apparently well behaved function may increase explosively as n in increases.

3 Cathode ray tubes and cellars

The path of an electron of mass m and charge e in an electric field $\mathbf{E}$ and a magnetic field $\mathbb{B}$ is given by

$$m\ddot{\mathbf{x}} = e(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}).$$

In the simple case when the fields are constant and $\mathbf{E} = (O, me^{-1}E, 0)$, $\mathbf{B} = (0, 0, me^{-1}B)$ the equation can be written coordinate-wise as

$$\begin{aligned}\ddot{x} &= B\dot{y}\\ \ddot{y} &= E - B\dot{x}\\ \ddot{z} &= 0.\end{aligned}$$

It is one thing to write down a set of equations like this. It is quite another thing to solve them. We can obtain $z = z_0 + w_0 t$ for some constants z_0 and w_0 from the third equation. We can simplify the first two equations by setting $u = \dot{x}$ and $v = \dot{y}$ to obtain

$$\begin{aligned}\dot{u} &= Bv\\ \dot{v} &= E - Bu.\end{aligned}$$

If we set $U = u + B^{-1}E$ and $V = v$ these equations take the simpler form

$$\begin{aligned}\dot{U} &= BV\\ \dot{V} &= -BU\end{aligned}$$

but we still have to solve them.

To do this we introduce complex numbers by adding i times the second equation to the first to obtain the single equation

$$\frac{d}{dt}(U + iV) = B(V - iU).$$

If we set $\Phi = U + iV$ this equation takes the form

$$\dot{\Phi} = -iB\Phi.$$

This is an equation that we can solve to obtain

$$\Phi(t) = Ae^{-iBt}$$

where A is a fixed complex number.

How does this solution fit in with our original problem. recall that we can write

$$A = r \exp i\alpha$$

where r is real and positive and α is real. We can thus write

$$\Phi(t) = re^{i(\alpha - Bt)}.$$

Taking real and imaginary parts this gives us

$$U = r\cos(Bt - \alpha)$$
$$V = -r\sin(Bt - \alpha).$$

Since $\dot{x} = U - B^{-1}E$ and $\dot{y} = V$ we obtain

$$x = x_0 - R\cos(Bt - \alpha)$$
$$y = y_0 - R\sin(Bt - \alpha)$$

where $R = r/B$ and x_0 and y_0 are constants.

Putting everything together we see that

$$x = x_0 - R\cos(Bt - \alpha) - B^{-1}Et$$
$$y = y_0 - R\sin(Bt - \alpha)$$
$$z = z_0 + w_0 t$$

so that the electron follows a spiral path. We note that small perturbations of the electron will have little effect on its path.

Here is another example of the use of complex numbers which brings us closer to the main topic of this course. Consider the temperature θ at a depth x in the ground. The equation for heat conduction is

$$\frac{\partial\theta}{\partial t} = K\frac{\partial^2\theta}{\partial x^2}.$$

It is natural to seek a solution for the case $\theta(0,t) = A\cos\omega t$ $[\omega > 0]$ in which the surface is periodically warmed and cooled (consider the surface temperature during a day). Let us try and solve the related complex problem

$$\frac{\partial Y}{\partial t} = K\frac{\partial^2 Y}{\partial x^2}. \qquad \bigstar$$

$Y(0,t) = Ae^{i\omega t}$. A natural guess is to try $Y(x,t) = f(x)e^{i\omega t}$. Substitution in $\bigstar$ yields

$$Kf''(x) = i\omega f(x)$$

and this in turn shows that

$$f(x) = a_1 e^{-\alpha x} + a_2 e^{\alpha x}$$

with

$$\alpha = (\omega/K)^{1/2} e^{i\pi/4} = \left(\frac{\omega}{K}\right)^{1/2} \frac{1+i}{2^{1/2}}$$

(where we take positive square roots of positive numbers).

Now we observe that $e^{-\alpha x} \to 0$ as $x \to \infty$ but $|e^{\alpha x}| \to \infty$. Thus the only physically plausible solutions for f will have $a_2 = 0$. Our initial guess thus gives

$$Y(x,t) = Ae^{-\alpha x} e^{i\omega t} = A \exp\left(-\left(\frac{\omega}{2K}\right)^{1/2} x\right) \exp\left(i\left(\omega t - \left(\frac{\omega}{2K}\right)^{1/2} x\right)\right)$$

as the solution of the complex problem. Taking real parts we obtain a solution for our original problem

$$\theta(x,t) = A \exp\left(\left(-\frac{\omega}{2K}\right)^{1/2} x\right) \cos\left(\omega t - \left(\frac{\omega}{2K}\right)^{1/2} x\right).$$

We can read off all sorts of interesting facts from this solution. First we note that the effects of periodic heating drop off exponentially with depth. Thus the annual heating and cooling of the arctic surface leaves the permafrost unaffected. We note also that the typical length in the exponential decrease is $(2K/\omega)^{1/2}$ so that low frequency effects are longer range than high frequency. The effects of daily heating only extend for 10's of centimetres below the surface but those of annual heating extend a few metres. Since a similar equation governs the penetration of radio-waves in water submarines can only be contacted by very low frequency radio waves. For similar reasons it would not make sense to use high frequencies in a microwave oven. It is worth noting the time lag of $(\omega/2K)^{1/2}$ which means that, for example, the soil temperature at a depth of about 2 metres is higher in winter than in summer.

4 Radars and such-like

We work in the (x,y) plane. Consider an array of $2N+1$ radio transmitters broadcasting at frequency ω. Let the kth transmitter be at $(0, kl)$ $[k = -N,\ -N+1,\ \ldots,\ 0,\ 1,\ \ldots,\ N]$. It is reasonable to take the signal at (x,y) due to the kth transmitter to be

$$A_k r_k^{-2} \exp(i(\omega t - \lambda^{-1} r_k - \phi_k))$$

where λ is the wavelength (thus $\omega\lambda = c$ the speed of light) and $r_k^2 = x^2 + (y - kl)^2$. The total signal at (x, y) is

$$S(x, y, t) = \sum_{k=-N}^{N} A_k r_k^{-2} \exp(i(\omega t - \lambda^{-1} r_k - \phi_k)).$$

Lemma 4.1. *If $x = R\cos\theta$, $y = R\sin\theta$ where R is very large then to a very good approximation*

$$S(R\cos\theta, R\sin\theta, t) = R^{-2} \exp(i(\omega t - \lambda R) Q(u)$$

where

$$Q(u) = \sum_{k=-N}^{N} A_k \exp(i(ku - \phi_k)),$$

and $u = \lambda^{-1} l \sin\theta$.

In the discussion that follows we use the notation of Lemma 4.1 and the discussion that preceded it. We set

$$P(u) = \sum_{k=-N}^{N} A_k \exp(iku),$$

that is $P = Q$ with $\phi_k = 0$ for all k. Since we could take the A_k to be complex, there was no real increase in generality in allowing $\phi_k \neq 0$ but I wished to make the following points.

Lemma 4.2. *(i) Given θ_0 we can find ϕ_k such that*

$$Q(u) = P(u - u_0).$$

(ii) $S(-x, -y, t) = S(x, y, t)$.
(iii) If $l > \lambda\pi/2$ then there exist $0 < \theta_1 < \pi/2$ such that

$$S(R\cos\theta_1, R\sin\theta_1, t) = S(0, R, t).$$

Bearing in mind that the equations governing the reception of signal at (x, y) transmitted from our array are essentially the same as those governing the reception of signal at our array, Lemma 4.2 (i) talks about electronic stearability of radar beams, and Lemma 4.2 (ii) and (iii) deal with ambiguity. It is worth noting that in practice the signal received by a radar corresponds to $|Q(u)|$.

Let us look at $P(u)$ in two interesting cases.

Lemma 4.3. *(i) If* $A_k = (2N+1)^{-1}$ *then, writing* $P_{N,l}(u) = P(u)$,

$$P_{N,l}(u) = \frac{1}{2N+1}\frac{\sin((N+\frac{1}{2})u)}{\sin(\frac{1}{2}u)}.$$

(ii) With the notation of (i)

$$P_{N,a/N}((a/N)^{-1}v) \to \frac{\sin av}{2av}$$

as $N \to \infty$.

Lemma 4.3 is usually interpreted as saying that a radar cannot discriminate between two targets if their angular distance is of the order of the size of λ/a where λ is the wave length used and a is the length of the array. It is natural to ask if a cleverer choice of A_k might enable us to avoid this problem. We shall see that, although the choice in Lemma 4.3 may not be the best, there is no way of avoiding the λ/a rule.

5 Towards rigour

I hope it is obvious that, so far, we have made no attempt at rigour. However, the deeper study of Fourier analysis makes little sense unless it is pursued rigorously. Much of what is often called 'a second course in analysis' was invented to aid the rigorisation of Fourier analysis and related topics.

Although I have tried to make the course accessible to that part of my audience unfamiliar with the ideas that follow, some parts will only become fully rigorous for those familiar with the following ideas. *If you are not familiar with these ideas do not worry and do not spend much time thinking about them. It is more important to reflect on the ideas of Fourier analysis and leave the details until later.* However, you should try to understand the notion of the uniform norm given in Definition 5.3.

The discussion that follows is thus intended to jog the memories of those who already know these ideas and (apart from Definition 5.3) may be ignored by the others.

Lemma 5.1. *If* $f : [a,b] \to \mathbb{R}$ *is continuous on the closed bounded interval* $[a,b]$ *then* f *is bounded and attains its bounds. In other words, we can find* $x_1, x_2 \in [a,b]$ *such that*

$$f(x_1) \geq f(x) \geq f(x_2) \text{ for all } x \in [a,b].$$

Lemma 5.2. *If $f : [a, b] \to \mathbb{C}$ is continuous then we can find an $x_1 \in [a, b]$ such that*

$$|f(x_1)| \geq |f(x)|.$$

Definition 5.3. *Using the notation of Lemma 5.2 we set*

$$\|f\|_\infty = |f(x_1)|.$$

In other words $\|f\|_\infty$ is the least K such that

$$|f(x)| \leq K \text{ for all } x \in [a, b].$$

Note that $\|f - g\|_\infty$ may be considered as the (or, more properly, a) distance between two continuous functions f and g.

We shall use various results about the uniform norm of which the following is the most important. (It is equivalent to the results known as 'the general principle of uniform convergence' and 'the completeness of the uniform norm'.)

Lemma 5.4. *If the functions $f_n : [a, b] \to \mathbb{C}$ are continuous and $\sum_{n=1}^\infty \|f_n\|_\infty$ converges, then there exists a continuous function $f : [a, b] \to \mathbb{C}$ such that*

$$\left\| \sum_{n=1}^{N} f_n - f \right\|_\infty \to 0$$

as $N \to \infty$.

6 Why is there a problem?

We now turn to to the question of whether Fourier expansions are always possible. It turns out to be simplest to work on the circle $\mathbb{T} = \mathbf{R}/2\pi\mathbb{Z}$ that is to work 'modulo 2π' so that $x + 2n\pi = x$. We ask whether a continuous function $f : \mathbb{T} \to \mathbb{C}$ can be represented in the form

$$f(t) \stackrel{?}{=} \sum_{n=-\infty}^{n=\infty} a_n \exp int.$$

Since

$$\frac{1}{2\pi} \int_{\mathbb{T}} (\exp int)(\exp -imt)\, dt = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{otherwise,} \end{cases}$$

the same arguments as we used in Lemma 1.1 show that we should ask whether

$$f(t) \stackrel{?}{=} \sum_{n=-\infty}^{n=\infty} \hat{f}(n) \exp int. \qquad \bigstar$$

where

$$\hat{f}(n) = \frac{1}{2\pi} \int_{\mathbb{T}} \exp(-int) f(t)\, dt.$$

Over the last two centuries we have learnt that the formula ★ can be interpreted in many different ways and that each way gives rise to new set of questions and answers but for the moment let us take the most obvious interpretation and ask whether

$$\sum_{n=-N}^{n=N} \hat{f}(n) \exp int \xrightarrow{?} f(t)$$

as $N \to \infty$ for each $t \in \mathbb{T}$?

Observe that

$$\begin{aligned} \sum_{n=-N}^{n=N} \hat{f}(n) \exp int &= \sum_{n=-N}^{n=N} \frac{1}{2\pi} \int_{\mathbb{T}} \exp(-inx) f(x)\, dx \exp(int) \\ &= \frac{1}{2\pi} \int_{\mathbb{T}} \sum_{n=-N}^{n=N} \exp(in(t-x)) f(x)\, dx. \end{aligned}$$

The same algebra that we used when considering the radar problem now gives us the result of the next lemma.

Lemma 6.1. (Dirichlet's kernel.) *(i) If $f : \mathbb{T} \to \mathbb{C}$ is continuous then*

$$\sum_{n=-N}^{n=N} \hat{f}(n) \exp int = \frac{1}{2\pi} \int_{\mathbb{T}} D_N(t-x) f(x)\, dx$$

where

$$D_N(s) = \sum_{n=-N}^{n=N} \exp(ins).$$

(ii) We have

$$D_N(s) = \begin{cases} 2N+1 & \text{if } s = 0, \\ \frac{\sin((N+\frac{1}{2})s)}{\sin(\frac{1}{2}s)} & \text{otherwise.} \end{cases}$$

(iii) $\displaystyle \frac{1}{2\pi} \int_{\mathbb{T}} D_N(x)\, dx = 1.$

We call D_N the Dirichlet kernel.

If we look at the graphs of $D_N(s)$ and $D_N(t-x)f(x)$ (where t is fixed) we see that $\frac{1}{2\pi}\int_{\mathbb{T}} D_N(t-x)f(x)\,dx$ may indeed tend to $f(t)$ as $N\to\infty$ but that, if it does so, it appears that this is the result of some quite complicated cancellation.

In order to make this statement more precise we need results which you may already know.

Lemma 6.2. *(i) If $f:[1,\infty)\to\mathbb{R}$ is a decreasing function then*

$$\sum_{n=1}^{N-1} f(n) \geq \int_1^N f(x)\,dx \geq \sum_{n=2}^{N} f(n).$$

(ii) $\displaystyle\frac{1}{\log N}\sum_{n=1}^{N}\frac{1}{n}\to 1$ *as* $N\to\infty$.

Lemma 6.3. *If $0\leq x\leq \pi/2$ then*

$$\frac{2x}{\pi}\leq \sin x\leq x.$$

Observing that

$$\begin{aligned}
\frac{1}{2\pi}\int_{\mathbb{T}}|D_N(x)|\,dx &\geq \frac{1}{\pi}\sum_{r=1}^{2N}\int_{r\pi/(2N+1)}^{(r+1)\pi/(2N+1)}\left|\frac{\sin\frac{(2N+1)x}{2}}{\sin\frac{x}{2}}\right|\,dx\\
&\geq \frac{1}{\pi}\sum_{r=1}^{2N}\int_{r\pi/(2N+1)}^{(r+1)\pi/(2N+1)}\frac{|\sin\frac{(2N+1)x}{2}|}{x}\,dx\\
&\geq \frac{1}{\pi}\sum_{r=1}^{2N}\frac{2N+1}{r+1}\int_{r\pi/(2N+1)}^{(r+1)\pi/(2N+1)}\left|\sin\frac{(2N+1)x}{2}\right|\,dx.,
\end{aligned}$$

we obtain the next lemmas. Here and elsewhere we adopt the the abbreviation

$$S_N(f,t)=\sum_{n=-N}^{n=N}\hat{f}(n)\exp int.$$

Lemma 6.4. *There exists a constant $A>0$ such that*

$$\frac{1}{\log N}\left(\frac{1}{2\pi}\int_{\mathbb{T}}|D_N(x)|\,dx\right)\geq A$$

for all $N\geq 1$.

Lemma 6.5. *There exists a constant $B > 0$ such that, given any $N \geq 1$ we can find a continuous function $f : \mathbb{T} \to \mathbb{R}$ with $\|f\|_\infty \leq 1$ and*

$$|S_N(f, 0)| \geq B \log N.$$

It thus comes as no surprise that the following theorem holds.

Theorem 6.6. *There exists a continuous function $f : \mathbb{T} \to \mathbb{R}$ such that $S_N(f, 0)$ fails to converge as $N \to \infty$.*

The full details of the proof require a good grasp of uniform convergence so I leave them as an exercise to be done (if it is done at all) after completion of the next section.

7 Fejér's theorem

It is, of course, true that, as we shall see later, the Fourier sum $S_N(f, t) \to f(t)$ for all sufficiently well behaved functions f but the fact that this result fails for some continuous f remained a serious bar to progress until the beginning of the 20th century. Then a young Hungarian mathematician realised that, although Fourier sums might behave badly, their averages

$$\sigma_N(f, t) = (N + 1)^{-1} \sum_{m=0}^{N} S_m(f, t) = \sum_{r=-N}^{N} \frac{N + 1 - |r|}{N + 1} \hat{f}(r) \exp irt$$

behave much better. (We call $\sigma_N(f, t)$ the Nth Fejér sum.)

We use the same procedure to study Fejér sums as we did Fourier sums. The following algebraic identity plays a very useful role.

$$\left(\sum_{r=0}^{N} \exp \left(i(r - \frac{N}{2})s \right) \right)^2 = \sum_{r=-N}^{N} (N + 1 - |r|) \exp irt.$$

The next result and its proof should be compared carefully with Lemma 6.1 and its proof.

Lemma 7.1. (Fejér's kernel.) *(i) If $f : \mathbb{T} \to \mathbb{C}$ is continuous then*

$$\sigma_N(f, t) = \sum_{n=-N}^{N} \frac{N + 1 - |n|}{N + 1} \hat{f}(n) \exp int = \frac{1}{2\pi} \int_{\mathbb{T}} K_N(t - x) f(x)\, dx$$

where

$$K_N(s) = \sum_{n=-N}^{n=N} \frac{N + 1 - |n|}{N + 1} \exp(ins).$$

(ii) We have

$$K_N(s) = \begin{cases} N+1 & \text{if } s = 0, \\ \frac{1}{N+1}\left(\frac{\sin(\frac{N+1}{2}s)}{\sin(\frac{1}{2}s)}\right)^2 & \text{otherwise.} \end{cases}$$

(iii) $\displaystyle\frac{1}{2\pi}\int_{\mathbb{T}} K_N(x)\,dx = 1.$

(iv) $K_n(s) \geq 0$ *for all* s.

(v) If $\eta > 0$ *then* $K_n \to 0$ *uniformly for* $|t| \geq \eta$ *as* $n \to \infty$.

We call K_N the Fejér kernel. Conditions (iv) and, to a lesser extent, (v) give the key differences between the Dirichlet and the Fejér kernels. If we look at the graphs of $K_N(s)$ and $K_N(t-x)f(x)$ (where t is fixed) we see that $\frac{1}{2\pi}\int_{\mathbb{T}} K_N(t-x)f(x)\,dx$ will indeed tend to $f(t)$ as $N \to \infty$ without any need for cancellation.

Using Lemma 7.1 we see that, if $0 < \eta < \pi$

$$\begin{aligned}
|\sigma_N(f,t) - f(t)| &\leq \left|\frac{1}{2\pi}\int_{\mathbb{T}} K_N(t-x)f(x)\,dx - f(t)\right| \\
&= \left|\frac{1}{2\pi}\int_{\mathbb{T}} K_N(x)f(t-x)\,dx - f(t)\right| \\
&= \left|\frac{1}{2\pi}\int_{\mathbb{T}} (K_N(x)f(t-x) - f(t))\,dx\right| \\
&\leq \frac{1}{2\pi}\int_{\mathbb{T}} K_N(x)|f(t-x) - f(t)|\,dx \\
&= \frac{1}{2\pi}\int_{|x|\leq\eta} K_N(x)|f(t-x) - f(t)|\,dx \\
&\quad + \frac{1}{2\pi}\int_{|x|>\eta} K_N(x)|f(t-x) - f(t)|\,dx \\
&\leq \sup_{|x|\leq\eta}|f(t-x) - f(t)|\frac{1}{2\pi}\int_{|x|\leq\eta} K_N(x)\,dx \\
&\quad + \sup_{|x|\geq\eta} K_N(x)\frac{1}{2\pi}\int_{|x|>\eta}|f(t-x) - f(t)|\,dx \\
&\leq \sup_{|x|\leq\eta}|f(t-x) - f(t)| + 2\|f\|_\infty \sup_{|x|\geq\eta} K_N(x).
\end{aligned}$$

This calculations immediately gives us the required result.

Theorem 7.2. *If* $f : \mathbb{T} \to \mathbb{C}$ *is continuous, then* $\sigma_N(f,t) \to f(t)$ *as* $N \to \infty$ *for each* $t \in \mathbb{T}$.

A little thought gives a still stronger and more useful theorem.

Theorem 7.3. (Fejér's theorem.) *If $f : \mathbb{T} \to \mathbb{C}$ is continuous then*

$$\|\sigma_N(f) - f\|_\infty \to 0$$

as $N \to \infty$.

Here $\sigma_N(f)(t) = \sigma_N(f, t)$.
Fejér's theorem has many important consequences.

Theorem 7.4. (Uniqueness.) *If $f, g : \mathbb{T} \to \mathbb{C}$ are continuous and $\hat{f}(n) = \hat{g}(n)$ for all n then $f = g$.*

Theorem 7.5. *If $f : \mathbb{T} \to \mathbb{C}$ is continuous and $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|$ converge then $\|S_N(f) - f\|_\infty \to 0$ as $N \to \infty$.*

The steps in the proof of Theorem 7.5 are set out in the next lemma.

Lemma 7.6. *Suppose that $f : \mathbb{T} \to \mathbb{C}$ is continuous and $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|$ converges.*
(i) $\sum_{n=-N}^{N} \hat{f}(n) \exp int$ converges uniformly to a continuous function g.
(ii) The function g of (i) satisfies $\hat{g}(n) = \hat{f}(n)$ for all n and so by the uniqueness of the Fourier coefficients $g = f$.

8 The trigonometric polynomials are uniformly dense

The 20th century made it clear that for many purposes convergence is less important than approximation. Fejér's theorem tells us that the trigonometric polynomials are uniformly dense in the continuous functions.

Theorem 8.1. *(i) If $f : \mathbb{T} \to \mathbb{C}$ is continuous then, given any $\epsilon > 0$ we can find a trigonometric polynomial*

$$P(t) = \sum_{n=-N}^{N} a_n \exp int$$

such that $\|P - f\|_\infty < \epsilon$.
(ii) If $g : \mathbb{T} \to \mathbb{R}$ is continuous then, given any $\epsilon > 0$ we can find a real trigonometric polynomial Q such that $\|Q - f\|_\infty < \epsilon$.

Here is a typical use of this result.

Theorem 8.2. *Suppose that $f : \mathbb{T} \to \mathbb{C}$ is continuous.*

(i) $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2$ converges and

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 = \frac{1}{2\pi} \int_{\mathbb{T}} |f(t)|^2 \, dt.$$

(ii) The expression

$$\frac{1}{2\pi} \int_{\mathbb{T}} \left| f(t) - \sum_{n=-N}^{N} a_n e^{int} \right|^2 \, dt$$

has a unique minimum (over choices of a_n) when $a_n = \hat{f}(n)$.

(iii) We have

$$\frac{1}{2\pi} \int_{\mathbb{T}} \left| f(t) - \sum_{n=-N}^{N} \hat{f}_n e^{int} \right|^2 \, dt \to 0$$

as $N \to \infty$.

In other words the Fourier sums are the best 'mean square' approximations and converge in 'mean square' to the original function. The following pretty formula (Parseval's identity) can be deduced from Theorem 8.2 or proved by using the same ideas.

Lemma 8.3. *Suppose that f, $g : \mathbb{T} \to \mathbb{C}$ are continuous. Then*

$$\sum_{n=-N}^{N} \hat{f}(n)\hat{g}(n)^* \to \frac{1}{2\pi} \int_{\mathbb{T}} f(t)g(t)^* \, dt$$

as $N \to \infty$.

Here is a beautiful application due to Weyl of Theorem 8.1. If x is real let us write $\langle x \rangle$ for the fractional part of x, that is, let us write

$$\langle x \rangle = x - [x].$$

Theorem 8.4. *If α is an irrational number and $0 \le a \le b \le 1$, then*

$$\frac{\operatorname{card}\{1 \le n \le N \mid \langle n\alpha \rangle \in [a, b]\}}{N} \to b - a$$

as $N \to \infty$. The result is false if α is rational.

The proof of Weyl's theorem can be split into stages as follows.

Lemma 8.5. *The following statements are equivalent.*

(i) If α is an irrational number and $0 \leq a \leq b \leq 1$, then

$$\frac{\operatorname{card}\{1 \leq n \leq N \mid \langle n\alpha\rangle \in [a,b]\}}{N} \rightarrow b-a$$

as $N \rightarrow \infty$.

(ii) If α is an irrational number and $0 \leq a \leq b \leq 2\pi$, then

$$\frac{\operatorname{card}\{1 \leq n \leq N \mid 2\pi n\alpha \in [a,b]\}}{N} \rightarrow \frac{b-a}{2\pi}$$

as $N \rightarrow \infty$.

(iii) If α is an irrational number and $f : \mathbb{T} \rightarrow \mathbb{C}$ is continuous then

$$\sum_{n=0}^{N} f(2\pi n\alpha) \rightarrow \frac{1}{2\pi}\int_{\mathbb{T}} f(x)\,dx$$

as $N \rightarrow \infty$.

Lemma 8.6. *If α is an irrational number and $f : \mathbb{T} \rightarrow \mathbb{C}$ is continuous let us write*

$$J_N f = \sum_{n=0}^{N} f(2\pi n\alpha) \text{ and } If = \frac{1}{2\pi}\int_{\mathbb{T}} f(x)\,dx.$$

(i) J_N and I are linear maps from $C(\mathbb{T})$ to $\mathbb{C}$ with $|J_N f| \leq \|f\|_\infty$ and $|If| \leq \|f\|_\infty$.

(ii) If we define $e_n : \mathbb{T} \rightarrow \mathbb{R}$ by $e_n(t) = \exp int$ then

$$J_N e_n \rightarrow I e_n$$

as $N \rightarrow \infty$ for all $n \in \mathbb{Z}$.

Much of this work can be extended to more dimensions. I shall probably leave the proofs of the following results as exercises for those who want to do them.

Lemma 8.7. *If we define $\tilde{K} : \mathbb{T}^m \rightarrow \mathbb{R}$ by $\tilde{K}(t_1, t_2, \ldots, t_m) = \prod_{j=1}^{m} K(t_j)$ then we have the following results.*

(i) $\frac{1}{(2\pi)^m}\int_{\mathbb{T}^m} \tilde{K}_n(\mathbf{t})\,d\mathbf{t} = 1$.

(ii) If $\eta > 0$ then $\int_{|\mathbf{t}|\geq\eta} \tilde{K}_n(\mathbf{t})\,d\mathbf{t} \rightarrow 0$ as $n \rightarrow \infty$.

(iii) $\tilde{K}_n(\mathbf{t}) \geq 0$ *for all* $\mathbf{t}$.
(iv) $\tilde{K}_n$ *is a (multidimensional) trigonometric polynomial.*
(v) If $f : \mathbb{T}^m \to \mathbb{C}$ *is continuous then*

$$\frac{1}{(2\pi)^m} \int_{\mathbb{T}^m} \tilde{K}_N(\mathbf{t} - \mathbf{x}) f(\mathbf{x})\, d\mathbf{x} \to f(\mathbf{t})$$

uniformly as $N \to \infty$.

Lemma 8.8. *If* $f : \mathbb{T}^m \to \mathbb{C}$ *is continuous then, given any* $\epsilon > 0$ *we can find a trigonometric polynomial*

$$P(t) = \sum_{|j(r)| \leq N} a_{j(1),j(2),\dots,j(m)} \exp\left(i \sum_{r=1}^{m} j(r) t_r \right)$$

such that $\|P - f\|_\infty < \epsilon$.

We immediately obtain a striking generalisation of Weyl's theorem (Theorem 8.4).

Lemma 8.9. *Suppose that* $\alpha_1,\ \alpha_2,\ \dots,\ \alpha_m$ *are real numbers. A necessary and sufficient condition that*

$$\frac{\operatorname{card}\{1 \leq n \leq N \mid (\langle n\alpha_1\rangle, \langle n\alpha_2\rangle, \dots, \langle n\alpha_m\rangle \in \prod_{j=1}^m [a_j, b_j]\}}{N} \to \prod_{j=1}^{m} (b_j - a_j)$$

as $N \to \infty$ *whenever* $0 \leq a_j \leq b_j \leq 1$ *is that*

$$\sum_{j=1}^{m} n_j \alpha_j \notin \mathbb{Z} \text{ for integer } n_j \text{ not all zero.} \qquad \bigstar$$

If $\alpha_1,\ \alpha_2,\ \dots,\ \alpha_m$ satisfy ★ we say that they are independent. The multidimensional version of Weyl's theorem has an important corollary.

Theorem 8.10. (Kronecker's theorem.) *Suppose that* $\alpha_1,\ \alpha_2,\ \dots,\ \alpha_m$ *are independent real numbers. Then given real numbers* $\beta_1,\ \beta_2,\ \dots,\ \beta_m$ *and* $\epsilon > 0$ *we can find integers* $N,\ r_1,\ r_2,\ \dots,\ r_m$ *such that*

$$|N\alpha_j - \beta_j - r_j| < \epsilon$$

for each $1 \leq j \leq M$.

The result is false if $\alpha_1,\ \alpha_2,\ \dots,\ \alpha_m$ *are not independent.*

9 First thoughts on Fourier transforms

We have seen that it very useful to look at functions $f : \mathbb{T} \to \mathbb{C}$ in the form

$$f(t) = \sum_{n=-\infty}^{n=\infty} a_n \exp int,$$

that is functions $f : \mathbb{T} \to \mathbb{C}$ which are weighted sums of simple waves (exponentials). However, many problems involve functions $g : \mathbb{R} \to \mathbb{C}$ so it is natural to investigate those $g : \mathbb{R} \to \mathbb{C}$ which are weighted *integrals* of simple waves (exponentials), that which can be written

$$g(t) = \int_{-\infty}^{\infty} G(\lambda) \exp(-i\lambda t)\, d\lambda.$$

(The minus sign is inserted for consistency with our other conventions.) We say that g is the *Fourier transform* of $G : \mathbb{R} \to \mathbb{C}$.

Unfortunately the rigorous treatment of 'integrals over an infinite range' raises certain problems.

Example 9.1. *(i) If we set*

$$a_{rs} = \begin{cases} 2^{-r} & \text{if } 2^r + 1 \leq s \leq 2^{r+1}, \\ -2^{-r-1} & \text{if } 2^{r+1} + 1 \leq s \leq 2^{r+2}, \\ 0 & \text{otherwise,} \end{cases}$$

then

$$\sum_{r=1}^{\infty} \left(\sum_{s=1}^{\infty} a_{rs} \right) = 0 \neq 1 = \sum_{s=1}^{\infty} \left(\sum_{r=1}^{\infty} a_{rs} \right).$$

(ii) We can find a continuous function $f : \mathbb{R}^2 \to \mathbb{R}$ with $f(\mathbf{x}) \to 0$ as $\|\mathbf{x}\| \to \infty$ with all the integrals in the next inequality well defined but

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x, y)\, dx \right) dy \neq \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x, y)\, dy \right) dx.$$

Provided the functions fall away sufficiently fast towards infinity the problem raised by the previous example will not occur.

Lemma 9.2. *If $f : \mathbb{R}^2 \to \mathbb{R}$ is such that we can find a constant A with*

$$|f(x, y)| \leq \frac{A}{(1 + x^2)(1 + y^2)}$$

for all $(x, y) \in \mathbb{R}^2$ *then all the integrals in the next equality are well defined and*

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x, y)\, dx \right) dy = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x, y)\, dy \right) dx.$$

In developing the theory of Fourier transforms we shall assume various theorems which depend on reasonably rapid decay towards infinity.

10 Fourier transforms

We now start the study of Fourier transforms in earnest.

Definition 10.1. *If* $f : \mathbb{R} \to \mathbb{C}$ *is reasonably well behaved, we define*

$$\hat{f}(\lambda) = \int_{-\infty}^{\infty} f(t) e^{-i\lambda t}\, dt,$$

and call the function $\hat{f} : \mathbb{R} \to \mathbb{C}$ *the* Fourier transform*.*

As we said in the previous section we require a certain amount of good behaviour from f. The condition that f, f' and f'' are continuous and $t^2 f(t),\ t^2 f'(t),\ t^2 f''(t) \to 0$ as $|t| \to \infty$ are amply sufficient for our purpose (much less is required, but there always has to be some control over behaviour towards infinity).

The following results form part of the grammar of Fourier transforms.

Lemma 10.2. *(i) If* $a \in \mathbb{R}$*, let us write* $f_a(t) = f(t - a)$*. Then*

$$\hat{f}_a(\lambda) = e^{-ia\lambda} \hat{f}(\lambda).$$

(Translation on one side gives phase change on other.)

(ii) If $K \in \mathbb{R}$ *and* $K > 0$*, let us write* $f_K(t) = f(Kt)$*. Then*

$$\hat{f}_K(\lambda) = K^{-1} \hat{f}(\lambda / K).$$

(Narrowing on one side gives broadening on the other.)

(iii) $\hat{f}(\lambda)^* = (f^*)\hat{\ }(-\lambda)$*.*

(iv) $(\hat{f})'(\lambda) = -i\hat{F}(\lambda)$ *where* $F(t) = tf(t)$*.*

(v) $(f')\hat{\ }(\lambda) = i\lambda \hat{f}(\lambda)$*.*

It is natural to hope that we could obtain results on Fourier transforms as limits (in some sense) of Fourier sums. This is not impossible and we shall see in Section 13 that there is a very elegant link between Fourier transforms and Fourier sums. However, there are technical difficulties and it is more straightforward to start afresh.

We pay particular attention to the Gaussian (or heat, or error) kernel $E(x) = (2\pi)^{-1/2}\exp(-x^2/2)$.

Lemma 10.3. *(i) The Fourier transform of E obeys the partial differential equation*

$$\hat{E}'(\lambda) = -\lambda \hat{E}(\lambda).$$

(ii) $\hat{E}(\lambda) = (2\pi)^{1/2}E(\lambda)$

(The fact that $\hat{E}(0) = 1$ is derived from a formula that is probably known to you. If not, consult Exercise 19.19.)

We use the following neat formula.

Lemma 10.4. *If $f, g : \mathbb{R} \to \mathbb{C}$ are well behaved then*

$$\int_{-\infty}^{\infty} \hat{g}(x)f(x)\,dx = \int_{-\infty}^{\infty} g(\lambda)\hat{f}(\lambda)\,d\lambda.$$

By taking $g(x) = E_R(x) = E(Rx)$ and allowing $R \to 0+$ we obtain an inversion formula.

Lemma 10.5. *If $f : \mathbb{R} \to \mathbb{C}$ is well behaved then*

$$f(0) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \hat{f}(\lambda)\,d\lambda.$$

Using Lemma 10.2, we see that translation gives us our full inversion result.

Theorem 10.6. *If f is well behaved, then $\hat{\hat{f}}(t) = 2\pi f(-t)$.*

(If we write $\mathcal{F}(f) = (2\pi)^{-1/2}\hat{f}$, $Jf(t) = f(-t)$ and $If = f$ then subject to appropriate conditions we obtain $\mathcal{F}^2 = J$ and $\mathcal{F}^4 = I$.)

The inversion formula gives a uniqueness result which is often more useful than the inversion formula itself.

Theorem 10.7. (Uniqueness.) *If f and g are well behaved and $\hat{f} = \hat{g}$ then $f = g$.*

One of the reasons that Fourier analysis works so well is that, although the inversion theorem does require some sort of good behaviour, it turns out that the uniqueness theorem is (almost) endlessly extendable.

Combining the inversion formula with Lemma 10.4 we obtain an elegant analogue of Theorem 8.2.

Lemma 10.8. *If $f : \mathbb{R} \to \mathbb{C}$ is well behaved, then*

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 \, d\lambda.$$

Fourier transforms are closely linked with the important operation of convolution.

Definition 10.9. *If $f,\ g : \mathbb{R} \to \mathbb{C}$ are well behaved, we define their* convolution *$f * g : \mathbb{R} \to \mathbb{C}$ by*

$$f * g(t) = \int_{-\infty}^{\infty} f(t-s)g(s) \, ds.$$

Lemma 10.10. *If f and g are well behaved, $\widehat{f * g}(\lambda) = \hat{f}(\lambda)\hat{g}(\lambda)$.*

For many mathematicians and engineers, Fourier transforms are important because they convert convolution into multiplication and convolution is important because it is transformed by Fourier transforms into multiplication. We shall see that convolutions occur naturally in the study of differential equations. It also occurs in probability theory where the sum $X + Y$ of two independent random variables X and Y with probability densities f_X and f_Y is $f_{X+Y} = f_X * f_Y$. In the next section we outline the connection of convolution with signal processing.

11 Signals and such-like

Suppose we have a black box $\mathcal{K}$. If we feed in a signal $f : \mathbb{R} \to \mathbb{C}$ we will get out a transformed signal $\mathcal{K}f : \mathbb{R} \to \mathbb{C}$. Simple black boxes will have the following properties:

(1) *Time invariance* If $\mathcal{T}_a f(t) = f(t-a)$, then $\mathcal{K}(\mathcal{T}_a f)(t) = (\mathcal{K}f)(t-a)$. In other words, $\mathcal{K}\mathcal{T}_a = \mathcal{T}_a\mathcal{K}$.

(2) *Causality* If $f(t) = 0$ for $t < 0$, then $(\mathcal{K}f)(t) = 0$ for $t < 0$. (The response to a signal cannot precede the signal.)

(3) *Stability* Roughly speaking, the black box should consume rather than produce energy. Roughly speaking, again, if there exists a R such that $f(t) = 0$ for $|t| \geq R$, then we should have $(\mathcal{K}f)(t) \to 0$ as $t \to \infty$. If conditions like this do not apply, both our mathematics and our black box have a tendency to explode. (Unstable systems may be investigated using a close relative of the Fourier transform called the Laplace transform.)

(4) *Linearity* In order for the methods of this course to work, our black box must be linear, that is

$$\mathcal{K}(af + bg) = a\mathcal{K}(f) + b\mathcal{K}(g).$$

(Engineers sometimes spend a lot of effort converting non-linear systems to linear for precisely this reason.)

As our first example of such a system, let us consider the differential equation

$$F''(t) + (a+b)F'(t) + abF(t) = f(t) \qquad \bigstar$$

(where $a > b > 0$), subject to the boundary condition $F(t),\ F'(t) \to 0$ as $t \to -\infty$. We take $\mathcal{K}f = F$.

Before we can solve the system using Fourier transforms we need a preliminary definition and lemma.

Definition 11.1. *The Heaviside function $H : \mathbb{R} \to \mathbb{R}$ is given by*

$$\begin{aligned} H(t) &= 0 \qquad \text{for } t < 0, \\ H(t) &= 1 \qquad \text{for } t \geq 0. \end{aligned}$$

Lemma 11.2. *Suppose that $\Re\alpha < 0$. Then, if we set $e_\alpha(t) = e^{\alpha t}H(t)$, we obtain*

$$\hat{e}_\alpha(\lambda) = \frac{1}{i\lambda - \alpha}.$$

(Some applied mathematicians would leave out the condition $\Re\alpha < 0$ in the lemma just given and most would write $\hat{H}(\lambda) = 1/(i\lambda)$. The study of Laplace transforms reveals why this reckless behaviour does not lead to disaster.)

Lemma 11.3. *The solution $F = \mathcal{K}f$ of*

$$F''(t) + (a+b)F'(t) + abF(t) = f(t) \qquad \bigstar$$

(where $a,\ b > 0$), subject to the boundary condition $F(t),\ F'(t) \to 0$ as $t \to -\infty$, is given by

$$\mathcal{K}f = K \star f \text{ where } K(t) = \frac{e^{-bt} - e^{-at}}{a-b} H(t).$$

Observe that $K(t) = 0$ for $t \leq 0$ and so, if $f(t) = 0$ for $t \leq 0$, we have

$$\begin{aligned} \mathcal{K}f(t) &= K \star f(t) = 0 \qquad \text{for } t \leq 0, \\ \mathcal{K}f(t) &= K \star f(t) = \int_0^t f(s)K(t-s)\,ds \qquad \text{for } t > 0. \end{aligned}$$

Thus $\mathcal{K}$ is indeed causal.

There is another way of analysing black boxes. Let g_n be a sequence of functions such that

(i) $g_n(t) \geq 0$ for all t,

(ii) $\displaystyle\int_{-\infty}^{\infty} g_n(t)\,dt = 1$,

(iii) $g_n(t) = 0$ for $|t| > 1/n$.

In some sense, the g_n 'converge' towards the 'idealised impulse function' δ whose defining property runs as follows.

Definition 11.4. *If $f : \mathbb{R} \to \mathbb{R}$ is a well behaved function then*

$$\int_{-\infty}^{\infty} f(t)\delta(t)\,dt = f(0).$$

If the black box is well behaved we expect $\mathcal{K}g_n$ to converge to some function E. We write

$$\mathcal{K}\delta = E$$

and say that the response of the black box to the delta function is the elementary solution E. Note that, since our black box is causal, $K(t) = 0$ for $t < 0$.

If f is a ordinary function, we define its translate by some real number a to be f_a where $f_a(t) = f(t-a)$. In the same way, we define the translate by a of the delta function by a to be δ_a where $\delta_a(t) = \delta(t-a)$ or, more formally, by

$$\int_{-\infty}^{\infty} f(t)\delta_a(t)\,dt = \int_{-\infty}^{\infty} f(t)\delta(t-a)\,dt = f(a).$$

Since our black box is time invariant, we have

$$\mathcal{K}\delta_a = E_a$$

and, since it is linear,

$$\mathcal{K}\sum_{j=1}^{n} \lambda_j \delta_{a_j}(t) = \sum_{j=1}^{n} \lambda_j E_{a_j}(t).$$

In particular, if F is a well behaved function,

$$\begin{aligned}\mathcal{K}\sum_{j=-MN}^{MN} N^{-1}F(j/N)\delta_{j/N}(t) &= \sum_{j=-MN}^{MN} N^{-1}F(j/N)E_{j/N}(t)\\ &= \sum_{j=-MN}^{MN} N^{-1}F(j/N)E(t-j/N).\end{aligned}$$

Crossing our fingers and allowing M and N to tend to infinity, we obtain

$$\mathcal{K}F(t) = \int_{-\infty}^{\infty} F(s)E(t-s)\,ds,$$

so

$$\mathcal{K}F = F * E.$$

Thus the response of the black box to a signal F is obtained by convolving F with the response of the black box to the delta function. (This is why the acoustics of concert halls are tested by letting off starting pistols.) We now understand the importance of convolution, delta functions and elementary solutions in signal processing and the study of partial differential equations. (The response of the black box to the delta function is often called the Green's function.)

We use two methods to find out what happens in our specific case.

First method Suppose

$$E''(t) + (a+b)E'(t) + abE(t) = \delta(t) \qquad \bigstar$$

and E is well behaved. Then we have

$$E''(t) + (a+b)E'(t) + abE(t) = 0$$

for $t > 0$ so $E(t) = Ae^{-at} + Be^{-bt}$ for $t > 0$ where A and B are constants to be determined. We also have

$$E''(t) + (a+b)E'(t) + abE(t) = 0$$

for $t < 0$ and the condition that $E(t),\ E'(t) \to 0$ as $t \to -\infty$ gives $E(t) = 0$ for $t < 0$. When a bat hits a ball the velocity of the ball changes (almost) instantaneously but the position does not. We thus expect E to be continuous at 0 even though E' is not. The continuity of E gives

$$0 = E(0-) = E(0+) = A + B$$

so $A = -B$ and $E(t) = Ae^{-at} - Ae^{-bt}$ for $t > 0$.

Next we observe that, if $\eta > 0$

$$\int_{-\eta}^{\eta} \big(E''(t) + (a+b)E'(t) + abE(t)\big)\,dt = \int_{-\eta}^{\eta} \delta(t)\,dt = 1$$

so

$$[E'(t) + (a+b)E(t)]_{-\eta}^{\eta} + ab\int_{-\eta}^{\eta} E(t)\,dt = 1.$$

Allowing $\eta \to 0+$ and using the continuity of E we get

$$E'(0+) - E'(0-) = 1$$

so $(b-a)A = 1$ and $A = (b-a)^{-1}$.
Second method Use Fourier transforms in the style of our treatment of Lemma Lemma, big star.

Fortunately both methods give the same result.

Lemma 11.5. *The solution $E = \mathcal{K}\delta$ of*

$$E''(t) + (a+b)E'(t) + abE(t) = \delta(t) \qquad \bigstar$$

(where, $a,\ b > 0$), subject to the boundary condition $E(t),\ E'(t) \to 0$ as $t \to -\infty$, is given by

$$E(t) = \frac{e^{-bt} - e^{-at}}{a-b} H(t).$$

Observe that Lemma 11.5 implies Lemma 11.3 and vice versa.

12 Heisenberg

If we strike a note on the piano the result can not be a pure tone (frequency) since it is limited in time. More generally, as hinted in our discussion of radar signals, we can not shape a function very sharply if it is made up of a limited band of frequencies. (Crudely 'thin functions' have 'fat transforms'.) Their are many different ways of expressing this insight. In this section we give one which has the advantage (and, perhaps, the disadvantage) of being mathematically precise.

We shall need to recall the notion of an inner product.

Definition 12.1. *If V is a vector space over $\mathbb{C}$ we say that the map from V^2 to $\mathbb{C}$ given by $(\mathbf{x},\mathbf{y}) \mapsto \langle \mathbf{x},\mathbf{y}\rangle$ is an inner product if*
(i) $\langle \mathbf{x},\mathbf{x}\rangle \geq 0$ with equality if and only if $\mathbf{x} = \mathbf{0}$,
(ii) $\langle \mathbf{x},\mathbf{y}\rangle = \langle \mathbf{y},\mathbf{x}\rangle^*$,
(iii) $\langle (\lambda\mathbf{x}),\mathbf{y}\rangle = \lambda\langle \mathbf{x},\mathbf{y}\rangle$,
(iv) $\langle (\mathbf{x}+\mathbf{y}),\mathbf{z}\rangle = \langle \mathbf{x},\mathbf{z}\rangle + \langle \mathbf{y},\mathbf{z}\rangle$.

As the audience has probably realised long ago we have met at least two inner products in the study of Fourier analysis.

Lemma 12.2. *(i) If we set*

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} f(t)g(t)^*\,dt$$

then $\langle\ ,\ \rangle$ *is an inner product on* $C(\mathbb{T})$.

(ii) If we set

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(t)g(t)^*\,dt,$$

then $\langle\ ,\ \rangle$ *is an inner product on the space of well behaved functions on* $\mathbb{R}$.

The reason that I draw this to the readers attention is that we need the Cauchy-Schwarz inequality.

Lemma 12.3. (The Cauchy-Schwarz inequality). *If* $\langle\ ,\ \rangle$ *is an inner product on a vector space* V *over* $\mathbb{C}$ *then*

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|.$$

where we write $\|\mathbf{a}\|$ *for the positive square root of* $\langle \mathbf{a}, \mathbf{a}\rangle$

Using Lemma 10.2 (v) (Fourier transform of a derivative), Theorem 8.2 (Parseval), the Cauchy-Schwarz inequality and a certain amount of ingenuity we see that, if f is well behaved.

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty} x^2|f(x)|^2\,dx\int_{-\infty}^{\infty}\lambda^2|\hat{f}(\lambda)|^2\,d\lambda &= \frac{1}{2\pi}\int_{-\infty}^{\infty}|xf(x)|^2\,dx\int_{-\infty}^{\infty}|\lambda\hat{f}(\lambda)|^2\,d\lambda\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}|xf(x)|^2\,dx\int_{-\infty}^{\infty}|\widehat{f'}(\lambda)|^2\,d\lambda\\
&= \int_{-\infty}^{\infty}|xf(x)|^2\,dx\int_{-\infty}^{\infty}|f'(x)|^2\,dx\\
&\geq \left(\int_{-\infty}^{\infty}|xf(x)f'(x)|\,dx\right)^2\\
&\geq \left(\int_{-\infty}^{\infty}\frac{x}{2}(f'(x)f^*(x)+f(x)f^{*\prime}(x))\,dx\right)^2\\
&= \frac{1}{4}\left(\int_{-\infty}^{\infty}x\left(\frac{d}{dx}|f(x)|^2\right)\,dx\right)^2\\
&= \frac{1}{4}\left(\int_{-\infty}^{\infty}|f(x)|^2\,dx\right)^2\\
&= \frac{1}{8\pi}\int_{-\infty}^{\infty}|f(x)|^2\,dx\int_{-\infty}^{\infty}|\hat{f}(\lambda)|^2\,d\lambda
\end{aligned}$$

Rewriting the result we obtain the desired theorem.

Theorem 12.4. (Heisenberg's inequality.) *If $f : \mathbb{R} \to \mathbb{C}$ is well behaved and non-trivial, then*

$$\frac{\int_{-\infty}^{\infty} x^2 |f(x)|^2 \, dx}{\int_{-\infty}^{\infty} |f(x)|^2 \, dx} \frac{\int_{-\infty}^{\infty} \lambda^2 |\hat{f}(\lambda)|^2 \, d\lambda}{\int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 \, d\lambda} \geq \frac{1}{4}.$$

Thus, if f is concentrated near the origin $\hat{f}$ cannot be.

13 Poisson's formula

The circle $\mathbb{T}$ is just the real line $\mathbb{R}$ rolled up. By reflecting on this we are led to a remarkable formula.

Theorem 13.1. (Poisson's formula.) *Suppose that $f : \mathbb{R} \to \mathbb{C}$ is a continuous function such that $\sum_{m=-\infty}^{\infty} |\hat{f}(m)|$ converges and $\sum_{n=-\infty}^{\infty} |f(2\pi n+x)|$ converges uniformly on $[-\pi, \pi]$. Then*

$$\sum_{m=-\infty}^{\infty} \hat{f}(m) = 2\pi \sum_{n=-\infty}^{\infty} f(2\pi n).$$

It is possible to adjust the hypotheses on f in Poisson's formula in various ways though some hypotheses there must be. The following rather simple lemma suffices for many applications.

Lemma 13.2. *If $f : \mathbb{R} \to \mathbb{C}$ is a twice continuously differentiable function such that $\int_{-\infty}^{\infty} |f(x)| \, dx$, $\int_{-\infty}^{\infty} |f'(x)| \, dx$ and $\int_{-\infty}^{\infty} |f''(x)| \, dx$ converge whilst $f'(x) \to 0$ and $x^2 f(x) \to 0$ as $|x| \to \infty$, then f satisfies the conditions of Theorem 13.1.*

To prove Theorem 13.1 we observe that $\sum_{n=-\infty}^{\infty} f(2\pi n + x)$ converges uniformly to a continuous 2π periodic function $g(x)$. We can now define a continuous function $G : \mathbb{T} \to \mathbb{C}$ by setting $G(x) = g(x)$ for $-\pi < x \leq \pi$. We now observe that

$$\begin{aligned} \hat{G}(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) \exp(-imx) \, dx \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \int_{-\pi}^{\pi} f(x + 2n\pi) \exp(-imx) \, dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \exp(-imx) \, dx = \frac{\hat{f}(m)}{2\pi}. \end{aligned}$$

Theorem 7.5 on absolutely convergent Fourier series now gives us the following result.

Lemma 13.3. *Suppose that $f : \mathbb{R} \to \mathbb{C}$ is a continuous function such that $\sum_{m=-\infty}^{\infty} |\hat{f}(m)|$ converges and $\sum_{n=-\infty}^{\infty} |f(2\pi n + x)|$ converges uniformly on $[-\pi, \pi]$. Then*

$$\sum_{m=-\infty}^{\infty} \hat{f}(m) \exp imt = 2\pi \sum_{n=-\infty}^{\infty} f(2\pi n + t).$$

Setting $t = 0$ gives Theorem 13.1.

Exercise 13.4. *Prove Lemma 13.3 from Theorem 13.1.*

Exercise 13.5. *Suppose that f satisfies the conditions of Lemma 13.2. (i) Show that, if $K > 0$, then*

$$K \sum_{m=-\infty}^{\infty} \hat{f}(Km) = 2\pi \sum_{n=-\infty}^{\infty} f(2\pi n/K).$$

What formula do you obtain if $K < 0$?

(ii) By allowing $K \to 0+$ obtain a new proof of the inversion formula

$$f(0) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \hat{f}(\lambda)\, d\lambda.$$

Deduce in the usual way that

$$\hat{\hat{f}}(t) = 2\pi f(-t)$$

for all t.

14 Shannon's theorem

Poisson's formula has a particularly interesting consequence.

Lemma 14.1. *If $g : \mathbb{R} \to \mathbb{C}$ is twice continuously differentiable and $g(t) = 0$ for $|t| \geq \pi$ then g is completely determined by the values of $\hat{g}(m)$ for integer m.*

Taking $g = \hat{f}$ and remembering the inversion formula we obtain the following result.

Theorem 14.2. *If $f : \mathbb{R} \to \mathbb{C}$ is a well behaved function with $\hat{f}(\lambda) = 0$ for $|\lambda| \geq \pi$ then f is determined by its values at integer points.*

The object of this final section is to give a constructive proof of this theorem.

The simplest approach is via the *sinc function*

$$\operatorname{sinc}(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-ix\lambda)\, d\lambda.$$

We state the most immediately useful properties of sinc.

Lemma 14.3. *(i)* $\operatorname{sinc}(0) = 1$,

(ii) $\operatorname{sinc}(n) = 0$ *if* $n \in \mathbb{Z}$ *but* $n \neq 0$.

(We note also that although, strictly speaking, $\widehat{\operatorname{sinc}}(\lambda)$ is not defined for us, since $\int |\operatorname{sinc}(x)|\, dx = \infty$ we are strongly tempted to say that $\widehat{\operatorname{sinc}}(\lambda) = 1$ if $|\lambda| < \pi$ and $\widehat{\operatorname{sinc}}(\lambda) = 0$ if $|\lambda| > \pi$.)

We can, at once, prove that Theorem 14.2 is best possible.

Lemma 14.4. *If* $\epsilon > 0$ *then we can find a well behaved non-zero* f *such that* $\hat{f}(\lambda) = 0$ *for* $|\lambda| > \pi + \epsilon$ *but* $f(n) = 0$ *for all* $n \in \mathbb{Z}$.

We now show how to recover the function of Theorem 14.2 from its values at integer points.

Theorem 14.5. *Suppose* $f : \mathbb{R} \to \mathbb{C}$ *is a continuous function with* $\int_{-\infty}^{\infty} |f(t)|\, dt < \infty$. *If* $\hat{f}(\lambda) = 0$ *for* $|\lambda| \geq \pi$ *then*

$$\sum_{n=-N}^{N} f(n) \operatorname{sinc}(t - n) \to f(t)$$

uniformly as $N \to \infty$.

To prove this we proceed as follows. Set $F = \hat{f}$. By the inversion theorem

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\lambda) \exp(i\lambda t)\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\lambda) \exp(i\lambda t)\, dt$$

and, more particularly,

$$f(n) = \frac{1}{2\pi} \hat{F}(-n).$$

This enables us to use results on $\mathbb{T}$ such as Schwarz's inequality and Parseval's equality as follows.

$$\begin{aligned}
&\left|\sum_{n=-N}^{N} f(n)\operatorname{sinc}(t-n) - f(t)\right| \\
&\qquad = \left|\sum_{n=-N}^{N} \frac{1}{2\pi}\hat{F}(-n)\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(i(t-n)\lambda)\,d\lambda - \frac{1}{2\pi}\int_{-\pi}^{\pi} F(\lambda)\exp(i\lambda t)\,d\lambda\right| \\
&\qquad = \frac{1}{2\pi}\left|\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\sum_{n=-N}^{N}\hat{F}(-n)\exp(i\lambda(t-n)) - F(\lambda)\exp(i\lambda t)\right)\,d\lambda\right| \\
&\qquad \leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{1}{2\pi}\sum_{n=-N}^{N}\hat{F}(-n)\exp(-i\lambda n) - F(\lambda)\right|\,d\lambda \\
&\qquad = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|F(\lambda) - \frac{1}{2\pi}\sum_{n=-N}^{N}\hat{F}(n)\exp(i\lambda n)\right|\,d\lambda \\
&\qquad \leq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|F(\lambda) - \frac{1}{2\pi}\sum_{n=-N}^{N}\hat{F}(n)\exp(i\lambda n)\right|^2\,d\lambda\right)^{1/2} \\
&\qquad = \frac{1}{2\pi}\left(\sum_{|n|\geq N+1}|\hat{F}(n)|^2\right)^{1/2} \to 0
\end{aligned}$$

as $N \to \infty$.

[At the level that this course is given we could have avoided the last two steps by assuming that f and thus F is sufficiently well behaved that $\frac{1}{2\pi}\sum_{n=-N}^{N}\hat{F}(n)\exp(i\lambda n) \to F(\lambda)$ uniformly, but the proof given here applies more generally.] We restate Theorem 14.2 in a very slightly generalised form.

Theorem 14.6. (Shannon's Theorem) *Suppose $f : \mathbb{R} \to \mathbb{C}$ is a continuous function with $\int_{-\infty}^{\infty}|f(t)|\,dt < \infty$ and that $K > 0$. If $\hat{f}(\lambda) = 0$ for $|\lambda| \geq K$ then then f is determined by its values at points of the form $n\pi K^{-1}$ with $n \in \mathbb{Z}$.*

We call πK^{-1} the 'Nyquist rate'. Since electronic equipment can only generate, transmit and receive in a certain band of frequencies and sampling more frequently than the Nyquist rate produces, in principle, no further information it is reasonable to suppose that the rate of transmission of information is is proportional to the Nyquist rate. We thus have

$$\frac{\text{rate of transmission of information}}{\text{band width of signal}} \leq \text{constant}$$

where the constant can be improved a little by elegant engineering but must remain of the same order of magnitude. Fibre optics gives a much broader bandwidth and therefore allows a much faster rate of information transmission than earlier systems.

We saw earlier that radio contact with a submerged submarine requires the use of very low frequencies and this means that the rate of transmission of information is very low indeed (reportedly more that a minute to transmit a couple of letters of Morse code).

The human ear is only sensitive to limited band of frequencies. Thus provided the sampling rate is high enough and the sampling done to sufficient precision sound can be recorded in digital form. This is the principle of the compact disc. (It is a comment on the ingenuity of engineers that the sampling for a CD is done fairly close to the appropriate Nyquist rate.)

15 Distributions on $\mathbb{T}$

For the moment we shall work on the circle $\mathbb{T}$. In Section 11 we introduced the notion of the delta function as follows. Let $h_n : \mathbb{T} \to \mathbb{R}$ be a sequence of continuous functions such that

(i) $h_n(t) \geq 0$ for all $t \in \mathbb{T}$,

(ii) $\displaystyle\int_{\mathbb{T}} h_n(t)\,dt = 1$,

(iii) $h_n(t) \to 0$ uniformly for all $\eta \leq |t| \leq \pi$ whenever $\eta > 0$.

Then we take $\int_{\mathbb{T}} f(t)\delta(t)\,dt$ to be the limit of $\int_{\mathbb{T}} f(t)h_n(t)\,dt$. The strengths of this approach are illustrated in the first part of the next exercise (done as Exercise 19.11) and the weaknesses in the second part.

Exercise 15.1. *(i) If h_n has the properties stated in the previous paragraph and $f : \mathbb{T} \to \mathbb{C}$ is continuous, then*

$$\int_{\mathbb{T}} f(t)h_n(t)\,dt \to f(0)$$

as $n \to \infty$.

(ii) Consider the functions k_n, $l_n : \mathbb{T} \to \mathbb{R}$ given by

$$k_n(t) = \begin{cases} 4(1-nt)/3 & \textit{if } 0 \leq t \leq n^{-1}, \\ 4(1-2nt)/3 & \textit{if } -2n^{-1} \leq t < 0, \\ 0 & \textit{otherwise,} \end{cases}$$

and $l_n(t) = k_n(-t)$. Show that k_n and l_n satisfy the condition placed on h_n in the previous paragraph. Show, however, that if we define $f : \mathbb{T} \to \mathbb{R}$ by

$$f(t) = \begin{cases} \pi - t & \text{if } 0 < t \leq \pi, \\ -\pi - t & \text{if } -\pi < t < 0, \\ 0 & \text{if } t = 0, \end{cases}$$

then

$$\int_{\mathbb{T}} f(t)k_n(t)\,dt \to -\frac{\pi}{3} \text{ and } \int_{\mathbb{T}} f(t)l_n(t)\,dt \to \frac{\pi}{3}$$

as $n \to \infty$.

Our approach and the more general approach via measure theory also fails to assign any meaning to the 'derivative δ' of the delta function' a concept used with considerable success by the physicist Dirac.

Exercise 15.1 seems to show that the delta function 'can be integrated against well behaved functions but not against less well behaved functions'. Laurent Schwarz had the happy idea of only pairing very well behaved objects with objects which (at least from the view point of classical analysis) might be rather badly behaved.

We first need a class of well behaved objects.

Definition 15.2. *We let $\mathcal{D}$ (read 'curly D') be the set of infinitely differentiable functions $f : \mathbb{T} \to \mathbb{C}$.*

In order to do analysis we need a notion of convergence.

Definition 15.3. *If f and f_n lie in $\mathcal{D}$, we say that $f_n \underset{\mathcal{D}}{\to} f$ if, for each fixed $r \geq 0$, we have $f_n^{(r)} \to f^{(r)}$ uniformly on $\mathbb{T}$.*

I suggest that the reader does to following short exercise to check that she understands the quantifiers in the definition just given.

Exercise 15.4. *Let $f_n(x) = 2^{-n} \sin nx$. Show that $f_n \underset{\mathcal{D}}{\to} 0$ but that $f_n^{(r)}(x)$ does not converge uniformly to 0 for $x \in \mathbb{T}$ and $r \geq 0$ as $n \to \infty$.*

We now have our collection of good objects $\mathcal{D}$ together with a notion of convergence and must use them to define our 'less classically good' objects.

Definition 15.5. *We write $\mathcal{D}'$ for the set of linear maps $T : \mathcal{D} \to \mathbb{C}$ which are continuous in the sense that if f and f_n lie in $\mathcal{D}$ and $f_n \underset{\mathcal{D}}{\to} f$, then $Tf_n \to Tf$.*

We call the set $\mathcal{D}'$ the space of distributions and $\mathcal{D}$ the space of test functions. We often write

$$Tf = \langle T, f\rangle.$$

To find out what a distribution T does we take a test function f and look at the value of $Tf = \langle T, f\rangle$.

Exercise 15.6. *Show that, if we set*

$$\langle \delta, f\rangle = f(0)$$

for all $f \in \mathcal{D}$ then $\delta \in \mathcal{D}'$.

The following is a simple but important observation.

Lemma 15.7. *If $\phi : \mathbb{T} \to \mathbb{C}$ and we set*

$$\langle T_\phi, f\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} \phi(t)f(t)\,dt$$

then $T_\phi \in \mathcal{D}'$.

We shall write $\langle \phi, f\rangle = \langle T_\phi, f\rangle$.

Whenever we define a new object T which we hope will be a distribution we must check that:-

(A) $\langle T, \lambda f + \mu g\rangle = \lambda\langle T, f\rangle + \mu\langle T, g\rangle$.

(B) $f_n \underset{\mathcal{D}}{\to} f$ implies $f_n \underset{\mathcal{D}}{\to} f$.

(C) Our definition is consistent when we use ordinary functions ϕ as distributions.

Conditions (A) and (B) are simply the definition. of a distribution. The meaning of condition (C) becomes clearer if we look at the following example.

Lemma 15.8. *Let T and S be distributions, λ, $\mu \in \mathbb{C}$ and let $\tau \in \mathcal{D}$. Then we may define distributions $\lambda T + \mu S$ and FT by*

(i) $\langle \lambda T + \mu S, f\rangle = \lambda\langle T, f\rangle + \mu\langle S, f\rangle$.

(ii) $\langle FT, f\rangle = \langle T, Ff\rangle$.

In order to check condition (C) we must prove the following lemma.

Lemma 15.9. *We use the definitions and notations of Lemmas 15.7 and 15.8.*

(i) If ϕ, $\psi \in \mathcal{D}$ and λ, $\mu \in \mathbb{C}$ then

$$T_{\lambda\phi+\mu\psi} = \lambda T_\phi + \mu T_\psi$$

.

(ii) If F, $\phi \in \mathcal{D}$ then

$$T_{F\phi} = FT_\phi.$$

An even clearer example of the use of condition (C) occurs when we seek to define the derivative of a distribution. Observe that if ϕ, $f \in \mathcal{D}$ then

$$\begin{aligned}\langle \phi', f\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi'(t)f(t)\,dt\\ &= \frac{1}{2\pi}\left[\phi(t)f(t)\right]_{-\pi}^{\pi} - \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(t)f'(t)\,dt\\ &= -\frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(t)f'(t)\,dt\\ &= -\langle \phi, f'\rangle.\end{aligned}$$

This fixes the form of our definition.

Definition 15.10. *If $T \in \mathcal{D}'$ then*

$$\langle T', f\rangle = -\langle T, f'\rangle$$

for all $f \in \mathcal{D}$.

Lemma 15.11. *(i) If $T \in \mathcal{D}'$ then $T' \in \mathcal{D}'$.*
(ii) If T, $S \in \mathcal{D}'$ and λ, $\mu \in \mathbb{C}$ then $(\lambda T + \mu S)' = \lambda T' + \mu S'$.

Exercise 15.12. *If*

$$f(t) = \begin{cases} \pi - t & \text{if } 0 < t \leq \pi,\\ -\pi - t & \text{if } -\pi < t < 0,\\ 0 & \text{if } t = 0,\end{cases}$$

then

$$T_f' = -T_1 + 2\pi\delta$$

or, more concisely

$$f' = -1 + 2\pi\delta.$$

We have defined convergence on $\mathcal{D}$ but not on $\mathcal{D}'$. The next definition remedies this deficiency.

Definition 15.13. *Let T_n, $T \in \mathcal{D}'$ we say that $T_n \underset{\mathcal{D}'}{\to} T$ if*

$$\langle T_n, f\rangle \to \langle T, f\rangle$$

for all $f \in \mathcal{D}$.

Exercise 15.14. *If g, $g_n \in C(\mathbb{T})$ and $g_n \to g$ uniformly, show that $g_n \underset{\mathcal{D}'}{\to} g$.*

Exercise 15.15. *(i) If $T_n, T \in \mathcal{D}'$ and $T_n \underset{\mathcal{D}'}{\to} T$ show that $T'_n \underset{\mathcal{D}'}{\to} T'$.*

(ii) Define $g_n : \mathbb{T} \to \mathbb{R}$ by $g_n(x) = A_n(x^2 - n^{-2})$ for $|x| \leq n^{-1}$ and $g_n(x) = 0$ otherwise where A_n is chosen so that $\int_{\mathbb{T}} g_n(t)\,dt = 1$. Explain why $g_n \underset{\mathcal{D}'}{\to} \delta$.

(iii) By (i) we have $g'_n \underset{\mathcal{D}'}{\to} \delta'$. Sketch the function g'_n for various values of n. Physicists like this way of visualising δ'. However the space $\mathcal{D}'$ contains much odder objects such as the 'distributional derivative' of classically nowhere differentiable functions.

We have seen that we can do many things with distributions. However we can not do everything that we wish. For example we can not (at least in the standard theory of distributions developed here) always multiply distributions. To see why this should be so, let us ask what meaning we should assign to δ^2 the square of the delta function. If h_n is the function discussed in the first paragraph of this section, then $h_n \underset{\mathcal{D}'}{\to} \delta$ so, presumably, we would want $h_n^2 \underset{\mathcal{D}'}{\to} \delta^2$. But, if $f : \mathbb{T} \to \mathbb{R}$ is continuous with $f(0) > 0$, then it is easy to see (Exercise 19.28) that

$$\int_{\mathbb{T}} h_n^2(x) f(x)\,dx \to \infty$$

as $n \to \infty$.

16 Distributions and Fourier series

[This was the last lecture and some proofs were merely sketched. None of the proofs are particularly hard but some require accurate argument.]

We require Theorem 16.2 which in turn requires Theorem 16.1 which we just quote.

Theorem 16.1. *Let $g_j : \mathbb{T} \to \mathbb{C}$ be differentiable with continuous derivative. If $\sum_{j=1}^n g_j(x)$ converges for each x and $\sum_{j=1}^n g'_j$ converges uniformly as $n \to \infty$, then $\sum_{j=1}^\infty g_j$ is differentiable and*

$$\frac{d}{dx}\left(\sum_{j=1}^{\infty} g_j(x)\right) = \sum_{j=1}^{\infty} g'_j(x).$$

Theorem 16.2. *(i) If $f \in \mathcal{D}$ then $n^k \hat{f}(n) \to 0$ as $|n| \to \infty$ for every integer $k \geq 0$.*

(ii) If $a_n \in \mathbb{C}$ and $n^k a_n \to 0$ as $|n| \to \infty$ for every integer $k \geq 0$ then there exists a unique $f \in \mathcal{D}$ such that $\hat{f}(n) = a_n$.

Theorems 7.5, 16.1 and 16.2 give us the following useful result. (As usual we write $S_n(f,t) = \sum_{j=-n} \hat{f}(n) \exp(ijt)$.)

Lemma 16.3. *If $f \in \mathcal{D}$ then*

$$S_n(f) \underset{\mathcal{D}}{\to} f$$

as $n \to \infty$.

(This result appears to put our previous results on convergence in the shade but it must be remembered that the space $\mathcal{D}$ is a very small subset of the classes of function that we considered earlier.)

How should we define the Fourier coefficients of a distribution. We observe that, if $\phi \in C(\mathbb{T})$, then

$$\hat{\phi}(n) = \frac{1}{2\pi} \int_{\mathbb{T}} \phi(t) e_{-n}(t)\, dt = (2\pi)^{-1} \langle \phi, e_{-n} \rangle$$

where $e_m(t) = \exp(imt)$. Our principle (C) thus leads us to the following definition.

Definition 16.4. *If $T \in \mathcal{D}'$ then*

$$\hat{T}(n) = \langle T, e_{-n} \rangle.$$

Lemma 16.3 now tells us that

$$\sum_{-n}^{n} \hat{T}(-j)\hat{f}(j) = \langle T, S_n f \rangle \to \langle T, f \rangle$$

as $n \to \infty$ whenever $T \in \mathcal{D}'$ and $f \in \mathcal{D}$. A little reflection gives the following theorem.

Theorem 16.5. *(i) If $T \in \mathcal{D}$ then there exists an integer $K \geq 0$ such that $n^{-K}\hat{T}(n) \to 0$ as $|n| \to \infty$.*

(ii) If $b_n \in \mathbb{C}$ and there exists an integer $K \geq 0$ such that $n^{-K} b_n \to 0$ as $|n| \to \infty$ for every integer $k \geq 0$ then there exists a unique $T \in \mathcal{D}'$ such that $\hat{T}(n) = b_n$.

(iii) If $T \in \mathcal{D}'$ and $f \in \mathcal{D}$ then

$$\langle T, f \rangle = \sum_{j=-\infty}^{\infty} \hat{T}(-j)\hat{f}(j).$$

The sum on the right is absolutely convergent.

Exercise 16.6. *If $T \in \mathcal{D}'$ show that*

$$\sum_{j=-n}^{n} \hat{T}(n)e_n \underset{\mathcal{D}}{\rightarrow} T.$$

What about convolution? We have not defined convolution on $\mathbb{T}$ even for ordinary functions but it is clear (apart from a choice of constants) what the definition should be in this case.

Definition 16.7. *If ϕ, $\psi \in C(\mathbb{T})$, we set*

$$\phi * \psi(t) = \frac{1}{2\pi}\int_{\mathbb{T}} \phi(t-s)\psi(s)\,ds.$$

Thus, if ϕ, $\psi \in C(\mathbb{T})$ and $f \in \mathcal{D}$, then

$$\begin{aligned}
\langle \phi * \psi, f\rangle &= \frac{1}{2\pi}\int_{\mathbb{T}}\left(\frac{1}{2\pi}\int_{\mathbb{T}} \phi(t-s)\psi(s)\,ds\right) f(t)\,dt \\
&= \frac{1}{(2\pi)^2}\int_{\mathbb{T}}\int_{\mathbb{T}} \phi(t-s)\psi(s)f(t)\,ds\,dt \\
&= \frac{1}{(2\pi)^2}\int_{\mathbb{T}}\int_{\mathbb{T}} \phi(t-s)\psi(s)f(t)\,dt\,ds \\
&= \frac{1}{(2\pi)^2}\int_{\mathbb{T}}\int_{\mathbb{T}} \phi(u)\psi(s)f(u+s)\,du\,ds \\
&= \langle \phi(u), \langle \psi(s), f(u+s)\rangle\rangle.
\end{aligned}$$

Principle (C) thus suggests the following definition.

Definition 16.8. *If T, $S \in \mathcal{D}'$ we define $T * S \in \mathcal{D}'$ by*

$$\langle T * S, f\rangle = \langle T(u), \langle S(s), f(u+s)\rangle\rangle.$$

Here our notation is slightly informal with s and u acting as *dummy variables*. It requires a fair amount of work to show that this definition actually makes sense but the reader may either take this on trust or do Exercise 19.29.

We have the following satisfying result.

Lemma 16.9. *If T, $S \in \mathcal{D}'$ then*

$$\widehat{T * S}(n) = \hat{T}(n)\hat{S}(n).$$

Theorem 16.5 is very satisfactory but raises the possibility that a distribution might be 'merely a formal trigonometric series'. The reason that this is not the case is that, although it makes no sense to talk about the value of a distribution at a point, distributions actually have locality.

Exercise 16.10. *Show that if $f \in \mathcal{D}$ and there exists an $\eta > 0$ such that $f(t) = 0$ for $|t| \leq \eta$ then*

$$\langle \delta', f \rangle = 0$$

Thus it makes sense to think of δ' as 'living at 0'. The next exercise shows the need for caution when using this idea.

Exercise 16.11. *Let $f(x) = \sin x$. Observe that that $f \in \mathcal{D}$ and $f(0) = 0$ but*

$$\langle \delta', f \rangle = -1.$$

Why does this not contradict Exercise 16.10?

We need a little topology (see Exercise 19.32) to exploit the idea of locality to the full but the following lemma and its proof give the main idea.

Lemma 16.12. *Suppose $T \in \mathcal{D}'$ is such that $\langle T, f_j \rangle = 0$ whenever $f_j \in \mathcal{D}$ and $f_j(x) = 0$ for $x \notin [a_j - \eta, b_j + \eta]$ for $j = 1, 2$ and some $\eta > 0$. Then $\langle T, f \rangle = 0$ whenever $f \in \mathcal{D}$ and $f(x) = 0$ for $x \notin [a_1, b_1] \cup [a_2, b_2]$.*

Proof. Choose $E_j \in \mathcal{D}$ such that $1 \geq E_j(x) \geq 0$ for all x, $E_j(x) = 1$ for all $x \in [a_j, b_j]$, $E_j(x) > 0$ for all $x \in (a_j - \eta, b_j + \eta)$ and $E_j(x) = 0$ for all $x \notin [a_j - \eta, b_j + \eta]$ $[j = 1, 2]$. Choose $E_3 \in \mathcal{D}$ such that $1 \geq E_3(x) \geq 0$ for all x, $E_3(x) = 1$ for all $x \notin (a_1 - \eta, b_1 + \eta) \cup (a_2 - \eta, b_2 + \eta)$, $E_j(x) > 0$ for all $x \notin [a_1, b_1] \cup [a_2, b_2]$ and $E_j(x) = 0$ for all $x \in [a_1, b_1] \cup [a_2, b_2]$. (See Exercises 19.30 and 19.31.)

We observe that $E_1(x) + E_2(x) + E_3(x) > 0$ for all x and so we may set

$$G_k(x) = \frac{E_k(x)}{E_1(x) + E_2(x) + E_3(x)}$$

obtaining $G_k \in \mathcal{D}$ for $1 \leq k \leq 3$. Note that

$$G_1 + G_2 + G_3 = 1$$

and so

$$\langle T, f \rangle = \langle T, (G_1 + G_2 + G_3) f \rangle = \sum_{k=1}^{3} \langle T, G_k \rangle.$$

If $f(x) = 0$ for $x \notin [a_1, b_1] \cup [a_2, b_2]$, then $G_3 f = 0$ so $\langle T, G_3 \rangle = 0$. Also $f_j = G_j f$ vanishes outside $x \notin [a_j - \eta, b_j + \eta]$ so $\langle T, G_j f \rangle = \langle T, f_j \rangle = 0$ for $j = 1$ and for $j = 2$. Thus $\langle T, f \rangle = 0$ as stated. □

17 Distributions on $\mathbb{R}$

So far as physicists and engineers are concerned $\mathbb{T}$ is a 'toy space'. What happens if we try to extend the theory of distributions to $\mathbb{R}$? The short answer is that the theory does extend but the fact that $\mathbb{R}$ is *unbounded* (or to speak more correctly, but more technically, *non-compact*) means that matters are less straightforward.

One way forward in inspired by Theorem 16.5.

Definition 17.1. *We let $\mathcal{S}$ (the 'Schwartz space') be the set of infinitely differentiable functions $f : \mathbb{R} \to \mathbb{C}$ such that*

$$(1 + x^2)^m f^r(x) \to 0$$

as $|x| \to \infty$ for all positive integer r and m. (We say that f and all its derivatives 'decrease faster than polynomial'.)

If f and f_n lie in $\mathcal{S}$ we say that $f_n \underset{\mathcal{S}}{\to} f$ if, for each fixed pair of positive integers r and m, we have $(1 + x^2)^m (f_n^{(r)}(x) - f^{(r)}(x)) \to 0$ uniformly on $\mathbb{R}$.

It turns out that the Schwartz space is beautifully adapted to the Fourier transform.

Theorem 17.2. *If $f \in \mathcal{S}$, let us write*

$$\mathcal{F}f(\lambda) = \hat{f}(\lambda) = \int_{-\infty}^{\infty} f(t) e^{-i\lambda t}\, dt.$$

Then $\mathcal{F}f$ is a well defined element of $\mathcal{S}$.

The map $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is linear and

$$\mathcal{F}^2 = 2\pi J$$

where $Jf(x) = f(-x)$. Thus $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is a bijection. Further $\mathcal{F}$ is continuous in the sense that $f_n \underset{\mathcal{S}}{\to} f$ implies $\mathcal{F}f_n \underset{\mathcal{S}}{\to} \mathcal{F}f$.

In is easy to define the appropriate space $\mathcal{S}'$ of distributions (called the Schwartz space of *tempered distributions*).

Definition 17.3. *We write $\mathcal{S}'$ for the set of linear maps $T : \mathcal{D} \to \mathbb{C}$ which are continuous in the sense that if f and f_n lie in $\mathcal{S}$ and $f_n \underset{\mathcal{S}}{\to} f$, then $Tf_n \to Tf$. We write $Tf = \langle T, f\rangle$.*

If T_n, $T \in \mathcal{S}'$ we say that $T_n \underset{\mathcal{S}'}{\to} T$ if

$$\langle T_n, f\rangle \to \langle T, f\rangle$$

for all $f \in \mathcal{S}$.

Much of of what we did for $\mathcal{D}$ such as the definition of the derivative of a distribution transfers directly. What about Fourier transforms? Writing $e_\lambda(t) = \exp i\lambda t$ we observe that $e_\lambda \notin \mathcal{S}$ so the expression $\langle T, e_\lambda \rangle$ makes no sense in this context. However, we recall from Lemma 10.4 that, if f, g : $\mathbb{R} \to \mathbb{C}$ are well behaved, then

$$\int_{-\infty}^{\infty} \hat{g}(x) f(x)\, dx = \int_{-\infty}^{\infty} g(\lambda) \hat{f}(\lambda)\, d\lambda,$$

and this hint gives us a way to define Fourier transforms of tempered distributions.

Definition 17.4. *If $T \in \mathcal{S}'$ then we define $\hat{T} \in \mathcal{S}'$ by the formula*

$$\langle \hat{T}, f \rangle = \langle T, \hat{f} \rangle.$$

Exercise 17.5. *Check that the definition works correctly.*

Theorem 17.2 which tells us that the Fourier transform works well on $\mathcal{S}$ now implies that the Fourier transform works well on $\mathcal{S}'$.

Theorem 17.6. *If $T \in \mathcal{S}'$ let us write $\mathcal{F}T = \hat{T}$. The map $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is linear and*

$$\mathcal{F}^2 = 2\pi J$$

where $\langle JT, f \rangle = \langle T(x), f(-x) \rangle$. Thus $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is a bijection. Further $\mathcal{F}$ is continuous in the sense that $T_n \underset{\mathcal{S}'}{\to} T$ implies $\mathcal{F}T_n \underset{\mathcal{S}'}{\to} \mathcal{F}T$.

The following result is in accordance with a formula often used in formal manipulation of Fourier transforms.

Exercise 17.7. *If we work in $\mathcal{S}'$ then*

$$\hat{\delta} = 1 \text{ and } \hat{1} = 2\pi\delta.$$

A major difference between $\mathbb{T}$ and the unbounded (that is non-compact) space $\mathbb{R}$ is that it is no longer possible to convolve every pair of distributions. Observe that no part of the formula

$$\widehat{1 * 1} \overset{?}{=} \hat{1}\hat{1} \overset{?}{=} (2\pi)^2 \delta^2$$

makes any sense in our theory since $\int_{-\infty}^{\infty} 1\, dt$ diverges. Convolution remains an important operation but we have to impose conditions on the objects convolved to make sure that we can perform it.

From the point of view of applications the space of tempered distributions is too small to deal with the functions of exponential growth which occur in the theory of differential equations.

Exercise 17.8. *If* $\phi(t) = \exp(t)$ *then setting* $f(t) = \exp(-(1+t^2)^{1/2}$ *we have* $f \in \mathcal{F}$ *but* $\int_{-\infty}^{\infty} \phi(t)f(t)\,dt$ *diverges.*

To get round this problem, Schwartz used a smaller space of test functions $\mathcal{D}(\mathbb{R})$, obtaining a larger space of distributions $\mathcal{D}'(\mathbb{R})$. But that is another story recounted, for example, in [2].

18 Further reading

On the whole, interesting books on Fourier analysis are at a higher level than this course. The excellent book of Dym and McKean [1] is, perhaps the most in the spirit of this course and I have also written a book [5] on Fourier analysis. There are two superb introductions to the study of Fourier Analysis for its own sake by Helson [3] and by Katznelson [4].

References

[1] H. Dym and H. P. McKean *Fourier Series and Integrals* Academic Press, 1972.

[2] F. G. Friedlander *Introduction to the Theory of Distributions* CUP, 1982. [There is a second edition also published by CUP in 1998 with an additional chapter by M. Joshi.]

[3] H. Helson *Harmonic Analysis* Adison–Wesley, 1983.

[4] Y. Katznelson *An Introduction to Harmonic Analysis* Wiley, 1963. [There is a Dover reprint, CUP hope to bring out a second edition.]

[5] T. W. Körner *Fourier Analysis* CUP, 1988.

19 Exercises

Here are some exercises. They are at various levels and you are not expected to do all of them. Just do the ones that interest you.

Exercise 19.1. (i) Let $f : \mathbb{R} \to \mathbb{R}$ be $n+1$ times differentiable. Show that there is a unique polynomial (to be exhibited) P of degree n such that $P^{(r)}(0) = f^{(r)}(0)$, for all $0 \leq r \leq n$.

(ii) Let $t > 0$. Set

$$E(t) = f(t) - P(t)$$

(so $E(t)$ is the 'error at point t) and write

$$g(x) = f(x) - P(x) - E(t)\left(\frac{x}{t}\right)^{n+1}.$$

By repeated use of Rolle's theorem show that there exists a $c_r \in (0,t)$ such that

$$g^{(r)}(0) = g^{(r)}(c_r)$$

for $1 \leq r \leq n$. Deduce that there exists a $c \in (0,t)$ such that we have the following 'Taylor theorem'

$$f(t) = P(t) + \frac{f^{(n+1)}(c)t^{n+1}}{(n+1)!}.$$

Exercise 19.2. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is $2n+2$ times differentiable. By considering polynomials of the form $x^k(1-x)^l$, or otherwise, show that there is a unique polynomial P of degree $2n+1$ such that

$$P^{(r)}(0) = f^{(r)}(0) \text{ and } P^{(r)}(1) = f^{(r)}(1) \text{ for all } 0 \leq r \leq n.$$

Show that the error $E(y) = f(y) - P(y)$ at $y \in [0,1]$ is given by

$$E(y) = \frac{f^{(2n+2)}(c)}{(2n+2)!}y^{n+1}(y-1)^{n+1},$$

for some $c \in (0,1)$.

Exercise 19.3. By taking imaginary parts in the de Moivre formula, or otherwise, show that there is a polynomial U_n of degree n such that $U_n(\cos\theta)\sin\theta = \sin(n+1)\theta$.

Show that $T_n'(x) = nU_{n-1}(x)$ for $n \geq 1$.

Exercise 19.4. By looking at the real part of $\sum_{n=0}^{\infty} t^n e^{in\theta}$, or otherwise, show that

$$\frac{1-tx}{1-2tx+t^2} = \sum_{n=0}^{\infty} T_n(x)t^n$$

for all $|t| < 1$ and $|x| \leq 1$.

Exercise 19.5. If $f : \mathbb{T} \to \mathbb{C}$ is continuous, we write

$$P_r(f,\theta) = \sum_{n=-\infty}^{\infty} r^{|n|}\hat{f}(n)\exp in\theta$$

for all $\theta \in \mathbb{T}$ and all real r with $0 < r < 1$. By modifying the proof of the Fejér theorem, show that $P_r(f,\theta) \to f(\theta)$ as $r \to 1$ from below.

Exercise 19.6. The ideas behind the vibrating string may be generalised. For example the equation of a two dimensional vibrating drum is

$$\nabla^2 \phi = K \frac{\partial^2 \phi}{\partial t^2}$$

where, by definition,

$$\nabla^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2}.$$

Suppose we have a circular drum of radius a. It is natural to seek a solution of the form

$$\phi = f(r)(A \cos \omega t + B \sin \omega t)$$

where r is the distance from the centre and $f(a) = 0$.

By using the chain rule, show that

$$\nabla^2 f(r) = f''(r) + \frac{1}{r} f'(r)$$

and deduce that

$$f''(r) + \frac{1}{r} f'(r) + \omega^2 f(r) = 0, \ f(a) = 0.$$

Exercise 19.7. Improve on Lemma 6.4 by showing that

$$\frac{1}{\log N} \left(\frac{1}{2\pi} \int_{\mathbb{T}} |D_N(x)| \, dx \right)$$

tends to a limit and find that limit. (This requires some thought.)

Exercise 19.8. Let $a_1, a_2, \ldots$ be a sequence of complex numbers.

(i) Show that, if $a_n \to a$ then

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \to a$$

as $n \to \infty$.

(ii) By taking an appropriate sequence of 0s and 1s, or otherwise, find a sequence a_n such that a_n does not tend to a limit as $n \to \infty$ but $(a_1 + a_2 + \cdots + a_n)/n$ does.

(iii) By taking an appropriate sequence of 0s and 1s, or otherwise, find a bounded sequence a_n such that $(a_1 + a_2 + \cdots + a_n)/n$ does not tend to a limit as $n \to \infty$.

Exercise 19.9. (i) Show that if $f : \mathbb{T} \to \mathbb{R}$ is continuous and $f(t) = f(-t)$ for all t, then, given any $\epsilon > 0$ we can find a trigonometric polynomial

$$P(t) = \sum_{n=-N}^{N} a_n \cos nt$$

with a_n real such that $\|P - f\|_\infty < \epsilon$.

(ii) By using Tchebychev polynomials (see Lemma 2.8), prove the theorem of Weierstrass which states that if $F : [0,1] \to \mathbb{R}$ is continuous then, given any $\epsilon > 0$ we can find a polynomial

$$Q(t) = \sum_{n=0}^{N} b_n t^n$$

with b_n real such that $\|Q - F\|_\infty < \epsilon$.

(iii) Suppose that $g : [0,1] \to \mathbb{R}$ is continuous and $\int_0^1 g(t)t^n \, dt = 0$ for all $n \geq 0$. Show that $g = 0$.

Exercise 19.10. Consider the heat equation on the circle. In other words consider well behaved functions $\theta : \mathbb{T} \times [0, \infty) \to \mathbb{C}$ satisfying the partial differential equation

$$\frac{\partial \theta}{\partial t} = K \frac{\partial^2 \theta}{\partial x^2}.$$

for all $t > 0$ and all $x \in \mathbb{T}$. Try to find solutions using separation of variables and then use the same kind of arguments as we used for the vibrating string to suggest that the general solution is

$$\theta(x,t) = \sum_{n=-\infty}^{\infty} a_n e^{inx} e^{-Kn^2 t}.$$

What happens as $t \to \infty$?

Exercise 19.11. Suppose $L_n ; \mathbb{T} \to \mathbb{R}$ is continuous and

(A) $\frac{1}{2\pi} \int_{\mathbb{T}} L_n(t) \, dt = 1$,

(B) If $\eta > 0$ then $L_n \to 0$ uniformly for $|t| \geq \eta$ as $n \to \infty$,

(C) $L_n(t) \geq 0$ for all t.

(i) Show that if $f : \mathbb{T} \to \mathbb{T}$ is continuous, then

$$\frac{1}{2\pi} \int_{\mathbb{T}} L_n(t) f(x-t) \, dt \to f(x)$$

uniformly as $n \to \infty$.

(ii) Show that condition (C) can be replaced by

(C') There exists a constant $A > 0$ such that

$$\int_{\mathbb{T}} |L_n(t)|\, dt \leq A.$$

Exercise 19.12. (In this question we write $S_n(f,t) = \sum_{r=-n}^{n} \hat{f}(n) \exp int.$)

(i) Use Theorem 8.2 to show that, if $f : \mathbb{T} \to \mathbb{C}$ is continuous, then $\hat{f}(n) \to 0$ as $|n| \to \infty$ (this is the very simplest version of the 'Riemann-Lebesgue lemma').

(ii) Suppose that f_1, $g_1 : \mathbb{T} \to \mathbb{C}$ are continuous and $f_1(t) = g_1(t) \sin t$ for all t. Show that $\hat{f}_1(j) = (\hat{g}_1(j-1) - \hat{g}_1(j+1)/2$ and deduce that $S_n(f_1, 0) \to 0$ as $n \to \infty$.

(iii) Suppose that $f_2 : \mathbb{T} \to \mathbb{C}$ is continuous, $f_2(n\pi) = 0$ and f_2 is differentiable at 0 and π. Show that there exists a continuous $g_2 : \mathbb{T} \to \mathbb{C}$ such that $f_2(t) = g_2(t) \sin t$ for all t and deduce that $S_n(f_2, 0) \to 0$ as $n \to \infty$.

(iv) Suppose that $f_3 : \mathbb{T} \to \mathbb{C}$ is continuous, $f_3(0) = 0$ and f_3 is differentiable at 0. Write $f_4(t) = f_3(2t)$. Compute $\hat{f}_4(j)$ in terms of the Fourier coefficients of f_3. Show that $S_n(f_4, 0) \to 0$ and deduce that $S_n(f_3, 0) \to 0$ as $n \to \infty$.

(v) Suppose that $f : \mathbb{T} \to \mathbb{C}$ is continuous, and f is differentiable at some point x. Show that $S_n(f, x) \to f(x)$ as $n \to \infty$.

Exercise 19.13. (i) Use Lemma 6.5 to show that given any $\epsilon_1 > 0$ and any $K_1 > 0$ we can find a continuous function $f_1 : \mathbb{T} \to \mathbb{R}$ with $\|f_1\|_\infty \leq \epsilon_1$ and and integer $M_1 > 0$

$$|S_{M_1}(f_1, 0)| \geq K_1.$$

(ii) Use part (i) to show that, given any $\epsilon_2 > 0$ and any $K_2 > 0$, we can find a real trigonometric polynomial $P_2 : \mathbb{T} \to \mathbb{R}$ with $\|P_2\|_\infty \leq \epsilon_2$ and and integer $M_2 > 0$

$$|S_{M_2}(P_2, 0)| \geq K_2.$$

(iii) Use part (ii) to show that, given any $\epsilon_3 > 0$, any $K_3 > 0$ and any integer $m_3 > 0$, we can find a real trigonometric polynomial $P_3 : \mathbb{T} \to \mathbb{R}$ with $\|P_3\|_\infty \leq \epsilon_3$, $\hat{P}_3(r) = 0$ for $|r| \leq m_3$ and and integer $M_3 > 0$

$$|S_{M_3}(P_3, 0)| \geq K_3.$$

(iv) Show that we can find a sequence of real trigonometric polynomials P_n and integers $q(n)$, $M(n)$ and $m(n)$ such that

(a) $\|P_n\|_\infty \leq 2^{-n}$ for all n.
(b) $\hat{P}_n(r) = 0$ if $|r| \leq mn$ or $|r| \geq q(n)$.
(c) $|S_{M(n)}(P_n, 0)| \geq 2^n$.
(d) $q(n-1) \leq m(n) \leq M_n < q(n)$
for all $n \geq 0$. (We take $q(0) = 0$.)

(v) Show carefully that $\sum_{n=1}^\infty P_n$ is uniformly convergent to some continuous function f and that $\hat{f}(r) = \hat{P}_n(r)$ if $m(n) \leq r \leq q(n)$.

(vi) Deduce that $|S_{M(n)}(f, 0) - S_{m(n)}(f, 0)| \geq 2^n$ for all $n \geq 1$ and that $S_N(f, 0)$ can not converge as $N \to \infty$.

Exercise 19.14. (i) Show that, if $f : \mathbb{T} \to \mathbb{C}$, then $\hat{f}(n) \to 0$ as $|n| \to \infty$.

(i) If $\kappa(n) > 0$ and $\kappa(n) \to \infty$ as $n \to \infty$, show that we can find $n(j) \to \infty$ such that $\sum_{j=1}^\infty 2^j/\kappa(n(j))$ converges. Deduce that we can find a continuous function f such that $\limsup_{n\to\infty} \kappa(n)\hat{f}(n) = \infty$.

Exercise 19.15. Show that

$$\frac{\operatorname{card}\{1 \leq n \leq N \mid \langle \log_{10} n \rangle \in [0, 1/2]\}}{N}$$

does not tend to limit as $N \to \infty$. Show however that given any $\epsilon > 0$ and any $x \in [0, 1]$ we can find a positive integer n such that

$$|\langle \log_{10} n \rangle - x| < \epsilon$$

as $N \to \infty$.

Prove the same results with $\log_{10}$ replaced by $\log_e$.

Exercise 19.16. Using the kind of ideas behind the proof of of Weyl's theorem (Theorem 8.4), or otherwise, prove the following results.

(i) If $f : \mathbb{T} \to \mathbb{C}$ is continuous, then $\hat{f}(n) \to 0$ as $|n| \to \infty$. (This is a version of the Lebesgue–Riemann lemma.)

(ii) If $f : \mathbb{T} \to \mathbb{R}$ is continuous, then

$$\int_0^{2\pi} f(t)|\sin nt|\, dt \to \frac{2}{\pi}\int_0^{2\pi} f(t)\, dt$$

as $n \to \infty$.

Exercise 19.17. Let R be a rectangle cut up into smaller rectangles $R(1)$, $R(2)$, ... , $R(k)$ each of which has sides parallel to the sides of R. Then, if each $R(j)$ has at least one pair of sides of integer length, it follows that R has at least one pair of sides of integer length.

First try and prove this without using Fourier analysis.

Then try and prove the result using what is, in effect, a Fourier transform

$$\iint_{R(j)} \exp\big(2\pi i(x+y)\big)\,dx\,dy.$$

Exercise 19.18. If $f,\ g : \mathbb{T} \to \mathbb{C}$ are well behaved, let us define their *convolution* $f * g : \mathbb{T} \to \mathbb{C}$ by

$$f * g(t) = \frac{1}{2\pi}\int_{\mathbb{T}} f(t-s)g(s)\,ds.$$

Show that $\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n)$.

Show that if P is a trigonometric polynomial $P * f$ is a trigonometric polynomial. Identify $D_N * f$ and $K_N * f$ where D_N and K_N are the Dirichlet and Fejér kernels.

Suppose that $L_N(t) = A_N K_N^2(t)$ with A_N chosen so that $\frac{1}{2\pi}\int_{\mathbb{T}} L_N(t)\,dt = 1$. Show that, if f is continuous, $L_N * f(t) \to f(t)$.

Exercise 19.19. . Let $E(x) = (2\pi)^{-1/2}\exp(-x^2/2)$. Show, by changing to polar coordinates, that

$$\begin{aligned}\left(\int_{-\infty}^{\infty} E(x)\,dx\right)^2 &= \int_{-\infty}^{\infty} E(x)\,dx \int_{-\infty}^{\infty} E(y)\,dy\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(-(x^2+y^2))\,dx\,dy\\ &= \int_0^{\infty} r\exp(-r^2/2)\,dr = 1.\end{aligned}$$

Kelvin once asked his class if they knew what a mathematician was. He wrote the formula

$$\int_{\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{\pi}$$

and the board and said. 'A mathematician is one to whom that is as obvious as that twice two makes four is to you. Liouville was a mathematician.'

Exercise 19.20. (i) Explain why

$$\sum_{j=-N}^{N} |a_j b_j| \leq \left(\sum_{j=-N}^{N} |a_j|^2\right)^{1/2} \left(\sum_{j=-N}^{N} |b_j|^2\right)^{1/2}$$

for all $a_j,\ b_j \in \mathbb{C}$.

(ii) Use (i) to show that, if $\sum_{j=-\infty}^{\infty} |a_j|^2$ and $\sum_{j=-\infty}^{\infty} |b_j|^2$ converges, then $\sum_{j=-\infty}^{\infty} |a_j b_j|$ converges and

$$\sum_{j=-\infty}^{\infty} |a_j b_j| \leq \left(\sum_{j=-\infty}^{\infty} |a_j|^2 \right)^{1/2} \left(\sum_{j=-\infty}^{\infty} |b_j|^2 \right)^{1/2}.$$

(iii) If $f : \mathbb{T} \to \mathbb{C}$ is continuously differentiable, explain why

$$\sum_{j=-\infty}^{\infty} j^2 |\hat{f}(j)|^2 \leq \frac{1}{2\pi} \int_{\mathbb{T}} |f'(t)|^2 \, dt.$$

(iv) Use (ii) to show that $\sum_{j=-\infty}^{\infty} |\hat{f}(j)|$ converges. Deduce that $\sum_{j=-\infty}^{\infty} \hat{f}(j) \exp ijt$ converges uniformly to $f(t)$.

Exercise 19.21. (i) If $u : \mathbb{T} \to \mathbb{R}$ is once continuously differentiable and $\frac{1}{2\pi} \int_{\mathbb{T}} u(t) \, dt = 0$, show that

$$\frac{1}{2\pi} \int_{\mathbb{T}} (u(t))^2 \, dt \leq \frac{1}{2\pi} \int_{\mathbb{T}} (u'(t))^2 \, dt$$

with equality if and only if $u(t) = C \cos(t + \phi)$ for some constants C and ϕ.

(ii) Use (i) to show that, if $v : [0, \pi/2] \to \mathbb{R}$ is once continuously differentiable with $v(0) = 0$ and $v'(\pi/2) = 0$ then

$$\int_0^{\pi/2} (v(t))^2 \, dt \leq \int_0^{\pi/2} (v'(t))^2 \, dt$$

with equality if and only if $u(t) = C \sin t$ for some constant C.

(iii) By approximating w by functions of the type considered in (iii) show that, if $w : [0, \pi/2] \to \mathbb{R}$ is once continuously differentiable with $w(0) = 0$, then

$$\int_0^{\pi/2} (w(t))^2 \, dt \leq \int_0^{\pi/2} (w'(t))^2 \, dt.$$

(This is Wirtinger's inequality.)

Exercise 19.22. (i) By applying Poisson's formula to the function f defined by $f(x) = \exp(-t|x|/2\pi)$ show that

$$2(1 - e^{-t})^{-1} = \sum_{n=-\infty}^{\infty} 2t(t^2 + 4\pi^2 n^2)^{-1}.$$

(ii) By expanding $(t^2 + 4\pi n^2)^{-1}$ and interchanging sums (justifying this, if you can, just interchanging, if not) deduce that

$$2(1 - e^{-t})^{-1} = 1 + 2t^{-1} + \sum_{m=0}^{\infty} c_m t^m$$

where $c_{2m} = 0$ and

$$c_{2m+1} = a_{2m+1} \sum_{n=1}^{\infty} n^{-2m}$$

for some value of a_{2m+1} to be given explicitly.

(iii) Hence obtain Euler's formula

$$\sum_{n=1}^{\infty} n^{-2m} = (-1)^{m-1} 2^{2m-1} b_{2m-1} \pi^{2m}/(2m-1)!$$

for $m \geq 1$, where the b_m are defined by the formula

$$(e^y - 1)^{-1} = y^{-1} - 2^{-1} + \sum_{n=1}^{\infty} b_n y^n/n!$$

(The b_n are called Bernoulli numbers.)

Exercise 19.23. (The Gibbs Phenomenon.) Ideally you should first look at what happens when we try to reconstruct a reasonable discontinuous function from its Fourier sums and then use this question to explain what you see. There are a number of questions linked to this one but you need not do them to understand what is going on.

We have only discussed Fourier series for continuous functions in this course. It is possible to use what we already know to discuss 'well behaved' discontinuous functions. Let $F : \mathbb{T} \to \mathbb{R}$ be defined by

$$F(t) = \begin{cases} \pi - t & \text{for } 0 \leq t \leq \pi \\ 0 & \text{for } t = 0 \\ pi - t & \text{for } -\pi < t \leq 0 \end{cases}$$

(i) Suppose that $f : \mathbb{T} \to \mathbb{R}$ is continuously differentiable on $\mathbb{T} \setminus \{0\}$ and that both the function f and its derivative have left and right limits at 0. Show that we can find a λ and and a continuous function $g : \mathbb{T} \to \mathbb{R}$ such that g' exists and is continuous except possibly at 0 and that g has left and right derivatives at 0 which are the left and right limits of g' at that point and

$$f = g + \lambda F.$$

Since the Fourier sums of g behave extremely well (see Exercise 19.24 or take my word for it) it follows that any bad behaviour will be due to F and study of the Fourier sums of F will tell us all we need to know about well behaved functions with a single well behaved discontinuity at 0. Why will this also tell us all we need to know about well behaved functions with a single well behaved discontinuity? Can the same idea be made to work for well behaved functions with a finite number of well behaved discontinuities?

(ii) Show that the nth Fourier sum of F

$$S_n(F,t) = \sum_{r=-n}^{n} \hat{F}(r) \exp(irt) = 2\sum_{r=1}^{n} \frac{1}{r} \sin rt.$$

(iii) Explain why

$$S_n(F,\tau/n) = 2\frac{\tau}{n}\sum_{r=1}^{n} \frac{1}{\frac{r\tau}{n}} \sin \tfrac{r\tau}{n} \rightarrow \int_0^\tau \frac{\sin x}{x}\,dx$$

as $n \rightarrow \infty$.

(iv) Sketch the behaviour of the function

$$G(\tau) = \int_0^\tau \frac{\sin x}{x}\,dx.$$

(The information in questions 19.25 and 19.26 including the fact that $\int_0^\infty \frac{\sin x}{x}\,dx = \pi/2$ is useful but not essential.)

(v) Sketch the behaviour of $S_n(F,\tau/n)$ for small τ and large n.
[General theorems show that $S_n(F,t) \rightarrow F(t)$ when t is fixed but if the the reader is unwilling to take my word for this they can do Question 19.27.]

Exercise 19.24. (This is just an extension of Question 19.20) (i) Suppose that $f : [-\pi,\pi] \rightarrow \mathbb{C}$ is continuous. We define

$$\hat{f}(r) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \exp(irt)\,dt.$$

Show that

$$\sum_{n=-N}^{N} |\hat{f}(n)|^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt.$$

(ii) Suppose that $f : [-\pi, \pi] \to \mathbb{C}$ has continuous first derivative (we use left and right derivatives at end points) and that $f(\pi) = f(-\pi)$. Show that

$$\widehat{f'}(r) = ir\hat{f}(r).$$

Show that

$$\sum_{r=-n}^{n} |\hat{f}(r)| \leq |\hat{f}(0)| + 2\sum_{r=1}^{n} r^{-2}\frac{1}{2\pi}\int_{-\pi}^{\pi} |f'(t)|^2\,dt.$$

Conclude that $\sum_{r=-\infty}^{\infty} |\hat{f}(r)|$ converges.

(iii) Suppose that $g : \mathbb{T} \to \mathbb{C}$ is continuous, that g' exists and is continuous except possibly at one point α and that g has left and right derivatives at α which are the left and right limits of g' at that point. Show that $\sum_{r=-\infty}^{\infty} |\hat{g}(r)|$ converges and deduce, using Theorem 7.5, that $\|S_N(g) - g\|_\infty \to 0$ as $N \to \infty$..

Exercise 19.25. (i) Explain why the function $f : [0, \infty) \to \mathbb{R}$ defined by $f(x) = (\sin x)/x$ for $x \neq 0$, $f(0) = 1$ is continuous at 0. It is traditional to write $f(x) = (\sin x)/x$ and ignore the fact that, strictly speaking, $(\sin 0)/0$ is meaningless. Sketch f.

(ii) If $I_n = \displaystyle\int_0^{n\pi} \frac{\sin x}{x}\,dx$, show, by using the alternating series test, that I_n tends to a strictly positive limit L, say. Deduce carefully that $\displaystyle\int_0^{\infty} \frac{\sin x}{x}\,dx$ exists with value L.

(iii) Let $I(t) = \displaystyle\int_0^{\infty} \frac{\sin tx}{x}\,dx$ for all $t \in \mathbb{R}$. Show using (i), or otherwise, that $I(t) = L$ for all $t > 0$, $I(0) = 0$, $I(t) = -L$ for $t < 0$.

(iv) Find a continuous function $g : [0, \pi] \to \mathbb{R}$ such that $g(t) \geq 0$ for all $t \in [0, \pi]$, $g(\pi/2) > 0$ and

$$\left|\frac{\sin x}{x}\right| \geq \frac{g(x - n\pi)}{n}$$

for all $n\pi \leq x \leq (n + 1)\pi$ and all integer $n \geq 1$. Hence, or otherwise, show that $\int_0^\infty |(\sin x)/x|\,dx$ fails to converge.

Exercise 19.26. Although the existence of the infinite integral $\int_0^\infty \frac{\sin x}{x}\,dx$ is very important, its actual value is less important. It is, however, reasonably easy to find using our knowledge of the Dirichlet kernel, in particular the fact that

$$2\pi = \int_{-\pi}^{\pi} \frac{\sin\left((n + \frac{1}{2})x\right)}{\sin\frac{x}{2}}\,dx$$

(see Lemma 6.1 (iii)).

(i) If $\epsilon > 0$, show that

$$\int_{-\epsilon}^{\epsilon} \frac{\sin \lambda x}{x}\, dx \to \int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx,$$

as $\lambda \to \infty$.

(ii) If $\pi \geq \epsilon > 0$, show, by using the estimates from the alternating series test, or otherwise, that

$$\int_{-\epsilon}^{\epsilon} \frac{\sin\big((n+\frac{1}{2})x\big)}{\sin\frac{x}{2}}\, dx \to \int_{-\pi}^{\pi} \frac{\sin\big((n+\frac{1}{2})x\big)}{\sin\frac{x}{2}}\, dx = 2\pi$$

as $n \to \infty$.

(iii) Show that

$$\left|\frac{2}{x} - \frac{1}{\sin\frac{1}{2}x}\right| \to 0$$

as $x \to 0$. and deduce that

$$\int_{0}^{\infty} \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

Exercise 19.27. This question refers back to Question 19.23. There we discussed the behaviour of $S_n(F,t)$ when t is small but did not show that $S_n(F,t)$ behaves well when t is far from 0. This follows from general theorems but we shall prove it directly. This brings us into direct contact with Fourier since he used F as a test case for his statement that any function[1] had a Fourier expansion.

(i) Show that

$$S_n(F,t) = \int_0^t \sum_{r=1}^{n} \cos rx\, dx$$

and

$$\sum_{r=1}^{n} \cos rx = \frac{\sin(n+\frac{1}{2})x}{2\sin\frac{x}{2}}.$$

(ii) Deduce that

$$S_n(F,t) = 2\int_0^t \frac{\sin(n+\frac{1}{2})x}{x}\, dx - t + \int_0^t g(x)\sin(n+\tfrac{1}{2})x\, dx$$

[1]We would now say any 'reasonable function' but Fourier and his contemporaries had a narrower view of what constituted a function.

where $g(0) = 0$ and

$$g(x) = \frac{x - 2\sin\frac{x}{2}}{x\sin\frac{x}{2}}.$$

for $0 < |x| < \pi$.

(iii) Show that g is continuous at 0. Show that g is differentiable at 0 and find its derivative there. Show that g is continuously differentiable on $(-T, T)$ for all $0 < T < \pi$. Note that, in particular, g and g' are bounded on any interval $[-|t|, |t|]$ with $|t| < \pi$. Use integration by parts to show that

$$\begin{aligned}\int_0^t & g(x)\sin(n+\tfrac{1}{2})x\,dx \\ &= -\frac{1}{n+\frac{1}{2}}g(t)\cos(n+\tfrac{1}{2})x + \frac{1}{n+\frac{1}{2}}\int_0^t g'(x)\cos(n+\tfrac{1}{2})x\,dx \\ &\to 0\end{aligned}$$

as $n \to \infty$ for all $0 < |t| < \pi$. (This really just another instance of the Riemann–Lebesgue lemma.

(iv) Show, using Question 19.26, that

$$2\int_0^t \frac{\sin(n+\frac{1}{2})x}{x}\,dx = 2\int_0^{(n+\frac{1}{2})t} \frac{\sin x}{x}\,dx \to \pi$$

and deduce that

$$S_n(F, t) \to F(t)$$

as $n \to \infty$ whenever $0 < |t| < \pi$. Show directly that the result is true when $t = 0$ and $t = \pi$ and so holds for all t.

(v) What does the result of (iv) tell us about the behaviour of the Fourier sums of the function f described in part (i) of Question 19.23?

Exercise 19.28. Suppose that $h_n : \mathbb{T} \to \mathbb{R}$ be a sequence of continuous functions such that

(i) $h_n(t) \geq 0$ for all $t \in \mathbb{T}$,

(ii) $\displaystyle\int_{\mathbb{T}} h_n(t)\,dt = 1$,

(iii) $h_n(t) \to 0$ uniformly for all $\eta \leq |t| \leq \pi$ whenever $\eta > 0$.

If $K > 0$, let us write

$$E_n = \{x \in \mathbb{T} : h_n(x) \geq K\}.$$

Show that

$$\int_{E_n} h_n(t)\,dt \to 1$$

as $n \to \infty$. Deduce that there exists an $N(K)$ such that

$$\int_{E_n} h_n(t)^2\,dt \geq \frac{K}{2}$$

for all $n \geq N(K)$.

If $f : \mathbb{T} \to \mathbb{R}$ is continuous with $f(0) > 0$, deduce that

$$\int_{\mathbb{T}} h_n(x)^2 f(x)\,dx \to \infty$$

as $n \to \infty$.

Exercise 19.29. (i) By using the mean value theorem or some other appropriate version of Taylors theorem,show that, if $f \in \mathcal{D}$,

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

uniformly in x as $h \to 0$.

(ii) If $h_n \neq 0$, $h_n \to 0$ and $f \in \mathcal{D}$, show that

$$\frac{f(x+h_n) - f(x)}{h_n} \underset{\mathcal{D}}{\to} f'(x).$$

Deduce that, if $S \in \mathcal{D}'$ and we write

$$g(x) = \langle S(s), f(x+s)\rangle$$

then

$$\frac{g(x+h_n) - g(x}{h_n} \to \langle S(s), f'(s+x)\rangle$$

as $n \to \infty$.

(iii) If $f \in \mathcal{D}$ and $S \in \mathcal{D}'$, show that

$$\frac{\langle S(s), f(x+h+s)\rangle - \langle S(s), f(x+s)\rangle}{h} \to \langle S(s), f'(s+x)\rangle$$

as $h \to 0$.

(iv) If $f \in \mathcal{D}$ and $S \in \mathcal{D}'$, show that, if $g(x) = \langle S(s), f(x+s)\rangle$ then $g \in \mathcal{D}$. Deduce that, if $T \in \mathcal{D}'$ $\langle T(u), \langle S(s), f(u+s)\rangle$ is a well defined object.

(v) If T, $S \in \mathcal{D}'$ we set

$$\langle T * S, f\rangle = \langle T(u), \langle S(s), f(u+s)\rangle.$$

for all $f \in \mathcal{D}$. Show that $T * S \in \mathcal{D}'$.

Exercise 19.30. Consider the function $E : \mathbb{R} \to \mathbb{R}$ defined by

$$\begin{aligned} E(0) &= 0 \\ E(x) &= \exp(-1/x^2) \qquad \text{otherwise.} \end{aligned}$$

(i) Prove by induction, using the standard rules of differentiation, that E is infinitely differentiable at all points $x \neq 0$ and that, at these points,

$$E^{(n)}(x) = P_n(1/x)\exp(-1/x^2)$$

where P_n is a polynomial which need not be found explicitly.

(ii) Explain why $x^{-1}P_n(1/x)\exp(-1/x^2) \to 0$ as $x \to 0$.

(iii) Show by induction, using the definition of differentiation, that E is infinitely differentiable at 0 with $E^{(n)}(0) = 0$ for all n. [Be careful to get this part of the argument right.]

(iv) Show that

$$E(x) = \sum_{j=0}^{\infty} \frac{E^{(j)}(0)}{j!} x^j$$

if and only if $x = 0$. (The reader may prefer to say that 'The Taylor expansion of E is only valid at 0'.)

(v) If you know some version of Taylor's theorem examine why it does not apply to E.

Exercise 19.31. The hard work for this question was done in Exercise 19.30.

(i) Let $F : \mathbb{R} \to \mathbb{R}$ be defined by $F(x) = 0$ for $x < 0$, $F(x) = E(x)$ for $x \geq 0$ where E is the function defined in Exercise 19.30. Show that F is infinitely differentiable.

(ii) Sketch the functions $f_1,\ f_2 : \mathbb{R} \to \mathbb{R}$ given by $f_1(x) = F(1-x)F(x)$ and $f_2(x) = \int_0^x f_1(t)\,dt$.

(iii) Show that given $a < \alpha < \beta < b$ we can find an infinitely differentiable function $f : \mathbb{R} \to \mathbb{R}$ with $1 \geq f(x) \geq 0$ for all x, $f(x) = 1$ for all $x \in [\alpha, \beta]$, $f(x) > 0$ for $x \in (a, b)$ and $f(x) = 0$ for all $x \notin [a, b]$.

Exercise 19.32. (This requires elementary topology, in particular knowledge of compactness and/or the Heine–Borel theorem.) (i) Let $T \in \mathcal{D}'$. We say that an open interval $(a, b) \in A$ if we can find an $\eta > 0$ such that, if $f \in \mathcal{D}$ and $f(x) = 0$ whenever $x \notin (a - \eta, b + \eta)$ then $\langle T, f\rangle = 0$.

Let $U = \bigcup_{(a,b)\in A}(a, b)$ and $\operatorname{supp} T = \mathbb{T} \setminus U$. Explain why $\operatorname{supp} T$ is closed. Show, by using compactness and an argument along the lines of our proof of Lemma 16.12, that if K is closed set with $K \cap \operatorname{supp} T = \emptyset$, $f \in \mathcal{D}$ and $f(x) = 0$ for all $x \notin K$, then $\langle T, f\rangle = 0$.

(ii) We continue with the notation of (i). Suppose L is a closed set with the property that, if K is closed set with $K \cap L = \emptyset$, $f \in \mathcal{D}$ and $f(x) = 0$ for all $x \notin K$, then $\langle T, f\rangle = 0$. Show that $L \supseteq \operatorname{supp} T$.

(iii) If S, $T \in \mathcal{D}'$ show that

$$\operatorname{supp}(T + S) \subseteq \operatorname{supp} T \cup \operatorname{supp} S.$$

(iv) If $T \in \mathcal{D}'$ show that

$$\operatorname{supp} T' \subseteq \operatorname{supp} T.$$

(v) If $f \in C(\mathbb{T})$ show that $\operatorname{supp} T_f$ (or, more briefly, $\operatorname{supp} f$ is the closure of $\{x : f(x) \neq 0\}$.

If $f \in \mathcal{D}$ and $T \in \mathcal{D}'$ show that

$$\operatorname{supp} fT \subseteq \operatorname{supp} T \cap \operatorname{supp} f.$$

Exercise 19.33. (Only if you know about metric spaces.)

(i) Show that, if we set

$$d(f, g) = \sum_{r=0}^{\infty} \frac{2^{-r}\|f^{(r)} - g^{(r)}\|_\infty}{1 + \|f^{(r)} - g^{(r)}\|_\infty},$$

then $(\mathcal{D}, d)$ is a metric space.

(ii) Show that $f_n \underset{\mathcal{D}}{\rightarrow} f$ if and only if $d(f_n, f) \to 0$.

(iii) (Only if you know what this means.) Show that $(\mathcal{D}, d)$ is complete.

(iv) Find a metric ρ on $\mathcal{S}$ such that $f_n \underset{\mathcal{S}}{\rightarrow} f$ if and only if $\rho(f_n, f) \to 0$

Exercise 19.34. Show that the following equality holds in the space of tempered distributions

$$2\pi \sum_{n=-\infty}^{\infty} \delta_{2\pi n} = \sum_{m=-\infty}^{\infty} e_m$$

where $\delta_{2\pi n}$ is the delta function at $2\pi n$ and e_n is the exponential function given by $e_n(t) = \exp(int)$. What formula results if we take the Fourier transform of both sides?