CHAPTER 4

Differential and Integral operators

The last part of the course includes some applications of Hilbert space and the spectral theorem – the completeness of the Fourier basis, some 'Sturm-Liouville' theory, which is to say the spectral theory for second-order differential operators on an interval or the circle (this case is traditionally called Hill's equation) and enough of a treatment of the eigenfunctions for the harmonic oscillator to show that the Fourier transform is an isomorphism on $L^2(\mathbb{R})$. Once one has all this, one can do a lot more, but there is no time left. Such is life.

1. Fourier series

Let us now try applying our knowledge of Hilbert space to a concrete Hilbert space such as $L^2(a,b)$ for a finite interval $(a,b) \subset \mathbb{R}$. Any such interval with $b > a$ can be mapped by a linear transformation onto $(0, 2\pi)$ and so we work with this special interval. You showed that $L^2(a,b)$ is indeed a Hilbert space. One of the reasons for developing Hilbert space techniques originally was precisely the following result.

Theorem 4.1. *If $u \in L^2(0,2\pi)$ then the Fourier series of u,*

$$\frac{1}{2\pi}\sum_{k\in\mathbb{Z}} c_k e^{ikx},\ c_k = \int_{(0,2\pi)} u(x) e^{-ikx} dx \tag{4.1}$$

converges in $L^2(0,2\pi)$ to u.

Notice that this does not say the series converges pointwise, or pointwise almost everywhere. In fact it is true that the Fourier series of a function in $L^2(0,2\pi)$ converges almost everywhere to u, but it is hard to prove! In fact it is an important result of L. Carleson. Here we are just claiming that

$$\lim_{n\to\infty} \int |u(x) - \frac{1}{2\pi}\sum_{|k|\le n} c_k e^{ikx}|^2 = 0 \tag{4.2}$$

for any $u \in L^2(0,2\pi)$.

Our abstract Hilbert space theory has put us quite close to proving this. First observe that if $e'_k(x) = \exp(ikx)$ then these elements of $L^2(0,2\pi)$ satisfy

$$\int e'_k \overline{e'_j} = \int_0^{2\pi} \exp(i(k-j)x) = \begin{cases} 0 & \text{if } k \ne j \\ 2\pi & \text{if } k = j. \end{cases} \tag{4.3}$$

Thus the functions

$$e_k = \frac{e'_k}{\|e'_k\|} = \frac{1}{\sqrt{2\pi}} e^{ikx} \tag{4.4}$$

form an orthonormal set in $L^2(0,2\pi)$. It follows that (4.1) is just the Fourier-Bessel series for u with respect to this orthonormal set:-

$$c_k = \sqrt{2\pi}(u, e_k) \Longrightarrow \frac{1}{2\pi} c_k e^{ikx} = (u, e_k)e_k. \tag{4.5}$$

So, we already know that this series converges in $L^2(0,2\pi)$ thanks to Bessel's inequality. So 'all' we need to show is

PROPOSITION 4.1. *The e_k, $k \in \mathbb{Z}$, form an orthonormal basis of $L^2(0,2\pi)$, i.e. are complete:*

$$\int u e^{ikx} = 0 \;\forall\; k \Longrightarrow u = 0 \text{ in } L^2(0,2\pi). \tag{4.6}$$

This however, is not so trivial to prove. An equivalent statement is that the finite linear span of the e_k is *dense* in $L^2(0,2\pi)$. I will prove this using Fejér's method. In this approach, we check that any continuous function on $[0,2\pi]$ satisfying the additional condition that $u(0) = u(2\pi)$ is the uniform limit on $[0,2\pi]$ of a sequence in the finite span of the e_k. Since uniform convergence of continuous functions certainly implies convergence in $L^2(0,2\pi)$ and we already know that the continuous functions which vanish near 0 and 2π are dense in $L^2(0,2\pi)$ this is enough to prove Proposition 4.1. However the proof is a serious piece of analysis, at least it seems so to me! There are other approaches, for instance we could use the Stone-Weierstrass Theorem; rather than do this we will deduce the Stone-Weierstrass Theorem from Proposition 4.1. Another good reason to proceed directly is that Fejér's approach is clever and generalizes in various ways as we will see.

So, the problem is to find the sequence in the span of the e_k which converges to a given continuous function and the trick is to use the Fourier expansion that we want to check! The idea of Cesàro is close to one we have seen before, namely to make this Fourier expansion 'converge faster', or maybe better. For the moment we can work with a general function $u \in L^2(0,2\pi)$ – or think of it as continuous if you prefer. The truncated Fourier series of u is a finite linear combination of the e_k :

$$U_n(x) = \frac{1}{2\pi} \sum_{|k|\le n} \Big(\int_{(0,2\pi)} u(t)e^{-ikt}dt\Big)e^{ikx} \tag{4.7}$$

where I have just inserted the definition of the c_k's into the sum. Since this is a finite sum we can treat x as a parameter and use the linearity of the integral to write it as

$$U_n(x) = \int_{(0,2\pi)} D_n(x-t)u(t),\; D_n(s) = \frac{1}{2\pi}\sum_{|k|\le n} e^{iks}. \tag{4.8}$$

Now this sum can be written as an explicit quotient, since, by telescoping,

$$2\pi D_n(s)(e^{is/2} - e^{-is/2}) = e^{i(n+\frac12)s} - e^{-i(n+\frac12)s}. \tag{4.9}$$

So in fact, at least where $s \neq 0$,

$$D_n(s) = \frac{e^{i(n+\frac12)s} - e^{-i(n+\frac12)s}}{2\pi(e^{is/2} - e^{-is/2})} \tag{4.10}$$

and the limit as $s \to 0$ exists just fine.

As I said, Cesàro's idea is to speed up the convergence by replacing U_n by its average

$$V_n(x) = \frac{1}{n+1}\sum_{l=0}^{n} U_l. \tag{4.11}$$

Again plugging in the definitions of the U_l's and using the linearity of the integral we see that

$$V_n(x) = \int_{(0,2\pi)} S_n(x-t)u(t),\ S_n(s) = \frac{1}{n+1}\sum_{l=0}^{n} D_l(s). \tag{4.12}$$

So again we want to compute a more useful form for $S_n(s)$ – which is the Fejér kernel. Since the denominators in (4.10) are all the same,

$$2\pi(n+1)(e^{is/2} - e^{-is/2})S_n(s) = \sum_{l=0}^{n} e^{i(l+\frac{1}{2})s} - \sum_{l=0}^{n} e^{-i(l+\frac{1}{2})s}. \tag{4.13}$$

Using the same trick again,

$$(e^{is/2} - e^{-is/2})\sum_{l=0}^{n} e^{i(l+\frac{1}{2})s} = e^{i(n+1)s} - 1 \tag{4.14}$$

so

$$\begin{gathered} 2\pi(n+1)(e^{is/2} - e^{-is/2})^2 S_n(s) = e^{i(n+1)s} + e^{-i(n+1)s} - 2 \\ \Longrightarrow S_n(s) = \frac{1}{n+1}\frac{\sin^2(\frac{(n+1)}{2}s)}{2\pi\sin^2(\frac{s}{2})}. \end{gathered} \tag{4.15}$$

Now, what can we say about this function? One thing we know immediately is that if we plug $u = 1$ into the discussion above, we get $U_n = 1$ for $n \ge 0$ and hence $V_n = 1$ as well. Thus in fact

$$\int_{(0,2\pi)} S_n(x - \cdot) = 1,\ \forall\ x \in (0, 2\pi). \tag{4.16}$$

Looking directly at (4.15) the first thing to notice is that $S_n(s) \ge 0$. Also, we can see that the denominator only vanishes when $s = 0$ or $s = 2\pi$ in $[0, 2\pi]$. Thus if we stay away from there, say $s \in (\delta, 2\pi - \delta)$ for some $\delta > 0$ then, $\sin(t)$ being a bounded function,

$$|S_n(s)| \le (n+1)^{-1}C_\delta \text{ on } (\delta, 2\pi - \delta). \tag{4.17}$$

We are interested in how close $V_n(x)$ is to the given $u(x)$ in supremum norm, where now we will take u to be continuous. Because of (4.16) we can write

$$u(x) = \int_{(0,2\pi)} S_n(x-t)u(x) \tag{4.18}$$

where t denotes the variable of integration (and x is fixed in $[0, 2\pi]$). This 'trick' means that the difference is

$$V_n(x) - u(x) = \int_{(0,2\pi)} S_n(x-t)(u(t) - u(x)). \tag{4.19}$$

For each x we split this integral into two parts, the set $\Gamma(x)$ where $x-t \in [0,\delta]$ or $x-t \in [2\pi-\delta, 2\pi]$ and the remainder. So

(4.20)
$$|V_n(x)-u(x)| \le \int_{\Gamma(x)} S_n(x-t)|u(t)-u(x)| + \int_{(0,2\pi)\setminus\Gamma(x)} S_n(x-t)|u(t)-u(x)|.$$

Now on $\Gamma(x)$ either $|t-x| \le \delta$ – the points are close together – or t is close to 0 and x to 2π so $2\pi - x + t \le \delta$ or conversely, x is close to 0 and t to 2π so $2\pi - t + x \le \delta$. In any case, by assuming that $u(0) = u(2\pi)$ and using the uniform continuity of a continuous function on $[0, 2\pi]$, given $\epsilon > 0$ we can choose δ so small that

$$|u(x)-u(t)| \le \epsilon/2 \text{ on } \Gamma(x). \tag{4.21}$$

On the complement of $\Gamma(x)$ we have (4.17) and since u is bounded we get the estimate

$$|V_n(x)-u(x)| \le \epsilon/2 \int_{\Gamma(x)} S_n(x-t) + (n+1)^{-1}q(\delta) \le \epsilon/2 + (n+1)^{-1}q(\delta) \tag{4.22}$$

where $q(\delta) = 2\sin(\delta/2)^{-2}\sup|u|$ is a positive constant depending on δ (and u). Here the fact that S_n is non-negative and has integral one has been used again to estimate the integral of $S_n(x-t)$ over $\Gamma(x)$ by 1. Having chosen δ to make the first term small, we can choose n large to make the second term small and it follows that

$$V_n(x) \to u(x) \text{ uniformly on } [0,2\pi] \text{ as } n \to \infty \tag{4.23}$$

under the assumption that $u \in \mathcal{C}([0,2\pi])$ satisfies $u(0) = u(2\pi)$.

So this proves Proposition 4.1 subject to the density in $L^2(0,2\pi)$ of the continuous functions which vanish near (but not of course in a fixed neighbourhood of) the ends. In fact we know that the L^2 functions which vanish near the ends are dense since we can chop off and use the fact that

$$\lim_{\delta\to 0}\left(\int_{(0,\delta)} |f|^2 + \int_{(2\pi-\delta,2\pi)} |f|^2\right) = 0. \tag{4.24}$$

This proves Theorem 4.1.

Notice that from what we have shown it follows that the finite linear combinations of the $\exp(ikx)$ are dense in the subspace of $\mathcal{C}([0,2\pi])$ consisting of the functions with equal values at the ends. Taking a general element $u \in \mathcal{C}([0,2\pi]$ we can choose constants so that

$$v = u - c - dx \in \mathcal{C}([0,2\pi]) \text{ satisfies } v(0) = v(2\pi) = 0. \tag{4.25}$$

Indeed we just need to take $c = u(0)$, $d = u(1) - c$. Then we know that v is the uniform limit of a sequence of finite sums of the $\exp(ikx)$. However, the Taylor series

$$e^{ikx} = \sum_l \frac{(ik)^l}{l!}x^l \tag{4.26}$$

converges uniformly to e^{ikx} in any (complex) disk. So it follows in turn that the polynomials are dense

THEOREM 4.2 (Stone-Weierstrass). *The polynomials are dense in $\mathcal{C}([a,b])$ for any $a < b$, in the uniform topology.*

Make sure you understand the change of variable argument to get to a general (finite) interval.

2. Toeplitz operators

Although the convergence of Fourier series was stated above for functions on an interval $(0, 2\pi)$ it can be immediately reinterpreted in terms of periodic functions on the line, or equivalently functions on the circle $\mathbb{S}$. Namely a 2π-periodic function

$$u : \mathbb{R} \longrightarrow \mathbb{C},\ u(x + 2\pi) = u(x)\ \forall\ x \in \mathbb{R} \tag{4.27}$$

is uniquely determined by its restriction to $[0, 2\pi)$ by just iterating to see that

$$u(x + 2\pi k) = u(x),\ x \in [0, 2\pi),\ k \in \mathbb{Z}. \tag{4.28}$$

Conversely a function on $[0, 2\pi)$ determines a 2π-periodic function this way. Thus a function on the circle

$$\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\} \tag{4.29}$$

is the same as a periodic function on the line in terms of the standard angular variable

$$\mathbb{S} \ni z = e^{2\pi i\theta},\ \theta \in [0, 2\pi). \tag{4.30}$$

In particular we can identify $L^2(\mathbb{S})$ with $L^2(0, 2\pi)$ in this way – since the missing end-point corresponds to a set of measure zero. Equivalently this identifies $L^2(\mathbb{S})$ as the *locally square integrable* functions on $\mathbb{R}$ which are 2π-periodic.

Since $\mathbb{S}$ is a compact Lie group (what is that you say? Look it up!) this brings us into the realm of harmonic analysis. Just restating the results above for any $u \in L^2(\mathbb{S})$ the Fourier series (thinking of each $\exp(ik\theta)$ as a 2π-periodic function on the line) converges in $L^2(I)$ for any bounded interval

$$u(x) = \sum_{k\in\mathbb{Z}} a_k e^{ikx},\ a_k = \int_{(0,2\pi)} u(x)e^{-ikx}dx. \tag{4.31}$$

After this adjustment of attitude, we follow G.H. Hardy (you might enjoy "A Mathematician's Apology") in thinking about:

Definition 4.1. Hardy space is

$$H = \{u \in L^2(\mathbb{S}); a_k = 0\ \forall\ k < 0\}. \tag{4.32}$$

There are lots of reasons to be interested in $H \subset L^2(\mathbb{S})$ but for the moment note that it is a closed subspace – since it is the intersection of the null spaces of the continuous linear functionals $H \longmapsto a_k,\ k < 0$. Thus there is a unique orthogonal projection

$$\pi_H : L^2(\mathbb{S}) \longrightarrow H \tag{4.33}$$

with range H.

If we go back to the definition of $L^2(\mathbb{S})$ we can see that a continuous function $\alpha \in \mathcal{C}(\mathbb{S})$ defines a bounded linear operator on $L^2(\mathbb{S})$ by multiplication. It is invertible if and only if $\alpha(\theta) \neq 0$ for all $\theta \in [0, 2\pi)$ which is the same as saying that α is a continuous map

$$\alpha : \mathbb{S} \longrightarrow \mathbb{C}^* = \mathbb{C} \setminus \{0\}. \tag{4.34}$$

For such a map there is a well-defined 'winding number' giving the number of times that the curve in the plane defined by α goes around the origin. This is easy to define using the properties of the logarithm. Suppose that α is once continuously differentiable and consider

$$\frac{1}{2\pi i}\int_{[0,2\pi]}\alpha^{-1}\frac{d\alpha}{d\theta}d\theta = \operatorname{wn}(\alpha). \tag{4.35}$$

If we can write

$$\alpha = \exp(2\pi i f(\theta)) \tag{4.36}$$

with $f : [0, 2\pi] \longrightarrow \mathbb{C}$ continuous then necessarily f is differentiable and

$$\operatorname{wn}(\alpha) = \int_0^{2\pi}\frac{df}{d\theta}d\theta = f(2\pi) - f(0) \in \mathbb{Z} \tag{4.37}$$

since $\exp(2\pi i(f(0) - f(2\pi)))) = 1$. In fact, even for a general $\alpha \in \mathcal{C}(\mathbb{S};\mathbb{C}^*)$, it is always possible to find a continuous f satisfying (4.36), using the standard properties of the logarithm as a local inverse to exp, but ill-determined up to addition of integral multiples of $2\pi i$. Then the winding number is given by the last expression in (4.37) and is independent of the choice of f.

Definition 4.2. A Toeplitz operator on H is an operator of the form

$$T_\alpha = \pi_H \alpha \pi_H : H \longrightarrow H,\ \alpha \in \mathcal{C}(\mathbb{S}). \tag{4.38}$$

The result I want is one of the first 'geometric index theorems' – it is a very simple case of the celebrated Atiyah-Singer index theorem (which it much predates).

Theorem 4.3 (Toeplitz). *If $\alpha \in \mathcal{C}(\mathbb{S};\mathbb{C}^*)$ then the Toeplitz operator (4.38) is Fredholm (on the Hardy space H) with index*

$$\operatorname{ind}(T_\alpha) = -\operatorname{wn}(\alpha) \tag{4.39}$$

given in terms of the winding number of α.

Proof. First we need to show that T_α is indeed a Fredholm operator. To do this we decompose the original, multiplication, operator into four pieces

$$\alpha = T_\alpha + H\alpha(\operatorname{Id} - H) + (\operatorname{Id} - H)\alpha H + (\operatorname{Id} - H)\alpha(\operatorname{Id} - H) \tag{4.40}$$

which you can think of a 2×2 matrix corresponding to writing

$$\begin{gathered} L^2(\mathbb{S}) = H \oplus H_-,\ H_- = (\operatorname{Id} - H)L^2(\mathbb{S}), \\ \alpha = \begin{pmatrix} T_\alpha & H\alpha(\operatorname{Id} - H) \\ (\operatorname{Id} - H)\alpha H & (\operatorname{Id} - H)\alpha(\operatorname{Id} - H) \end{pmatrix}. \end{gathered} \tag{4.41}$$

Now, we will show that the two 'off-diagonal' terms are compact operators (on $L^2(\mathbb{S})$). Consider first $(\operatorname{Id} - H)\alpha H$. Now, we showed above, as a form of the Stone-Weierstrass Theorem, that the finite Fourier sums are dense in $\mathcal{C}(\mathbb{S})$ in the supremum norm. This is not the convergence of the Fourier series but there is a sequence $\alpha_k \to \alpha$ in supremum norm, where each

$$\alpha_k = \sum_{j=-N_k}^{N_k} a_{kj} e^{ij\theta}. \tag{4.42}$$

It follows that

$$\|(\operatorname{Id} - H)\alpha_k H - (\operatorname{Id} - H)\alpha H\|_{\mathcal{B}(L^2(\mathbb{S}))} \to 0. \tag{4.43}$$

Now by (4.42) each $(\operatorname{Id} - H)\alpha_k H$ is a finite linear combination of terms

$$(\operatorname{Id} - H)e^{ij\theta}H,\ |j| \le N_k. \tag{4.44}$$

However, each of these operators is of finite rank. They actually vanish if $j \geq 0$ and for $j < 0$ the rank is exactly $-j$. So each $(\mathrm{Id} - H)\alpha_k H$ is of finite rank and hence $(\mathrm{Id} - H)\alpha H$ is compact. A very similar argument works for $H\alpha(\mathrm{Id} - H)$ (or you can use adjoints).

Now, again assume that $\alpha \neq 0$. Then the whole multiplication operator in (4.40) is invertible. If we remove the two compact terms we see that

$$T_\alpha + (\mathrm{Id} - H)\alpha(\mathrm{Id} - H) \text{ is Fredholm.} \tag{4.45}$$

Now the first part maps H to H and the second maps H_- to H_-. It follows that the null space and range of T_α are the projections of the null space and range of the sum (4.45) – so it must have finite dimensional null space and closed range with a finite-dimensional complement as a map from H to itself:-

$$\alpha \in \mathcal{C}(\mathbb{S}; \mathbb{C}^*) \Longrightarrow T_\alpha \text{ is Fredholm in } \mathcal{B}(H). \tag{4.46}$$

So it remains to compute its index. Note that the index of the sum (4.45) acting on $L^2(\mathbb{S})$ vanishes, so that does not really help! The key here is the stability of both the index and the winding number.

Lemma 4.1. *If $\alpha \in \mathcal{C}(\mathbb{S}; \mathbb{C}^*)$ has winding number $p \in \mathbb{Z}$ then there is a curve*

$$\alpha_t : [0, 1] \longrightarrow \mathcal{C}(\mathbb{S}; \mathbb{C}^*),\ \alpha_1 = \alpha,\ \alpha_0 = e^{ip\theta}. \tag{4.47}$$

Proof. If you take a continuous function $f : [0, 2\pi] \longrightarrow \mathbb{C}$ then

$$\alpha = \exp(2\pi i f) \in \mathcal{C}(\mathbb{S}; \mathbb{C}^*) \text{ iff } f(2\pi) = f(0) + p,\ p \in \mathbb{Z} \tag{4.48}$$

(so that $\alpha(2\pi) = \alpha(0)$) where p is precisely the winding number of α. So to construct a continuous family as in (4.47) we can deform f instead provided we keep the difference between the end values constant. Clearly

$$\alpha_t = \exp(2\pi i f_t),\ f_t(\theta) = p\frac{\theta}{2\pi}(1 - t) + f(\theta)t,\ t \in [0, 1] \tag{4.49}$$

does this since $f_t(0) = f(0)t,\ f_t(2\pi) = p(1 - t) + f(2\pi)t = f(0)t + p,\ f_0 = p\frac{\theta}{2\pi}$, $f_1(\theta) = f(\theta)$. □

It was shown above that the index of a Fredholm operator is constant on the components – so along any norm continuous curve such as T_{α_t} where α_t is as in (4.47). Thus the index of T_α, where α has winding number p is the same as the index of the Toeplitz operator defined by $\exp(ip\theta)$, which has the same winding number (note that the winding number is also constant under deformations of α). So we are left to compute the index of the operator $He^{ip\theta}H$ acting on H. This is just a p-fold 'shift up'. If $p \leq 0$ it is actually surjective and has null space spanned by the $\exp(ij\theta)$ with $0 \leq j < -p$ – since these are mapped to $\exp(i(j + p)\theta)$ and hence killed by H. Thus indeed the index of T_α for $\alpha = \exp(ip\theta)$ is k in thus case. For $p > 0$ we can take the adjoint so we have proved Theorem 4.3. □

Why is this important? Suppose you have a function $\alpha \in \mathcal{C}(\mathbb{S}; \mathbb{C}^*)$ and you know it has winding number $-k$ for $k \in \mathbb{N}$. Then you know that the operator T_α must have null space *at least* of dimension k. It could be bigger but this is an existence theorem hence useful. The index is generally *relatively* easy to compute and from that one can tell quite a lot about a Fredholm operator.

3. Cauchy problem

Most, if not all, of you will have had a course on ordinary differential equations so the results here are probably familiar to you at least in outline. I am not going to try to push things very far but I will use the Cauchy problem to introduce 'weak solutions' of differential equations.

So, here is a form of the Cauchy problem. Let me stick to the standard interval we have been using but as usual it does not matter. So we are interested in solutions u of the equation, for some positive integer k

$$
\begin{aligned}
Pu(x) = \frac{d^k u}{dx^k}(x) + \sum_{j=0}^{k-1} a_k(x) \frac{d^j u}{dx^j}(x) &= f(x) \text{ on } [0, 2pi] \\
\frac{d^j u}{dx^j}(0) = 0,\ j = 0, \dots, k-1 & \\
a_j \in \mathcal{C}^j([0, 2\pi]),\ j = 0, \dots, k-1. &
\end{aligned}
\tag{4.50}
$$

So, the a_j are fixed (corresponding if you like to some physical system), u is the 'unknown' function and f is also given. Recall that $\mathcal{C}^j([0, 2\pi])$ is the space (complex valued here) of functions on $[0, 2\pi]$ which have j continuous derivatives. The middle line consists of the 'homogeneous' Cauchy conditions – also called initial conditions – where homogeneous just means zero. The general case of non-zero initial conditions follows from this one.

If we want the equation to make 'classical sense' we need to assume for instance that u has continuous derivatives up to order k and f is continuous. I have written out the first term, involving the highest order of differentiation, in (4.50) separately to suggest the following observation. Suppose u is just k times differentiable, but without assuming the kth derivative is continous. The equation still makes sense but if we assume that f is continuous then it actually follows that u is k times continuously differentiable. In fact each of the terms in the sum is continuous, since this only invovles derivatives up to order $k-1$ multiplied by continuous functions. We can (mentally if you like) move these to the right side of the equation, so together with f this becomes a continuous function. But then the equation itself implies that $\frac{d^k u}{dx^k}$ is continuous and so u is actually k times continuously differentiable. This is a rather trivial example of 'elliptic' regularity which we will push much further.

So, the problem is to prove

THEOREM 4.4. *For each $f \in \mathcal{C}([0, 2\pi])$ there is a unique k times continuously differentiable solution, u, to* (4.50).

Note that in general there is no way of 'writing the solution down'. We can show it exists, and is unique, and we can say a lot about it but there is no formula – although we will see that it is the sum of a reasonable series.

How to proceed? There are many ways but to adopt the one I want to use I need to manipulate the equation in (4.50). There is a certain discriminatory property of the way I have written the equation. Although it seems rather natural, writing the 'coefficients' a_k on the left involves an element of 'handism' if that is a legitimate concept. Instead we could try for the 'rigthist' approach and look at the similar equation

$$\begin{gathered}\frac{d^k u}{dx^k}(x) + \sum_{j=0}^{k-1} \frac{d^j(b_j(x)u)}{dx^j}(x) = f(x) \text{ on } [0, 2\pi] \\ \frac{d^j u}{dx^j}(0) = 0,\ j = 0, \dots, k-1 \\ b_j \in \mathcal{C}^j([0, 2\pi]),\ j = 0, \dots, k-1.\end{gathered} \tag{4.51}$$

As already written in (4.50) we think of P as an operator, sending u to this sume.

LEMMA 4.2. *For any functions $a_j \in \mathcal{C}^j([0, 2\pi])$ there are unique functions $b_j \in \mathcal{C}^j([0, 2\pi])$ so that* (4.51) *gives the same operator as* (4.50).

PROOF. Here we can simply write down a formula for the b_j in terms of the a_j. Namely the product rule for derivatives means that

$$\frac{d^j(b_j(x)u)}{dx^j} = \sum_{p=0}^{j} \binom{j}{p} \frac{d^{j-p} b_j}{dx^{j-p}} \cdot \frac{d^p u}{dx^p}. \tag{4.52}$$

If you are not quite confident that you know this, you do know it for $j = 1$ which is just the usual product rule. So proceed by induction over j and observe that the formul for $j + 1$ follows from the formula for j using the properties of the binomial coefficients.

Pulling out the coefficients of a fixed derivative of u show that we need b_j to satisfy

$$a_p = b_p + \sum_{j=p+1}^{k-1} \binom{j}{p} \frac{d^{j-p} b_j}{dx^{j-p}}. \tag{4.53}$$

This shows the uniquness since we can recover the a_j from the b_j. On the other hand we can solve (4.53) for the b_j too. The 'top' equation says $a_{k-1} = b_{k-1}$ and then successive equations determine b_p in terms of a_p and the b_j with $j > p$ which we already know iteratively.

Note that the $b_j \in \mathcal{C}^j([0, 2\pi])$. □

So, what has been achieved by 'writing the coefficients on the right'? The important idea is that we can solve (4.50) in one particular case, namely when all the a_j (or equivalently b_j) vanish. Then we would just integrate k times. Let us denote Riemann integration by

$$I : \mathcal{C}([0, 2\pi]) \longrightarrow \mathcal{C}([0, 2\pi]),\ If(x) = \int_0^x f(s)ds. \tag{4.54}$$

Of course we can also think of this as Lebesgue integration and then we know for instance that

$$I : L^2(0, 2\pi) \longrightarrow \mathcal{C}([0, 2\pi]) \tag{4.55}$$

is a bounded linear operator. Note also that

$$(If)(0) = 0 \tag{4.56}$$

satisfies the first of the Cauchy conditions.

Now, we can apply the operator I to (4.51) and repeat k times. By the fundamental theorem of calculus

$$u \in \mathcal{C}^j([0,2\pi]),\ \frac{d^p u}{dx^p}(0) = 0,\ p = 0,\dots,j \Longrightarrow I^j(\frac{d^j u}{dx^j}) = u. \tag{4.57}$$

Thus (4.51) becomes

$$(\mathrm{Id} + B)u = u + \sum_{j=0}^{k-1} I^{k-j}(b_j u) = I^k f. \tag{4.58}$$

Notice that this argument is reversible. Namely if $u \in \mathcal{C}^k([0,2\pi])$ satisfies (4.58) for $f \in \mathcal{C}([0,2\pi])$ then $u \in \mathcal{C}^k([0,2\pi])$ does indeed satisfy (4.58). In fact even more is true

Proposition 4.2. *The operator* $\mathrm{Id} + B$ *is invertible on* $L^2(0,2\pi)$ *and if* $f \in \mathcal{C}([0,2\pi])$ *then* $u = (\mathrm{Id} + B)^{-1} I^k f \in \mathcal{C}^k([0,2\pi])$ *is the unique solution of* (4.51).

Proof. From (4.58) we see that B is given as a sum of operators of the form $I^p \circ b$ where b is multiplcation by a continuous function also denoted $b \in \mathcal{C}([0,2\pi])$ and $p \geq 1$. Writing out I^p as an iterated (Riemann) integral

$$I^p v(x) = \int_0^x \int_0^{y_1} \cdots \int_0^{y_{p-1}} v(y_p) dy_p \cdots dy_1. \tag{4.59}$$

For the case of $p = 1$ we can write

$$(I \cdot b_{k-1})v(x) = \int_0^{2\pi} \beta_{k-1}(x,t) v(t) dt,\ \beta_{k-2}(x,t) = H(x-t) b_{k-1}(x) \tag{4.60}$$

where the Heaviside function restricts the integrand to $t \leq x$. Similarly in the next case by reversing the order of integration

$$\begin{aligned} (I^2 \cdot b_{k-2})v(x) &= \int_0^x \int_0^s b(t) v(t) dt ds \\ &= \int_0^x \int^x b_{k-2}(t) v(t) ds dt = \int_0^{2\pi} \beta_{k-2}(x,t) v(t) dt, \\ &\qquad \beta_{k-2} = (x-t)_+ b_{k-2}(x). \end{aligned} \tag{4.61}$$

In general

$$(I^p \cdot b_{k-p})v(x) = \int_0^{2\pi} \beta_{k-p}(x,t) v(t) dt,\ \beta_{k-p} = \frac{1}{(p-1)!}(x-t)_+^{p-1} b_{k-p}(x). \tag{4.62}$$

The explicit formula here is not that important, but (throwing away a lot of information) all the $\beta_*(t,x)$ have the property that they are of the form

$$\beta(x,t) = H(x-t) e(x,t),\ e \in \mathcal{C}([0,2\pi]^2). \tag{4.63}$$

This is a *Volterra operator*

$$Bv(x) = \int_0^{2\pi} \beta(x,t) v(t) \tag{4.64}$$

with β as in (4.63).

So now the point is that for any Volterra operator B, $\mathrm{Id} + B$ is invertible on $L^2(0,2\pi)$. In fact we can make a stronger statement that

$$B \text{ Volterra} \Longrightarrow \sum_j (-1)^j B^j \text{ converges in } \mathcal{B}(L^2(0,2\pi)). \tag{4.65}$$

This is just the Neumann series, but notice we are *not* claiming that $\|B\| < 1$ which would give the convergence as we know from earlier. Rather the key is that the powers of B behave very much like the operators I^k computed above.

LEMMA 4.3. *For a Volterra operator in the sense of* (4.63) *and* (4.64)
(4.66)
$$B^j v(x) = \int_0^{2\pi} H(x-t)e_j(x,t)v(t),\ e_j \in \mathcal{C}([0,2\pi]^2),\ e_j \le \frac{C^j}{(j-1)!}(x-t)_+{}^{j-1},\ j > 1.$$

PROOF. Proceeding inductively we can assume (4.66) holds for a given j. Then $B^{j+1} = B \circ B^j$ is of the form in (4.66) with

$$\begin{aligned} e_{j+1}(x,t) &= \int_0^{2\pi} H(x-s)e(x,s)H(s-t)e_j(s-x)ds \\ &= \int_t^x e(x,s)e_j(s-t)ds \le \sup|e|\frac{C^j}{(j-1)!}\int_t^x (s-t)_+^{j-1}ds \le \frac{C^{j+1}}{j!}(x-t)_+{}^j \end{aligned} \tag{4.67}$$

provided $C \ge \sup|e|$. □

The estimate (4.67) means that, for a different constant

$$\|B^j\|_{L^2} \le \frac{C^j}{j-1}, j > 1 \tag{4.68}$$

which is summable, so the Neumann series (4.58) does converge.

To see the regularity of $u = (\mathrm{Id} + B)^{-1}I^k f$ when $f \in \mathcal{C}([0,2\pi])$ consider (4.58). Each of the terms in the sum maps $L^2(0,2\pi)$ into $\mathcal{C}([0,2\pi])$ so $u \in \mathcal{C}([0.2\pi])$. Proceeding iteratively, for each $p = 0, \dots, k-1$, each of these terms, $I^{k-j}(b_j u)$ maps $\mathcal{C}^p([0,e\pi])$ into $\mathcal{C}^{p+1}([0,2\pi])$ so $u \in \mathcal{C}^k([0,2\pi])$. Similarly for the Cauchy conditions. Differentiating (4.58) recovers (4.51). □

As indicated above, the case of non-vanishing Cauchy data follows from Theorem 4.4. Let

$$\Sigma : \mathcal{C}^k([0,2\pi]) \longrightarrow \mathcal{C}^k \tag{4.69}$$

denote the Cauchy data map – evaluating the function and its first $k-1$ derivatives at 0.

PROPOSITION 4.3. *The combined map*

$$(\Sigma, P) : \mathcal{C}^k([0,2\pi]) \longrightarrow \mathbb{C}^k \oplus \mathcal{C}([0,2\pi]) \tag{4.70}$$

is an isomorphism.

PROOF. The map Σ in (4.69) is certainly surjective, since it is surjective even on polynomials of degree $k-1$. Thus given $z \in \mathbb{C}^k$ there exists $v \in \mathcal{C}^k([0,2\pi])$ with $\Sigma v = z$. Now, given $f \in \mathcal{C}([0,2\pi])$ Theorem 4.4 allows us to find $w \in \mathcal{C}^k([0,2\pi])$ with $Pw = f - Pv$ and $\Sigma w = 0$. So $u = v + w$ satisfies $(\Sigma, P)u = (z, f)$ and we have shown the surjectivity of (4.70). The injectivity again follows from Theorem 4.4 so (Σ, P) is a bijection and hence and isomorphism using the Open Mapping Theorem (or directly). □

Let me finish this discussion of the Cauch problem by introducing the notion of a *weak solution.* let $\Sigma_{2\pi} : \mathcal{C}^k([0,2\pi]) \longrightarrow \mathcal{C}^k$ be the evaluation of the Cauchy data at the top end of the interval. Then if $u \in \mathcal{C}^k([0,2\pi])$ satisfies $\Sigma u = 0$ and

$v \in \mathcal{C}([90, 2\pi])$ satisfies $\Sigma_{2\pi} v = 0$ there are no boundary terms in integration by parts for derivatives up to order k and it follows that

$$\int_{(0,2\pi)} Pu\overline{v} = \int_{(0,2\pi)} u\overline{Qv},\ Qv = (-1)^k \frac{d^k v}{dx^k} + \sum_{j=0}^{k-1} \frac{d^j \overline{a_j} v}{dx^j} \tag{4.71}$$

Thus Q is another operator just like P called the 'formal adjoint' of P.

If $Pu = f$ then (4.71) is just

$$\langle u, Qv\rangle_{L^2} = \langle f, v\rangle_{L^2}\ \forall\ v \in \mathcal{C}^k([0, 2\pi]) \text{ with } \Sigma_{2\pi} v = 0. \tag{4.72}$$

The significant point here is that (4.72) makes sense even if $u,\ f \in L^2([0, 2\pi])$.

DEFINITION 4.3. If $u,\ f \in L^2([0, 2\pi])$ satisfy (4.72) then u is said to be a weak solution of (4.51).

EXERCISE 2. Prove that 'weak=strong' meaning that if $f \in \mathcal{C}([0, 2\pi])$ and $u \in L^2(0, 2\pi)$ is a weak solution of (4.72) then in fact $u \in \mathcal{C}^k([0, 2\pi])$ satisifes (4.51) 'in the classical sense'.

4. Dirichlet problem on an interval

I want to do a couple more 'serious' applications of what we have done so far. There are many to choose from, and I will mention some more, but let me first consider the Diriclet problem on an interval. I will choose the interval $[0, 2\pi]$ because we looked at it before but of course we could work on a general bounded interval instead. So, we are supposed to be trying to solve

$$-\frac{d^2u(x)}{dx^2} + V(x)u(x) = f(x) \text{ on } (0, 2\pi),\ u(0) = u(2\pi) = 0 \tag{4.73}$$

where the last part are the Dirichlet boundary conditions. I will assume that the 'potential'

$$V : [0, 2\pi] \longrightarrow \mathbb{R} \text{ is continuous and real-valued.} \tag{4.74}$$

Now, it certainly makes sense to try to solve the equation (4.73) for say a given $f \in \mathcal{C}([0, 2\pi])$, looking for a solution which is twice continuously differentiable on the interval. It may not exist, depending on V but one thing we can shoot for, which has the virtue of being explicit, is the following:

PROPOSITION 4.4. *If $V \geq 0$ as in (4.74) then for each $f \in \mathcal{C}([0, 2\pi])$ there exists a unique twice continuously differentiable solution, u, to (4.73).*

There are in fact various approaches to this but we want to go through L^2 theory – not surprisingly of course. How to start?

Well, we do know how to solve (4.73) if $V \equiv 0$ since we can use (Riemann) integration. Thus, ignoring the boundary conditions for the moment, we can find a solution to $-d^2v/dx^2 = f$ on the interval by integrating twice:

$$v(x) = -\int_0^x \int_0^y f(t)dtdy \text{ satifies } -d^2v/dx^2 = f \text{ on } (0, 2\pi). \tag{4.75}$$

Moroever v really is twice continuously differentiable if f is continuous. So, what has this got to do with operators? Well, we can change the order of integration in (4.75) to write v as

$$v(x) = -\int_0^x \int_t^x f(t)dydt = \int_0^{2\pi} a(x,t)f(t)dt,\ a(x,t) = (t-x)H(x-t) \tag{4.76}$$

where the Heaviside function $H(y)$ is 1 when $y \geq 0$ and 0 when $y < 0$. Thus $a(x,t)$ is actually continuous on $[0,2\pi]\times[0,2\pi]$ since the $t-x$ factor vanishes at the jump in $H(t-x)$. So (4.76) shows that v is given by applying an integral operator, with continuous kernel on the square, to f.

Before thinking more seriously about this, recall that there is also the matter of the boundary conditions. Clearly, $v(0) = 0$ since we integrated from there. On the other hand, there is no particular reason why

$$v(2\pi) = \int_0^{2\pi} (t - 2\pi) f(t) dt \tag{4.77}$$

should vanish. However, we can always add to v any linear function and still satisfy the differential equation. Since we do not want to spoil the vanishing at $x = 0$ we can only afford to add cx but if we choose the constant c correctly this will work. Namely consider

$$c = \frac{1}{2\pi}\int_0^{2\pi} (2\pi - t) f(t) dt, \text{ then } (v + cx)(2\pi) = 0. \tag{4.78}$$

So, finally the solution we want is

$$w(x) = \int_0^{2\pi} b(x,t) f(t) dt, \ b(x,t) = \min(t,x) - \frac{tx}{2\pi} \in \mathcal{C}([0,2\pi]^2) \tag{4.79}$$

with the formula for b following by simple manipulation from

$$b(x,t) = a(x,t) + x - \frac{tx}{2\pi} \tag{4.80}$$

Thus there is a unique, twice continuously differentiable, solution of $-d^2w/dx^2 = f$ in $(0,2\pi)$ which vanishes at both end points and it is given by the *integral operator* (4.79).

Lemma 4.4. *The integral operator* (4.79) *extends by continuity from* $\mathcal{C}([0,2\pi])$ *to a compact, self-adjoint operator on* $L^2(0,2\pi)$.

Proof. Since w is given by an integral operator with a continuous real-valued kernel which is even in the sense that (check it)

$$b(t,x) = b(x,t) \tag{4.81}$$

we might as well give a more general result. □

Proposition 4.5. *If* $b \in \mathcal{C}([0,2\pi]^2)$ *then*

$$Bf(x) = \int_0^{2\pi} b(x,t) f(t) dt \tag{4.82}$$

defines a compact operator on $L^2(0,2\pi)$ *if in addition* b *satisfies*

$$\overline{b(t,x)} = b(x,t) \tag{4.83}$$

then B *is self-adjoint.*

Proof. If $f \in L^2((0,2\pi))$ and $v \in \mathcal{C}([0,2\pi])$ then the product $vf \in L^2((0,2\pi))$ and $\|vf\|_{L^2} \leq \|v\|_\infty \|f\|_{L^2}$. This can be seen for instance by taking an absolutely summable approximation to f, which gives a sequence of continuous functions converging a.e. to f and bounded by a fixed L^2 function and observing that $vf_n \to vf$

a.e. with bound a constant multiple, $\sup |v|$, of that function. It follows that for $b \in \mathcal{C}([0,2\pi]^2)$ the product

$$b(x,y)f(y) \in L^2(0,2\pi) \tag{4.84}$$

for each $x \in [0,2\pi]$. Thus $Bf(x)$ is well-defined by (4.83) since $L^2((0,2\pi) \subset L^1((0,2\pi))$.

Not only that, but $Bf \in \mathcal{C}([0,2\pi])$ as can be seen from the Cauchy-Schwarz inequality,

(4.85)
$$|Bf(x') - Bf(x)| = |\int (b(x',y) - b(x,y))f(y)| \le \sup_y |b(x',y - b(x,y)|(2\pi)^{\frac{1}{2}} \|f\|_{L^2}.$$

Essentially the same estimate shows that

$$\sup_x |Bf(x)| \le (2\pi)^{\frac{1}{2}} \sup_{(x,y)} |b| \|f\|_{L^2} \tag{4.86}$$

so indeed, $B : L^2(0,2\pi) \longrightarrow \mathcal{C}([0,2\pi])$ is a bounded linear operator.

When b satisfies (4.83) and f and g are continuous

$$\int Bf(x)\overline{g(x)} = \int f(x)\overline{Bg(x)} \tag{4.87}$$

and the general case follows by approximation in L^2 by continuous functions.

So, we need to see the compactness. If we fix x then $b(x,y) \in \mathcal{C}([0,2\pi])$ and then if we let x vary,

$$[0,2\pi] \ni x \longmapsto b(x,\cdot) \in \mathcal{C}([0,2\pi]) \tag{4.88}$$

is continuous as a map into this Banach space. Again this is the uniform continuity of a continuous function on a compact set, which shows that

$$\sup_y |b(x',y) - b(x,y)| \to 0 \text{ as } x' \to x. \tag{4.89}$$

Since the inclusion map $\mathcal{C}([0,2\pi]) \longrightarrow L^2((0,2\pi))$ is bounded, i.e continuous, it follows that the map (I have reversed the variables)

$$[0,2\pi] \ni y \longmapsto b(\cdot,y) \in L^2((0,2\pi)) \tag{4.90}$$

is continuous and so has a compact range.

Take the Fourier basis e_k for $[0,2\pi]$ and expand b in the first variable. Given $\epsilon > 0$ the compactness of the image of (4.90) implies that the Fourier Bessel series converges uniformly (has uniformly small tails), so for some N

$$\sum_{|k|>N} |(b(x,y),e_k(x))|^2 < \epsilon \;\forall\; y \in [0,2\pi]. \tag{4.91}$$

The finite part of the Fourier series is continuous as a function of both arguments

$$b_N(x,y) = \sum_{|k|\le N} e_k(x)c_k(y),\; c_k(y) = (b(x,y),e_k(x)) \tag{4.92}$$

and so defines another bounded linear operator B_N as before. This operator can be written out as

$$B_N f(x) = \sum_{|k|\le N} e_k(x) \int c_k(y)f(y)dy \tag{4.93}$$

and so is of finite rank – it always takes values in the span of the first $2N+1$ trigonometric functions. On the other hand the remainder is given by a similar operator with corresponding to $q_N = b - b_N$ and this satisfies

$$\sup_y \|q_N(\cdot, y)\|_{L^2((0,2\pi))} \to 0 \text{ as } N \to \infty. \tag{4.94}$$

Thus, q_N has small norm as a bounded operator on $L^2((0,2\pi))$ so B is compact – it is the norm limit of finite rank operators. □

Now, recall from Problem# that $u_k = \pi^{-\frac{1}{2}} \sin(kx/2)$, $k \in \mathbb{N}$, is also an orthonormal basis for $L^2(0,2\pi)$ (it is not the Fourier basis!) Moreover, differentiating we find straight away that

$$-\frac{d^2 u_k}{dx^2} = \frac{k^2}{4} u_k. \tag{4.95}$$

Since of course $u_k(0) = 0 = u_k(2\pi)$ as well, from the uniqueness above we conclude that

$$Bu_k = \frac{4}{k^2} u_k \ \forall\ k. \tag{4.96}$$

Thus, in this case we know the orthonormal basis of eigenfunctions for B. They are the u_k, each eigenspace is 1 dimensional and the eigenvalues are $4k^{-2}$.

Remark 4.1. As noted by Pavel Etingof it is worthwhile to go back to the discussion of trace class operators to see that B is indeed of trace class. Its trace can be computed in two ways. As the sum of its eigenvalues and as the integral of its kernel on the diagonal. This gives the well-known formula

$$\frac{\pi^2}{6} = \sum_{k \in \mathbb{N}} \frac{1}{k^2}. \tag{4.97}$$

This is a simple example of a 'trace formula'; you might like to look up some others!

So, this happenstance allows us to decompose B as the square of another operator defined directly on the othornormal basis. Namely

$$Au_k = \frac{2}{k} u_k \Longrightarrow B = A^2. \tag{4.98}$$

Here again it is immediate that A is a compact self-adjoint operator on $L^2(0,2\pi)$ since its eigenvalues tend to 0. In fact we can see quite a lot more than this.

Lemma 4.5. *The operator A maps $L^2(0,2\pi)$ into $\mathcal{C}([0,2\pi])$ and $Af(0) = Af(2\pi) = 0$ for all $f \in L^2(0,2\pi)$.*

Proof. If $f \in L^2(0,2\pi)$ we may expand it in Fourier-Bessel series in terms of the u_k and find

$$f = \sum_k c_k u_k,\ \{c_k\} \in l^2. \tag{4.99}$$

Then of course, by definition,

$$Af = \sum_k \frac{2c_k}{k} u_k. \tag{4.100}$$

Here each u_k is a bounded continuous function, with the bound $|u_k| \leq C$ being independent of k. So in fact (4.100) converges uniformly and absolutely since it is uniformly Cauchy, for any $q > p$,

$$|\sum_{k=p}^{q} \frac{2c_k}{k} u_k| \leq 2C \sum_{k=p}^{q} |c_k| k^{-1} \leq 2C \left(\sum_{k=p}^{q} k^{-2} \right)^{\frac{1}{2}} \|f\|_{L^2} \tag{4.101}$$

where Cauchy-Schwarz has been used. This proves that

$$A : L^2(0, 2\pi) \longrightarrow \mathcal{C}([0, 2\pi])$$

is bounded and by the uniform convergence $u_k(0) = u_k(2\pi) = 0$ for all k implies that $Af(0) = Af(2\pi) = 0$. □

So, going back to our original problem we try to solve (4.73) by moving the Vu term to the right side of the equation (don't worry about regularity yet) and hope to use the observation that

$$u = -A^2(Vu) + A^2 f \tag{4.102}$$

should satisfy the equation and boundary conditions. In fact, let's anticipate that $u = Av$, which has to be true if (4.102) holds with $v = -AVu + Af$, and look instead for

$$v = -AVAv + Af \Longrightarrow (\mathrm{Id} + AVA)v = Af. \tag{4.103}$$

So, we know that multiplication by V, which is real and continuous, is a bounded self-adjoint operator on $L^2(0, 2\pi)$. Thus AVA is a self-adjoint compact operator so we can apply our spectral theory to it and so examine the invertibility of $\mathrm{Id} + AVA$. Working in terms of a complete orthonormal basis of eigenfunctions e_i of AVA we see that $\mathrm{Id} + AVA$ is invertible if and only if it has trivial null space, i.e. if -1 is *not* an eigenvalue of AVA. Indeed, an element of the null space would have to satisfy $u = -AVAu$. On the other hand we know that AVA is *positive* since

$$(AVAw, w) = (VAv, Av) = \int_{(0,2\pi)} V(x)|Av|^2 \geq 0 \Longrightarrow \int_{(0,2\pi)} |u|^2 = 0, \tag{4.104}$$

using the assumed non-negativity of V. So, there can be no null space – all the eigenvalues of AVA are at least non-negative and the inverse is the bounded operator given by its action on the basis

$$(\mathrm{Id} + AVA)^{-1} e_i = (1 + \tau_i)^{-1} e_i, \ AVAe_i = \tau_i e_i. \tag{4.105}$$

Thus $\mathrm{Id} + AVA$ is invertible on $L^2(0, 2\pi)$ with inverse of the form $\mathrm{Id} + Q$, Q again compact and self-adjoint since $(1 + \tau_i)^{-1} - 1 \to 0$. Now, to solve (4.103) we just need to take

$$v = (\mathrm{Id} + Q)Af \Longleftrightarrow v + AVAv = Af \text{ in } L^2(0, 2\pi). \tag{4.106}$$

Then indeed

$$u = Av \text{ satisfies } u + A^2 Vu = A^2 f. \tag{4.107}$$

In fact since $v \in L^2(0, 2\pi)$ from (4.106) we already know that $u \in \mathcal{C}([0, 2\pi])$ vanishes at the end points.

Moreover if $f \in \mathcal{C}([0, 2\pi])$ we know that $Bf = A^2 f$ is twice continuously differentiable, since it is given by two integrations – that is where B came from. Now, we know that u in L^2 satisfies $u = -A^2(Vu) + A^2 f$. Since $Vu \in L^2((0, 2\pi) \Longrightarrow$

$A(Vu) \in L^2(0, 2\pi)$ and then, as seen above, $A(A(Vu)$ is continuous. So combining this with the result about $A^2 f$ we see that u itself is continuous and hence so is Vu. But then, going through the routine again

$$u = -A^2(Vu) + A^2 f \tag{4.108}$$

is the sum of two twice continuously differentiable functions. Thus it is so itself. In fact from the properties of $B = A^2$ it satisifes

$$-\frac{d^2u}{dx^2} = -Vu + f \tag{4.109}$$

which is what the result claims. So, we have proved the existence part of Proposition 4.4.

The uniqueness follows pretty much the same way. If there were two twice continuously differentiable solutions then the difference w would satisfy

$$-\frac{d^2w}{dx^2} + Vw = 0,\ w(0) = w(2\pi) = 0 \Longrightarrow w = -Bw = -A^2Vw. \tag{4.110}$$

Thus $w = A\phi$, $\phi = -AVw \in L^2(0, 2\pi)$. Thus ϕ in turn satisfies $\phi = AVA\phi$ and hence is a solution of $(\mathrm{Id} + AVA)\phi = 0$ which we know has none (assuming $V \geq 0$). Since $\phi = 0$, $w = 0$.

This completes the proof of Proposition 4.4. To summarize, what we have shown is that $\mathrm{Id} + AVA$ is an invertible bounded operator on $L^2(0, 2\pi)$ (if $V \geq 0$) and then the solution to (4.73) is precisely

$$u = A(\mathrm{Id} + AVA)^{-1}Af \tag{4.111}$$

which is twice continuously differentiable and satisfies the Dirichlet conditions for each $f \in \mathcal{C}([0, 2\pi])$.

This may seem a 'round-about' approach but it is rather typical of methods from Functional Analysis. What we have done is to separate the two problems of 'existence' and 'regularity'. We first get existence of what is often called a 'weak solution' of the problem, in this case given by (4.111), which is in $L^2(0, 2\pi)$ for $f \in L^2(0, 2\pi)$ and then show, given regularity of the right hand side f, that this is actually a 'classical solution'.

Even if we do not assume that $V \geq 0$ we can see fairly directly what is happening.

Theorem 4.5. *For any $V \in \mathcal{C}([0, 2\pi])$ real-valued, there is an orthonormal basis w_k of $L^2(0, 2\pi)$ consisting of twice-continuously differentiable functions on $[0, 2\pi]$, vanishing at the end-points and satisfying $-\frac{d^2w_k}{dx^2} + Vw_k = T_k w_k$ where $T_k \to \infty$ as $k \to \infty$. The equation (4.73) has a (twice continuously differentiable) solution for given $f \in \mathcal{C}([0, 2\pi])$ if and only if*

$$T_k = 0 \Longrightarrow \int_{(0,2\pi)} f w_k = 0. \tag{4.112}$$

Proof. For a real-valued V we can choose a constant c such that $V + c \geq 0$. Then the eigenvalue equation we are trying to solve can be rewritten

$$-\frac{d^2w}{dx^2} + Vw = Tw \Longleftrightarrow -\frac{d^2w}{dx^2} + (V + c)w = (T + c)w. \tag{4.113}$$

Proposition 4.4 shows that there is indeed an orthonormal basis of solutions of the second equation, w_k as in the statement above with positive eigenvalues $T_k+c \to \infty$ with k.

So, only the solvability of (4.73) remains to be checked. What we have shown above is that if $f \in \mathcal{C}([0,2\pi])$ then a twice continuously differentiable solution to

$$-\frac{d^2w}{dx^2} + Vw = f,\ w(0) = w(2\pi) = 0 \tag{4.114}$$

is precisely of the form $w = Av$ where

$$v \in L^2(0,2\pi),\ (\mathrm{Id} + AVA)v = Af. \tag{4.115}$$

Going from (4.114) to (4.115) involves the properties of $B = A^2$ since (4.114) implies that

$$w + A^2Vw = A^2f \Longrightarrow w = Av \text{ with } v \text{ as in (4.115)}$$

Conversely if v satisfies (4.115) then $w = Av$ satisfies $w + A^2Vw = A^2f$ which implies that w has the correct regularity and satisfies (4.114).

Applying this to the case $f = 0$ shows that for twice continuously differentiable functions on $[0, 2\pi]$,

$$-\frac{d^2w}{dx^2} + Vw = 0,\ w(0) = w(2\pi) = 0 \Longleftrightarrow \\ w \in A\{v \in L^2(0,2\pi); (\mathrm{Id} + AVA)v = 0\}. \tag{4.116}$$

Since AVA is compact and self-adjoint we see that

$$(\mathrm{Id} + AVA)v = Af \text{ has a solution in } L^2(0,2\pi) \Longrightarrow \\ Af \perp \{v' \in L^2(0,2\pi); (\mathrm{Id} + AVA)v' = 0\}. \tag{4.117}$$

However this last condition is equivalent to $f \perp A\{v \in L^2(0,2\pi); (\mathrm{Id} + AVA)v = 0\}$ which is by the equivalence of (4.114) and (4.115) reduces precisely to (4.112). $\square$

So, ultimately the solution of the differential equation (4.73) is just like the solution of a finite dimensional problem for a self-adjoint matrix. There is a solution if and only if the right side is orthogonal to the null space; it just requires a bit more work. Lots of 'elliptic' problems turn out to be like this.

We can also say (a great deal) more about the eigenvalues T_k and eigenfunctions w_k. For instance, the derivative

$$w_k'(0) \neq 0. \tag{4.118}$$

Indeed, were this to vanish w_k would be a solution of the Cauchy problem (4.50) for the second-order operator $P = \frac{d^2}{dx^2} - q + T_k$ with 'forcing term' $f = 0$ and hence, by Theorem 4.4 must itself vanish on the interval, which is a contradiction.

From this in turn it follows that the (non-trivial) eigenspaces

$$E_k(q) = \{w \in \mathcal{C}^2([0,2\pi]); -\frac{d^2w}{dx^2} + qw = T_kw\} \tag{4.119}$$

are exactly one-dimensional. Indeed if w_k is one non-zero element, so satisfing (4.118) and w is another, then $w - w'(0)w_k/w_k'(0) \in E_k(q)$ again must satisfy the Cauchy conditions at 0 so $w = w'(0)w_k/w_k'(0)$.

EXERCISE 3. Show that the eigenfunctions functions normalized to have $w_k'(0) = 1$ are all real and w_k has exactly $k-1$ zeros in the interior of the interval.

You could try your hand at proving Borg's Theorem – if $q \in \mathcal{C}([0, 2\pi])$ and the eigenvalues $T_k = \frac{k^2}{4}$ are the same as those for $q = 0$ then $q = 0$! This is the beginning of a large theory of the inverse problem – to what extent can one recover q from the knowledge of the T_k? In brief the answer is that one cannot do so in general. However q *is* determined if one knows the T_k and the 'norming constants' $w'_k(2\pi)/w'_k(0)$.

5. Harmonic oscillator

As a second 'serious' application of our Hilbert space theory I want to discuss the harmonic oscillator, the corresponding Hermite basis for $L^2(\mathbb{R})$. Note that so far we have not found an explicit orthonormal basis on the whole real line, even though we know $L^2(\mathbb{R})$ to be separable, so we certainly know that such a basis exists. How to construct one explicitly and with some handy properties? One way is to simply orthonormalize – using Gram-Schmidt – some countable set with dense span. For instance consider the basic Gaussian function

$$\exp(-\frac{x^2}{2}) \in L^2(\mathbb{R}). \tag{4.120}$$

This is so rapidly decreasing at infinity that the product with any polynomial is also square integrable:

$$x^k \exp(-\frac{x^2}{2}) \in L^2(\mathbb{R}) \ \forall\ k \in \mathbb{N}_0 = \{0, 1, 2, \dots\}. \tag{4.121}$$

Orthonormalizing this sequence gives an orthonormal basis, where completeness can be shown by an appropriate approximation technique but as usual is not so simple. This is in fact the Hermite basis as we will eventually show.

Rather than proceed directly we will work up to this by discussing the eigenfunctions of the harmonic oscillator

$$P = -\frac{d^2}{dx^2} + x^2 \tag{4.122}$$

which we want to think of as an operator – although for the moment I will leave vague the question of what it operates *on*.

As you probably already know, and we will show later, it is straightforward to show that P has a lot of eigenvectors using the 'creation' and 'annihilation' operators. We want to know a bit more than this and in particular I want to apply the abstract discussion above to this case but first let me go through the 'formal' theory. There is nothing wrong here, just that we cannot easily conclude the completeness of the eigenfunctions.

The first thing to observe is that the Gaussian is an eigenfunction of H

$$\begin{aligned} Pe^{-x^2/2} &= -\frac{d}{dx}(-xe^{-x^2/2} + x^2e^{-x^2/2}) \\ &= -(x^2 - 1)e^{-x^2/2} + x^2e^{-x^2/2} = e^{-x^2/2} \end{aligned} \tag{4.123}$$

with eigenvalue 1. It is an eigenfunction but not, for the moment, of a bounded operator on any Hilbert space – in this sense it is only a formal eigenfunction.

In this special case there is an essentially algebraic way to generate a whole sequence of eigenfunctions from the Gaussian. To do this, write

$$Pu = (-\frac{d}{dx}+x)(\frac{d}{dx}+x)u + u = (\operatorname{Cr}\operatorname{An}+1)u,\qquad \operatorname{Cr} = (-\frac{d}{dx}+x),\ \operatorname{An} = (\frac{d}{dx}+x) \tag{4.124}$$

again formally as operators. Then note that

$$\operatorname{An} e^{-x^2/2} = 0 \tag{4.125}$$

which again proves (4.123). The two operators in (4.124) are the 'creation' operator and the 'annihilation' operator. They almost commute in the sense that

$$[\operatorname{An},\operatorname{Cr}]u = (\operatorname{An}\operatorname{Cr} - \operatorname{Cr}\operatorname{An})u = 2u \tag{4.126}$$

for say any twice continuously differentiable function u.

Now, set $u_0 = e^{-x^2/2}$ which is the 'ground state' and consider $u_1 = \operatorname{Cr} u_0$. From (4.126), (4.125) and (4.124),

$$Pu_1 = (\operatorname{Cr}\operatorname{An}\operatorname{Cr} + \operatorname{Cr})u_0 = \operatorname{Cr}^2\operatorname{An} u_0 + 3\operatorname{Cr} u_0 = 3u_1. \tag{4.127}$$

Thus, u_1 is an eigenfunction with eigenvalue 3.

Lemma 4.6. *For $j \in \mathbb{N}_0 = \{0,1,2,\dots\}$ the function $u_j = \operatorname{Cr}^j u_0$ satisfies $Pu_j = (2j+1)u_j$.*

Proof. This follows by induction on j, where we know the result for $j = 0$ and $j = 1$. Then

$$P\operatorname{Cr} u_j = (\operatorname{Cr}\operatorname{An}+1)\operatorname{Cr} u_j = \operatorname{Cr}(P-1)u_j + 3\operatorname{Cr} u_j = (2j+3)u_j. \tag{4.128}$$

□

Again by induction we can check that $u_j = (2^jx^j + q_j(x))e^{-x^2/2}$ where q_j is a polynomial of degree at most $j-2$. Indeed this is true for $j = 0$ and $j = 1$ (where $q_0 = q_1 \equiv 0$) and then

$$\operatorname{Cr} u_j = (2^{j+1}x^{j+1} + \operatorname{Cr} q_j)e^{-x^2/2}. \tag{4.129}$$

and $q_{j+1} = \operatorname{Cr} q_j$ is a polynomial of degree at most $j-1$ – one degree higher than q_j.

From this it follows in fact that the finite span of the u_j consists of all the products $p(x)e^{-x^2/2}$ where $p(x)$ is any polynomial.

Now, all these functions are in $L^2(\mathbb{R})$ and we want to compute their norms. First, a standard integral computation[1] shows that

$$\int_{\mathbb{R}} (e^{-x^2/2})^2 = \int_{\mathbb{R}} e^{-x^2} = \sqrt{\pi} \tag{4.130}$$

[1]To compute the Gaussian integral, square it and write as a double integral then introduce polar coordinates

$$(\int_{\mathbb{R}} e^{-x^2}dx)^2 = \int_{\mathbb{R}^2} e^{-x^2-y^2}dxdy = \int_0^\infty\int_0^{2\pi} e^{-r^2}rdrd\theta = \pi\big[-e^{-r^2}\big]_0^\infty = \pi.$$

For $j > 0$, integration by parts (easily justified by taking the integral over $[-R, R]$ and then letting $R \to \infty$) gives

$$\int_{\mathbb{R}} (\mathrm{Cr}^j u_0)^2 = \int_{\mathbb{R}} \mathrm{Cr}^j u_0(x) \,\mathrm{Cr}^j u_0(x) dx = \int_{\mathbb{R}} u_0 \,\mathrm{An}^j \,\mathrm{Cr}^j u_0. \tag{4.131}$$

Now, from (4.126), we can move one factor of An through the j factors of Cr until it emerges and 'kills' u_0

$$\begin{aligned} \mathrm{An}\,\mathrm{Cr}^j u_0 &= 2\,\mathrm{Cr}^{j-1} u_0 + \mathrm{Cr}\,\mathrm{An}\,\mathrm{Cr}^{j-1} u_0 \\ &= 2\,\mathrm{Cr}^{j-1} u_0 + \mathrm{Cr}^2 \,\mathrm{An}\,\mathrm{Cr}^{j-2} u_0 = 2j\,\mathrm{Cr}^{j-1} u_0. \end{aligned} \tag{4.132}$$

So in fact,

$$\int_{\mathbb{R}} (\mathrm{Cr}^j u_0)^2 = 2j \int_{\mathbb{R}} (\mathrm{Cr}^{j-1} u_0)^2 = 2^j j! \sqrt{\pi}. \tag{4.133}$$

A similar argument shows that

$$\int_{\mathbb{R}} u_k u_j = \int_{\mathbb{R}} u_0 \,\mathrm{An}^k \,\mathrm{Cr}^j u_0 = 0 \text{ if } k \neq j. \tag{4.134}$$

Thus the functions

$$e_j = 2^{-j/2} (j!)^{-\frac{1}{2}} \pi^{-\frac{1}{4}} C^j e^{-x^2/2} \tag{4.135}$$

form an orthonormal sequence in $L^2(\mathbb{R})$.

We would like to show this orthonormal sequence is complete. Rather than argue through approximation, we can guess that in some sense the operator

$$\mathrm{An}\,\mathrm{Cr} = (\frac{d}{dx} + x)(-\frac{d}{dx} + x) = -\frac{d^2}{dx^2} + x^2 + 1 \tag{4.136}$$

should be invertible, so one approach is to use the ideas above of Friedrichs' extension to construct its 'inverse' and show this really exists as a compact, self-adjoint operator on $L^2(\mathbb{R})$ and that its only eigenfunctions are the e_i in (4.135). Another, more indirect approach is described below.

6. Fourier transform

The Fourier transform for functions on $\mathbb{R}$ is in a certain sense the limit of the definition of the coefficients of the Fourier series on an expanding interval, although that is not generally a good way to approach it. We know that if $u \in L^1(\mathbb{R})$ and $v \in \mathcal{C}_\infty(\mathbb{R})$ is a bounded continuous function then $vu \in L^1(\mathbb{R})$ – this follows from our original definition by approximation. So if $u \in L^1(\mathbb{R})$ the integral

$$\hat{u}(\xi) = \int e^{-ix\xi} u(x) dx, \ \xi \in \mathbb{R} \tag{4.137}$$

exists for each $\xi \in \mathbb{R}$ as a Lebesgue integral. Note that there are many different normalizations of the Fourier transform in use. This is the standard 'analyst's' normalization.

Proposition 4.6. *The Fourier tranform,* (4.137)*, defines a bounded linear map*

$$\mathcal{F} : L^1(\mathbb{R}) \ni u \longmapsto \hat{u} \in \mathcal{C}_0(\mathbb{R}) \tag{4.138}$$

into the closed subspace $\mathcal{C}_0(\mathbb{R}) \subset \mathcal{C}_\infty(\mathbb{R})$ of continuous functions which vanish at infinity (with respect to the supremum norm).

PROOF. We know that the integral exists for each ξ and from the basic properties of the Lebesgue integal

$$|\hat{u}(\xi)| \le \|u\|_{L^1}, \text{ since } |e^{-ix\xi}u(x)| = |u(x)|. \tag{4.139}$$

To investigate its properties we restrict to $u \in \mathcal{C}_{\mathrm{c}}(\mathbb{R})$, a compactly-supported continuous function, say with support in $-R, R]$. Then the integral becomes a Riemann integral and the integrand is a continuous function of both variables. It follows that the Fourier transform is uniformly continuous since

$$|\hat{u}(\xi) - \hat{u}(\xi')| \le \int_{|x|\le R} |e^{-ix\xi} - e^{-ix\xi'}||u(x)|dx \le 2R\sup|u| \sup_{|x|\le R} |e^{-ix\xi} - e^{-ix\xi'}| \tag{4.140}$$

with the right side small by the uniform continuity of continuous functions on compact sets. From (4.139), if $u_n \to u$ in $L^1(\mathbb{R})$ with $u_n \in \mathcal{C}_{\mathrm{c}}(\mathbb{R})$ it follows that $\hat{u}_n \to \hat{u}$ uniformly on $\mathbb{R}$. Thus, as the uniform limit of uniformly continuous functions, the Fourier transform is uniformly continuous on $\mathbb{R}$ for any $u \in L^1(\mathbb{R})$ (you can also see this from the continuity-in-the-mean of L^1 functions).

Now, we know that even the compactly-supported once continuously differentiable functions, forming $\mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$ are dense in $L^1(\mathbb{R})$ so we can also consider (4.137) where $u \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$. Then the integration by parts as follows is justified

$$\xi\hat{u}(\xi) = i\int (\frac{de^{-ix\xi}}{dx})u(x)dx = -i\int e^{-ix\xi}\frac{du(x)}{dx}dx. \tag{4.141}$$

Since $du/dx \in \mathcal{C}_{\mathrm{c}}(\mathbb{R})$ (by assumption) the estimate (4.139) now gives

$$\sup_{\xi\in\mathbb{R}} |\xi\hat{u}(\xi)| \le 2R\sup_{x\in\mathbb{R}} |\frac{du}{dx}|. \tag{4.142}$$

This certainly implies the weaker statement that

$$\lim_{|\xi|\to\infty} |\hat{u}(\xi)| = 0 \tag{4.143}$$

which is 'vanishing at infinity'. Now we again use the density, this time of $\mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$, in $L^1(\mathbb{R})$ and the uniform estimate (4.139), plus the fact that if a sequence of continuous functions on $\mathbb{R}$ converges uniformly on $\mathbb{R}$ and each element vanishes at infinity then the limit vanishes at infinity to complete the proof of the Proposition. □

7. Fourier inversion

We could use the completeness of the orthonormal sequence of eigenfunctions for the harmonic oscillator discussed above to show that the Fourier tranform extends by continuity from $\mathcal{C}_{\mathrm{c}}(\mathbb{R})$ to define an isomorphism

$$\mathcal{F} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R}) \tag{4.144}$$

with inverse given by the corresponding continuous extension of

$$\mathcal{G}v(x) = (2\pi)^{-1}\int e^{ix\xi}v(\xi). \tag{4.145}$$

Instead, we will give a direct proof of the Fourier inversion formula, via Schwartz space and an elegant argument due to Hörmander. Then we will use this to prove the completeness of the eigenfunctions we have found.

We have shown above that the Fourier transform is defined as an integral if $u \in L^1(\mathbb{R})$. Suppose that in addition we know that $xu \in L^1(\mathbb{R})$. We can summarize the combined information as

$$(1 + |x|)u \in L^1(\mathbb{R}). \tag{4.146}$$

LEMMA 4.7. *If u satisfies* (4.146) *then $\hat{u}$ is continuously differentiable and $d\hat{u}/d\xi = \mathcal{F}(-ixu)$ is bounded.*

PROOF. Consider the difference quotient for the Fourier transform:

$$\frac{\hat{u}(\xi + s) - \hat{u}(\xi)}{s} = \int D(x, s)e^{-ix\xi}u(x), \ D(x, s) = \frac{e^{-ixs} - 1}{s}. \tag{4.147}$$

We can use the standard proof of Taylor's formula to write the difference quotient inside the integral as

$$D(x, s) = -ix \int_0^1 e^{-itxs} dt \Longrightarrow |D(x, s)| \le |x|. \tag{4.148}$$

It follows that as $s \to 0$ (along a sequence if you prefer) $D(x, s)e^{-ix\xi}u(x)$ is bounded by the $L^1(\mathbb{R})$ function $|x||u(x)|$ and converges pointwise to $-ie^{-ix\xi}xu(x)$. Dominated convergence therefore shows that the integral converges showing that the derivative exists and that

$$\frac{d\hat{u}(\xi)}{d\xi} = \mathcal{F}(-ixu). \tag{4.149}$$

From the earlier results it follows that the derivative is continuous and bounded, proving the lemma. $\square$

Now, we can iterate this result and so conclude:

$$\begin{gathered} (1 + |x|)^k u \in L^1(\mathbb{R}) \ \forall \ k \Longrightarrow \\ \hat{u} \text{ is infinitely differentiable with bounded derivatives and} \\ \frac{d^k \hat{u}}{d\xi^k} = \mathcal{F}((-ix)^k u). \end{gathered} \tag{4.150}$$

This result shows that from 'decay' of u we deduce smoothness of $\hat{u}$. We can go the other way too. One way to ensure the assumption in (4.150) is to make the stronger assumption that

$$x^k u \text{ is bounded and continuous } \forall \ k. \tag{4.151}$$

Indeed, Dominated Convergence shows that if u is continuous and satisfies the bound

$$|u(x)| \le (1 + |x|)^{-r}, \ r > 1$$

then $u \in L^1(\mathbb{R})$. So the integrability of $x^j u$ follows from the bounds in (4.151) for $k \le j + 2$. This is throwing away information but simplifies things below.

In the opposite direction, suppose that u is continuously differentiable and satisfies the estimates for some $r > 1$

$$|u(x)| + |\frac{du(x)}{dx}| \le C(1 + |x|)^{-r}.$$

Then consider

$$\xi\hat{u} = i \int \frac{de^{-ix\xi}}{dx} u(x) = \lim_{R \to \infty} i \int_{-R}^{R} \frac{de^{-ix\xi}}{dx} u(x). \tag{4.152}$$

We may integrate by parts in this integral to get

$$\xi\hat{u} = \lim_{R\to\infty}\left(i\big[e^{-ix\xi}u(x)\big]_{-R}^{R} - i\int_{-R}^{R} e^{-ix\xi}\frac{du}{dx}\right). \tag{4.153}$$

The decay of u shows that the first term vanishes in the limit and the second integral converges so

$$\xi\hat{u} = \mathcal{F}(-i\frac{du}{dx}). \tag{4.154}$$

Iterating this in turn we see that if u has continuous derivatives of all orders and for all j

$$|\frac{d^j u}{dx^j}| \le C_j(1+|x|)^{-r},\ r>1 \text{ then the } \xi^j\hat{u} = \mathcal{F}((-i)^j\frac{d^j u}{dx^j}) \tag{4.155}$$

are all bounded.

Laurent Schwartz defined a space which handily encapsulates these results.

Definition 4.4. Schwartz space, $\mathcal{S}(\mathbb{R})$, consists of all the infinitely differentiable functions $u:\mathbb{R}\longrightarrow\mathbb{C}$ such that

$$\|u\|_{j,k} = \sup|x^j\frac{d^k u}{dx^k}| < \infty\ \forall\ j,\ k\ge 0. \tag{4.156}$$

This is clearly a linear space. In fact it is a complete metric space in a natural way. All the $\|\cdot\|_{j,k}$ in (4.156) are norms on $\mathcal{S}(\mathbb{R})$, but none of them is stronger than the others. So there is no natural norm on $\mathcal{S}(\mathbb{R})$ with respect to which it is complete. In the problems below you can find some discussion of the fact that

$$d(u,v) = \sum_{j,k\ge 0} 2^{-j-k}\frac{\|u-v\|_{j,k}}{1+\|u-v\|_{j,k}} \tag{4.157}$$

is a complete metric. We will not use this here but it is the right way to understand what is going on.

Notice that there is some prejudice on the order of multiplication by x and differentiation in (4.156). This is only apparent, since these estimates (taken together) are equivalent to

$$\sup|\frac{d^k(x^j u)}{dx^k}| < \infty\ \forall\ j,\ k\ge 0. \tag{4.158}$$

To see the equivalence we can use induction over N where the inductive statement is the equivalence of (4.156) and (4.158) for $j+k\le N$. Certainly this is true for $N=0$ and to carry out the inductive step just differentiate out the product to see that

$$\frac{d^k(x^j u)}{dx^k} = x^j\frac{d^k u}{dx^k} + \sum_{l+m<k+j} c_{l,m,k,j}x^m\frac{d^l u}{dx^l}$$

where one can be much more precise about the extra terms, but the important thing is that they all are lower order (in fact both degrees go down). If you want to be careful, you can of course prove this identity by induction too! The equivalence of (4.156) and (4.158) for $N+1$ now follows from that for N.

Theorem 4.6. *The Fourier transform restricts to a bijection on $\mathcal{S}(\mathbb{R})$ with inverse*

$$\mathcal{G}(v)(x) = \frac{1}{2\pi}\int e^{ix\xi}v(\xi). \tag{4.159}$$

PROOF. The proof (due to Hörmander as I said above) will take a little while because we need to do some computation, but I hope you will see that it is quite clear and elementary.

First we need to check that $\mathcal{F} : \mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\mathbb{R})$, but this is what I just did the preparation for. Namely the estimates (4.156) imply that (4.155) applies to all the $\frac{d^k(x^j u)}{dx^k}$ and so

$$\xi^k \frac{d^j \hat{u}}{d\xi^j} \text{ is continuous and bounded } \forall\ k,\ j \Longrightarrow \hat{u} \in \mathcal{S}(\mathbb{R}). \tag{4.160}$$

This indeed is why Schwartz introduced this space.

So, what we want to show is that with $\mathcal{G}$ defined by (4.159), $u = \mathcal{G}(\hat{u})$ for all $u \in \mathcal{S}(\mathbb{R})$. Notice that there is only a sign change and a constant factor to get from $\mathcal{F}$ to $\mathcal{G}$ so certainly $\mathcal{G} : \mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\mathbb{R})$. We start off with what looks like a small part of this. Namely we want to show that

$$I(\hat{u}) = \int \hat{u} = 2\pi u(0). \tag{4.161}$$

Here, $I : \mathcal{S}(\mathbb{R}) \longrightarrow \mathbb{C}$ is just integration, so it is certainly well-defined. To prove (4.161) we need to use a version of Taylor's formula and then do a little computation.

LEMMA 4.8. *If $u \in \mathcal{S}(\mathbb{R})$ then*

$$u(x) = u(0)\exp(-\frac{x^2}{2}) + xv(x),\ \ v \in \mathcal{S}(\mathbb{R}). \tag{4.162}$$

PROOF. Here I will leave it to you (look in the problems) to show that the Gaussian

$$\exp(-\frac{x^2}{2}) \in \mathcal{S}(\mathbb{R}). \tag{4.163}$$

Observe then that the difference

$$w(x) = u(x) - u(0)\exp(-\frac{x^2}{2}) \in \mathcal{S}(\mathbb{R}) \text{ and } w(0) = 0.$$

This is clearly a necessary condition to see that $w = xv$ with $v \in \mathcal{S}(\mathbb{R})$ and we can then see from the Fundamental Theorem of Calculus that

$$w(x) = \int_0^x w'(y)dy = x\int_0^1 w'(tx)dt \Longrightarrow v(x) = \int_0^1 w'(tx) = \frac{w(x)}{x}. \tag{4.164}$$

From the first formula for v it follows that it is infinitely differentiable and from the second formula the derivatives decay rapidly since each derivative can be written in the form of a finite sum of terms $p(x)\frac{d^l w}{dx^l}/x^N$ where the p's are polynomials. The rapid decay of the derivatives of w therefore implies the rapid decay of the derivatives of v. So indeed we have proved Lemma 4.8. □

Let me set $\gamma(x) = \exp(-\frac{x^2}{2})$ to simplify the notation. Taking the Fourier transform of each of the terms in (4.162) gives

$$\hat{u} = u(0)\hat{\gamma} + i\frac{d\hat{v}}{d\xi}. \tag{4.165}$$

Since $\hat{v} \in \mathcal{S}(\mathbb{R})$,

$$\int \frac{d\hat{v}}{d\xi} = \lim_{R\to\infty} \int_{-R}^{R} \frac{d\hat{v}}{d\xi} = \lim_{R\to\infty} \left[\hat{v}(\xi)\right]_{-R}^{R} = 0. \tag{4.166}$$

So now we see that

$$\int \hat{u} = cu(0), \ c = \int \hat{\gamma}$$

being a constant that we still need to work out!

LEMMA 4.9. *For the Gaussian,* $\gamma(x) = \exp(-\frac{x^2}{2})$,

$$\hat{\gamma}(\xi) = \sqrt{2\pi}\gamma(\xi). \tag{4.167}$$

PROOF. Certainly, $\hat{\gamma} \in \mathcal{S}(\mathbb{R})$ and from the identities for derivatives above

$$\frac{d\hat{\gamma}}{d\xi} = -i\mathcal{F}(x\gamma), \ \xi\hat{\gamma} = \mathcal{F}(-i\frac{d\gamma}{dx}). \tag{4.168}$$

Thus, $\hat{\gamma}$ satisfies the same differential equation as γ :

$$\frac{d\hat{\gamma}}{d\xi} + \xi\hat{\gamma} = -i\mathcal{F}(\frac{d\gamma}{dx} + x\gamma) = 0.$$

This equation we can solve and so we conclude that $\hat{\gamma} = c'\gamma$ where c' is also a constant that we need to compute. To do this observe that

$$c' = \hat{\gamma}(0) = \int \gamma = \sqrt{2\pi} \tag{4.169}$$

which gives (4.167). The computation of the integral in (4.169) is a standard clever argument which you probably know. Namely take the square and work in polar coordinates in two variables:

$$\begin{aligned}(\int \gamma)^2 &= \int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dxdy \\ &= \int_0^{2\pi} \int_0^\infty e^{-r^2/2} r dr d\theta = 2\pi\big[-e^{-r^2/2}\big]_0^\infty = 2\pi.\end{aligned} \tag{4.170}$$

□

So, finally we need to get from (4.161) to the inversion formula. Changing variable in the Fourier transform we can see that for any $y \in \mathbb{R}$, setting $u_y(x) = u(x+y)$, which is in $\mathcal{S}(\mathbb{R})$ if $u \in \mathcal{S}(\mathbb{R})$,

$$\mathcal{F}(u_y) = \int e^{-ix\xi} u_y(x) dx = \int e^{-i(s-y)\xi} u(s) ds = e^{iy\xi}\hat{u}. \tag{4.171}$$

Now, plugging u_y into (4.161) we see that

$$\int \hat{u}_y(0) = 2\pi u_y(0) = 2\pi u(y) = \int e^{iy\xi}\hat{u}(\xi) d\xi \Longrightarrow u(y) = \mathcal{G}u, \tag{4.172}$$

the Fourier inversion formula. So we have proved the Theorem. □

8. Convolution

There is a discussion of convolution later in the notes, I have inserted a new (but not very different) treatment here to cover the density of $\mathcal{S}(\mathbb{R})$ in $L^2(\mathbb{R})$ needed in the next section.

Consider two continuous functions of compact support $u, v \in \mathcal{C}_c(\mathbb{R})$. Their convolution is

$$u * v(x) = \int u(x-y)v(y)dy = \int u(y)v(x-y)dy. \tag{4.173}$$

The first integral is the definition, clearly it is a well-defined Riemann integral since the integrand is continuous as a function of y and vanishes whenever $v(y)$ vanishes – so has compact support. In fact if both u and v vanish outside $[-R, R]$ then $u * v = 0$ outside $[-2R, 2R]$.

From standard properties of the Riemann integral (or Dominated convergence if you prefer!) it follows easily that $u*v$ is continuous. What we need to understand is what happens if (at least) one of u or v is smoother. In fact we will want to take a very smooth function, so I pause here to point out

LEMMA 4.10. *There exists a ('bump') function $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ which is infinitely differentiable, i.e. has continuous derivatives of all orders, vanishes outside $[-1, 1]$, is strictly positive on $(-1, 1)$ and has integral 1.*

PROOF. We start with an explicit function,

$$\phi(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \le 0. \end{cases} \tag{4.174}$$

The exponential function grows faster than any polynomial at $+\infty$, since

$$\exp(x) > \frac{x^k}{k!} \text{ in } x > 0 \ \forall\ k. \tag{4.175}$$

This can be seen directly from the Taylor series which converges on the whole line (indeed in the whole complex plane)

$$\exp(x) = \sum_{k \ge 0} \frac{x^k}{k!}.$$

From (4.175) we deduce that

$$\lim_{x \downarrow 0} \frac{e^{-1/x}}{x^k} = \lim_{R \to \infty} \frac{R^k}{e^R} = 0 \ \forall\ k \tag{4.176}$$

where we substitute $R = 1/x$ and use the properties of $\exp$. In particular ϕ in (4.174) is continuous across the origin, and so everywhere. We can compute the derivatives in $x > 0$ and these are of the form

$$\frac{d^l \phi}{dx^l} = \frac{p_l(x)}{x^{2l}} e^{-1/x},\ x > 0,\ p_l \text{ a polynomial.} \tag{4.177}$$

As usual, do this by induction since it is true for $l = 0$ and differetiating the formula for a given l one finds

$$\frac{d^{l+1}\phi}{dx^{l+1}} = \left(\frac{p_l(x)}{x^{2l+2}} - 2l \frac{p_l(x)}{x^{2l+1}} + \frac{p_l'(x)}{x^{2l}} \right) e^{-1/x} \tag{4.178}$$

where the coefficient function is of the desired form p_{l+1}/x^{2l+2}.

Once we know (4.177) then we see from (4.176) that all these functions are continuous down to 0 where they vanish. From this it follows that ϕ in (4.174) is infinitely differentiable. For ϕ itself we can use the Fundamental Theorem of Calculus to write

$$\phi(x) = \int_\epsilon^x U(t)dt + \phi(\epsilon),\ x > \epsilon > 0. \tag{4.179}$$

Here U is the derivative in $x > 0$. Taking the limit as $\epsilon \downarrow 0$ both sides converge, and then we see that

$$\phi(x) = \int_0^x U(t)dt.$$

From this it follows that ϕ is continuously differentiable across 0 and it derivative is U, the continuous extension of the derivative from $x > 0$. The same argument applies to succssive derivatives, so indeed ϕ is infinitely differentiable.

From ϕ we can construct a function closer to the desired bump function. Namely

$$\Phi(x) = \phi(x+1)\phi(1-x).$$

The first factor vanishes when $x \le -1$ and is otherwise positive while the second vanishes when $x \ge 1$ but is otherwise positive, so the product is infinitely differentiable on $\mathbb{R}$ and positive on $(-1, 1)$ but otherwise 0. Then we can normalize the integral to 1 by taking

$$\psi(x) = \Phi(x) / \int \Phi. \tag{4.180}$$

□

In particular from Lemma 4.10 we conclude that the space $\mathcal{C}_c^\infty(\mathbb{R})$, of infinitely differentiable functions of compact support, is not empty. Going back to convolution in (4.173) suppose now that v is smooth. Then

$$u \in \mathcal{C}_{\mathrm{c}}(\mathbb{R}),\ v \in \mathcal{C}_c^\infty(\mathbb{R}) \Longrightarrow u * v \in \mathcal{C}_c^\infty(\mathbb{R}). \tag{4.181}$$

As usual this follows from properties of the Riemann integral or by looking directly at the difference quotient

$$\frac{u * v(x+t) - u * v(x)}{t} = \int u(y)\frac{v(x+t-y) - v(x-y)}{t}dt.$$

As $t \to 0$, the difference quotient for v converges uniformly (in y) to the derivative and hence the integral converges and the derivative of the convolution exists,

$$\frac{d}{dx} u * v(x) = u * (\frac{dv}{dx}). \tag{4.182}$$

This result allows immediate iteration, showing that the convolution is smooth and we know that it has compact support

Proposition 4.7. *For any $u \in \mathcal{C}_c(\mathbb{R})$ there exists $u_n \to u$ uniformly on $\mathbb{R}$ where $u_n \in \mathcal{C}_c^\infty(\mathbb{R})$ with supports in a fixed compact set.*

Proof. For each $\epsilon > 0$ consider the rescaled bump function

$$\psi_\epsilon = \epsilon^{-1}\psi(\frac{x}{\epsilon}) \in \mathcal{C}_c^\infty(\mathbb{R}). \tag{4.183}$$

In fact, ψ_ϵ vanishes outside the interval (ϵ, ϵ), is positive within this interval and has integral 1 – which is what the factor of ϵ^{-1} does. Now set

$$u_\epsilon = u * \psi_\epsilon \in \mathcal{C}_c^\infty(\mathbb{R}),\ \epsilon > 0, \tag{4.184}$$

from what we have just seen. From the supports of these functions, u_ϵ vanishes outside $[-R-\epsilon, R+\epsilon]$ if u vanishes outside $[-R, R]$. So only the convergence remains. To get this we use the fact that the integral of ψ_ϵ is equal to 1 to write

$$u_\epsilon(x) - u(x) = \int (u(x-y)\psi_\epsilon(y) - u(x)\psi_\epsilon(y))dy. \tag{4.185}$$

Estimating the integral using the positivity of the bump function

$$|u_\epsilon(x) - u(x)| = \int_{-\epsilon}^{\epsilon} |u(x-y) - u(x)|\psi_\epsilon(y)dy. \tag{4.186}$$

By the uniformity of a continuous function on a compact set, given $\delta > 0$ there exists $\epsilon > 0$ such that

$$\sup_{[-\epsilon,\epsilon]} |u(x-y) - y(x)| < \delta \ \forall \ x \in \mathbb{R}.$$

So the uniform convergence follows:-

$$\sup |u_\epsilon(x) - u(x)| \le \delta \int \phi_\epsilon = \delta \tag{4.187}$$

Pass to a sequence $\epsilon_n \to 0$ if you wish, □

Corollary 4.1. *The spaces $\mathcal{C}_c^\infty(\mathbb{R})$ and $\mathcal{S}(\mathbb{R})$ are dense in $L^2(\mathbb{R})$.*

Uniform convegence of continuous functions with support in a fixed subset is stronger than L^2 convergence, so the result follows from the Proposition above for $\mathcal{C}_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$.

9. Plancherel and Parseval

But which is which?

We proceed to show that $\mathcal{F}$ and $\mathcal{G}$, defined in (4.137) and (4.145), both extend to isomorphisms of $L^2(\mathbb{R})$ which are inverses of each other. The main step is to show that

$$\int u(x)\hat{v}(x)dx = \int \hat{u}(\xi)v(\xi)d\xi,\ u,\ v \in \mathcal{S}(\mathbb{R}). \tag{4.188}$$

Since the integrals are rapidly convergent at infinity we may substitute the definite of the Fourier transform into (4.188), write the result out as a double integral and change the order of integration

$$\begin{aligned} \int u(x)\hat{v}(x)dx &= \int u(x) \int e^{-ix\xi} v(\xi)d\xi dx \\ &= \int v(\xi) \int e^{-ix\xi} u(x)dxd\xi = \int \hat{u}(\xi)v(\xi)d\xi. \end{aligned} \tag{4.189}$$

Now, if $w \in \mathcal{S}(\mathbb{R})$ we may replace $v(\xi)$ by $\overline{\hat{w}}(\xi)$, since it is another element of $\mathcal{S}(\mathbb{R})$. By the Fourier Inversion formula,

$$w(x) = (2\pi)^{-1} \int e^{-ix\xi}\hat{w}(\xi) \Longrightarrow \overline{w(x)} = (2\pi)^{-1} \int e^{ix\xi}\overline{\hat{w}(\xi)} = (2\pi)^{-1}\hat{v}. \tag{4.190}$$

Substituting these into (4.188) gives Parseval's formula

$$\int u\overline{w} = \frac{1}{2\pi} \int \hat{u}\overline{\hat{w}},\ u,\ w \in \mathcal{S}(\mathbb{R}). \tag{4.191}$$

Proposition 4.8. *The Fourier transform $\mathcal{F}$ extends from $\mathcal{S}(\mathbb{R})$ to an isomorphism on $L^2(\mathbb{R})$ with $\frac{1}{\sqrt{2\pi}}\mathcal{F}$ an isometric isomorphism with adjoint, and inverse, $\sqrt{2\pi}\mathcal{G}$.*

PROOF. Setting $u = w$ in (4.191) shows that

$$\|\mathcal{F}(u)\|_{L^2} = \sqrt{2\pi}\|u\|_{L^2} \tag{4.192}$$

for all $u \in \mathcal{S}(\mathbb{R})$. The density of $\mathcal{S}(\mathbb{R})$, established above, then implies that $\mathcal{F}$ extends by continuity to the whole of $L^2(\mathbb{R})$ as indicated. $\square$

This isomorphism of $L^2(\mathbb{R})$ has many implications. For instance, we would like to define the Sobolev space $H^1(\mathbb{R})$ by the conditions that $u \in L^2(\mathbb{R})$ and $\frac{du}{dx} \in L^2(\mathbb{R})$ but to do this we would need to make sense of the derivative. However, we can 'guess' that if it exists, the Fourier transform of du/dx should be $i\xi\hat{u}(\xi)$. For a function in L^2, such as $\hat{u}$ given that $u \in L^2$, we *do* know what it means to require $\xi\hat{u}(\xi) \in L^2(\mathbb{R})$. We can then define the Sobolev spaces of any positive, even non-integral, order by

$$H^r(\mathbb{R}) = \{u \in L^2(\mathbb{R}); |\xi|^r\hat{u} \in L^2(\mathbb{R})\}. \tag{4.193}$$

Of course it would take us some time to investigate the properties of these spaces!

10. Weak and strong derivatives

In approaching the issue of the completeness of the eigenbasis for harmonic oscillator more directly, rather than by the kernel method discussed above, we run into the issue of weak and strong solutions of differential equations. Suppose that $u \in L^2(\mathbb{R})$, what does it *mean* to say that $\frac{du}{dx} \in L^2(\mathbb{R})$. For instance, we will want to understand what the 'possible solutions of'

$$\operatorname{An} u = f,\ u,\ f \in L^2(\mathbb{R}),\ \operatorname{An} = \frac{d}{dx} + x \tag{4.194}$$

are. Of course, if we assume that u is continuously differentiable then we know what this means, but we need to consider the possibilities of giving a meaning to (4.194) under more general conditions – without assuming too much regularity on u (or any at all).

Notice that there is a difference between the two terms in $\operatorname{An} u = \frac{du}{dx} + xu$. If $u \in L^2(\mathbb{R})$ we can assign a meaning to the second term, xu, since we know that $xu \in L^2_{\text{loc}}(\mathbb{R})$. This is not a normed space, but it is a perfectly good vector space, in which $L^2(\mathbb{R})$ 'sits' – if you want to be pedantic it naturally injects into it. The point however, is that we *do know* what the statement $xu \in L^2(\mathbb{R})$ means, given that $u \in L^2(\mathbb{R})$, it means that there exists $v \in L^2(\mathbb{R})$ so that $xu = v$ in $L^2_{\text{loc}}(\mathbb{R})$ (or $L^2_{\text{loc}}(\mathbb{R})$). The derivative can actually be handled in a similar fashion using the Fourier transform but I will not do that here.

Rather, consider the following three 'L^2-based notions' of derivative.

DEFINITION 4.5. (1) We say that $u \in L^2(\mathbb{R})$ has a *Sobolev derivative* if there exists a sequence $\phi_n \in \mathcal{C}^1_{\text{c}}(\mathbb{R})$ such that $\phi_n \to u$ in $L^2(\mathbb{R})$ and $\phi_n' \to v$ in $L^2(\mathbb{R})$, $\phi_n' = \frac{d\phi_n}{dx}$ in the usual sense of course.

(2) We say that $u \in L^2(\mathbb{R})$ has a *strong derivative* (in the L^2 sense) if the limit

$$\lim_{0 \neq s \to 0} \frac{u(x+s) - u(x)}{s} = \tilde{v} \text{ exists in } L^2(\mathbb{R}). \tag{4.195}$$

(3) Thirdly, we say that $u \in L^2(\mathbb{R})$ has a *weak derivative* in L^2 if there exists $w \in L^2(\mathbb{R})$ such that

$$(u, -\frac{df}{dx})_{L^2} = (w, f)_{L^2} \ \forall \ f \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R}). \tag{4.196}$$

In all cases, we will see that it is justified to write $v = \tilde{v} = w = \frac{du}{dx}$ because these defintions turn out to be equivalent. Of course *if* $u \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$ then u is differentiable in each sense and the derivative is always du/dx – note that the integration by parts used to prove (4.196) is justified in that case. In fact we are most interested in the first and third of these definitions, the first two are both called 'strong derivatives.'

It is easy to see that the existence of a Sobolev derivative implies that this is also a weak derivative. Indeed, since ϕ_n, the approximating sequence whose existence is the definition of the Sobolev derivative, is in $\mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$ the integration by parts implicit in (4.196) is valid and so for all $f \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$,

$$(\phi_n, -\frac{df}{dx})_{L^2} = (\phi'_n, f)_{L^2}. \tag{4.197}$$

Since $\phi_n \to u$ in L^2 and $\phi'_n \to v$ in L^2 both sides of (4.197) converge to give the identity (4.196).

Before proceeding to the rest of the equivalence of these definitions we need to do some preparation. First let us investigate a little the consequence of the existence of a Sobolev derivative.

LEMMA 4.11. *If $u \in L^2(\mathbb{R})$ has a Sobolev derivative then $u \in \mathcal{C}(\mathbb{R})$ and there exists a uniquely defined element $w \in L^2(\mathbb{R})$ such that*

$$u(x) - u(y) = \int_y^x w(s)ds \ \forall \ y \le x \in \mathbb{R}. \tag{4.198}$$

PROOF. Suppose u has a Sobolev derivative, determined by some approximating sequence ϕ_n. Consider a general element $\psi \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$. Then $\tilde{\phi}_n = \psi\phi_n$ is a sequence in $\mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$ and $\tilde{\phi}_n \to \psi u$ in L^2. Moreover, by the product rule for standard derivatives

$$\frac{d}{dx}\tilde{\phi}_n = \psi'\phi_n + \psi\phi'_n \to \psi' u + \psi w \text{ in } L^2(\mathbb{R}). \tag{4.199}$$

Thus in fact ψu also has a Sobolev derivative, namely $\phi' u + \psi w$ if w is the Sobolev derivative for u given by ϕ_n – which is to say that the product rule for derivatives holds under these conditions.

Now, the formula (4.198) comes from the Fundamental Theorem of Calculus which in this case really does apply to $\tilde{\phi}_n$ and shows that

$$\psi(x)\phi_n(x) - \psi(y)\phi_n(y) = \int_y^x \frac{d\tilde{\phi}_n}{ds}(s)ds. \tag{4.200}$$

For any given $x = \bar{x}$ we can choose ψ so that $\psi(\bar{x}) = 1$ and then we can take y below the lower limit of the support of ψ so $\psi(y) = 0$. It follows that for this choice of ψ,

$$\phi_n(\bar{x}) = \int_y^{\bar{x}} (\psi'\phi_n(s) + \psi\phi'_n(s))ds. \tag{4.201}$$

Now, we can pass to the limit as $n \to \infty$ and the left side converges for each fixed $\bar{x}$ (with ψ fixed) since the integrand converges in L^2 and hence in L^1 on this compact interval. This actually shows that the limit $\phi_n(\bar{x})$ must exist for each fixed $\bar{x}$. In fact we can always choose ψ to be constant near a particular point and apply this argument to see that

$$\phi_n(x) \to u(x) \text{ locally uniformly on } \mathbb{R}. \tag{4.202}$$

That is, the limit exists locally uniformly, hence represents a continuous function but that continuous function must be equal to the original u almost everywhere (since $\psi\phi_n \to \psi u$ in L^2).

Thus in fact we conclude that '$u \in \mathcal{C}(\mathbb{R})$' (which really means that u *has a representative* which is continuous). Not only that but we get (4.198) from passing to the limit on both sides of

$$u(x) - u(y) = \lim_{n\to\infty}(\phi_n(x) - \phi_n(y)) = \lim_{n\to\infty}\int_y^s(\phi'(s))ds = \int_y^s w(s)ds. \tag{4.203}$$

□

One immediate consequence of this is

$$\text{The Sobolev derivative is unique if it exists.} \tag{4.204}$$

Indeed, if w_1 and w_2 are both Sobolev derivatives then (4.198) holds for both of them, which means that $w_2 - w_1$ has vanishing integral on any finite interval and we know that this implies that $w_2 = w_1$ a.e.

So at least for Sobolev derivatives we are now justified in writing

$$w = \frac{du}{dx} \tag{4.205}$$

since w is unique and behaves like a derivative in the integral sense that (4.198) holds.

Lemma 4.12. *If u has a Sobolev derivative then u has a strong derivative and if u has a strong derivative then this is also a weak derivative.*

Proof. If u has a Sobolev derivative then (3.17) holds. We can use this to write the difference quotient as

$$\frac{u(x+s) - u(x)}{s} - w(x) = \frac{1}{s}\int_0^s (w(x+t) - w(x))dt \tag{4.206}$$

since the integral in the second term can be carried out. Using this formula twice the square of the L^2 norm, which is finite, is

$$\begin{aligned}\|\frac{u(x+s) - u(x)}{s} - w(x)\|^2_{L^2} \\ = \frac{1}{s^2}\int\int_0^s\int_0^s (w(x+t) - w(x)\overline{(w(x+t') - w(x))}dtdt'dx.\end{aligned} \tag{4.207}$$

There is a small issue of manupulating the integrals, but we can always 'back off a little' and replace u by the approximating sequence ϕ_n and then everything is fine – and we only have to check what happens at the end. Now, we can apply the Cauchy-Schwarz inequality as a triple integral. The two factors turn out to be the same so we find that

$$\|\frac{u(x+s)-u(x)}{s} - w(x)\|^2_{L^2} \le \frac{1}{s^2}\int\int_0^s\int_0^s |w(x+t)-w(x)|^2 dt dt' dx \\ = \frac{1}{s}\int_0^s \int |w(x+t)-w(x)|^2 dx dt \tag{4.208}$$

since the integrand does not depend on t'.

Now, something we checked long ago was that L^2 functions are 'continuous in the mean' in the sense that

$$\lim_{0\ne t\to 0}\int |w(x+t)-w(x)|^2 dx = 0. \tag{4.209}$$

Applying this to (4.208) and then estimating the t integral shows that

$$\frac{u(x+s)-u(x)}{s} - w(x) \to 0 \text{ in } L^2(\mathbb{R}) \text{ as } s \to 0. \tag{4.210}$$

By definition this means that u has w as a strong derivative. I leave it up to you to make sure that the manipulation of integrals is okay.

So, now suppose that u has a strong derivative, $\tilde{v}$. Observe that if $f \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$ then the limit defining the derivative

$$\lim_{0\ne s\to 0}\frac{f(x+s)-f(x)}{s} = f'(x) \tag{4.211}$$

is *uniform*. In fact this follows by writing down the Fundamental Theorem of Calculus, as in (4.198), again using the properties of Riemann integrals. Now, consider

$$(u(x), \frac{f(x+s)-f(x)}{s})_{L^2} = \frac{1}{s}\int u(x)\overline{f(x+s)}dx - \frac{1}{s}\int u(x)\overline{f(x)}dx \\ = (\frac{u(x-s)-u(x)}{s}, f(x))_{L^2} \tag{4.212}$$

where we just need to change the variable of integration in the first integral from x to $x+s$. However, letting $s\to 0$ the left side converges because of the uniform convergence of the difference quotient and the right side converges because of the assumed strong differentiability and as a result (noting that the parameter on the right is really $-s$)

$$(u, \frac{df}{dx})_{L^2} = -(w,f)_{L^2} \ \forall\ f \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R}) \tag{4.213}$$

which is weak differentiability with derivative $\tilde{v}$. □

So, at this point we know that Sobolev differentiabilty implies strong differentiability and either of the stong ones implies the weak. So it remains only to show that weak differentiability implies Sobolev differentiability and we can forget about the difference!

Before doing that, note again that a weak derivative, if it exists, is unique – since the difference of two would have to pair to zero in L^2 with all of $\mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$ which is dense. Similarly, if u has a weak derivative then so does ψu for any $\psi \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$

since we can just move ψ around in the integrals and see that

$$
\begin{aligned}
(\psi u, -\frac{df}{dx}) &= (u, -\overline{\psi}\frac{df}{dx}) \\
&= (u, -\frac{d\overline{\psi} f}{dx}) + (u, \overline{\psi'} f) \\
&= (w, \overline{\psi} f + (\psi' u, f) = (\psi w + \psi' u, f)
\end{aligned}
\tag{4.214}
$$

which also proves that the product formula holds for weak derivatives.

So, let us consider $u \in L^2_c(\mathbb{R})$ which does have a weak derivative. To show that it has a Sobolev derivative we need to construct a sequence ϕ_n. We will do this by convolution.

LEMMA 4.13. *If $\mu \in \mathcal{C}_c(\mathbb{R})$ then for any $u \in L^2_c(\mathbb{R})$,*

$$
\mu * u(x) = \int \mu(x-s)u(s)ds \in \mathcal{C}_c(\mathbb{R})
\tag{4.215}
$$

and if $\mu \in \mathcal{C}^1_c(\mathbb{R})$ then

$$
\mu * u(x) \in \mathcal{C}^1_c(\mathbb{R}), \ \frac{d\mu * u}{dx} = \mu' * u(x).
\tag{4.216}
$$

It folows that if μ has more continuous derivatives, then so does $\mu * u$.

PROOF. Since u has compact support and is in L^2 it in L^1 so the integral in (4.215) exists for each $x \in \mathbb{R}$ and also vanishes if $|x|$ is large enough, since the integrand vanishes when the supports become separate – for some R, $\mu(x-s)$ is supported in $|s-x| \le R$ and $u(s)$ in $|s| < R$ which are disjoint for $|x| > 2R$. It is also clear that $\mu * u$ is continuous using the estimate (from uniform continuity of μ)

$$
|\mu * u(x') - \mu * u(x)| \le \sup |\mu(x-s) - \mu(x'-s)| \|u\|_{L^1}.
\tag{4.217}
$$

Similarly the difference quotient can be written

$$
\frac{\mu * u(x') - \mu * u(x)}{t} = \int \frac{\mu(x'-s) - \mu(x-s)}{s} u(s)ds
\tag{4.218}
$$

and the uniform convergence of the difference quotient shows that

$$
\frac{d\mu * u}{dx} = \mu' * u.
\tag{4.219}
$$

□

One of the key properties of thes convolution integrals is that we can examine what happens when we 'concentrate' μ. Replace the one μ by the family

$$
\mu_\epsilon(x) = \epsilon^{-1}\mu(\frac{x}{\epsilon}), \ \epsilon > 0.
\tag{4.220}
$$

The singular factor here is introduced so that $\int \mu_\epsilon$ is independent of $\epsilon > 0$,

$$
\int \mu_\epsilon = \int \mu \ \forall \ \epsilon > 0.
\tag{4.221}
$$

Note that since μ has compact support, the support of μ_ϵ is concentrated in $|x| \le \epsilon R$ for some fixed R.

LEMMA 4.14. *If $u \in L^2_c(\mathbb{R})$ and $0 \le \mu \in \mathcal{C}^1_c(\mathbb{R})$ then*

$$\lim_{0\neq\epsilon\to0} \mu_\epsilon * u = (\int \mu)u \text{ in } L^2(\mathbb{R}). \tag{4.222}$$

In fact there is no need to assume that u has compact support for this to work.

PROOF. First we can change the variable of integration in the definition of the convolution and write it intead as

$$\mu * u(x) = \int \mu(s)u(x-s)ds. \tag{4.223}$$

Now, the rest is similar to one of the arguments above. First write out the difference we want to examine as

$$\mu_\epsilon * u(x) - (\int \mu)(x) = \int_{|s|\le\epsilon R} \mu_\epsilon(s)(u(x-s)-u(x))ds. \tag{4.224}$$

Write out the square of the absolute value using the formula twice and we find that

$$\begin{aligned}\int |\mu_\epsilon * u(x) - (\int \mu)(x)|^2 dx \\ = \int\int_{|s|\le\epsilon R}\int_{|t|\le\epsilon R} \mu_\epsilon(s)\mu_\epsilon(t)(u(x-s)-u(x))\overline{(u(x-s)-u(x))}dsdtdx\end{aligned} \tag{4.225}$$

Now we can write the integrand as the product of two similar factors, one being

$$\mu_\epsilon(s)^{\frac12}\mu_\epsilon(t)^{\frac12}(u(x-s)-u(x)) \tag{4.226}$$

using the non-negativity of μ. Applying the Cauchy-Schwarz inequality to this we get two factors, which are again the same after relabelling variables, so

$$\int |\mu_\epsilon*u(x)-(\int \mu)(x)|^2dx \le \int\int_{|s|\le\epsilon R}\int_{|t|\le\epsilon R} \mu_\epsilon(s)\mu_\epsilon(t)|u(x-s)-u(x)|^2. \tag{4.227}$$

The integral in x can be carried out first, then using continuity-in-the mean bounded by $J(s) \to 0$ as $\epsilon \to 0$ since $|s| < \epsilon R$. This leaves

$$\begin{aligned}\int |\mu_\epsilon * u(x) - (\int \mu)u(x)|^2dx \\ \le \sup_{|s|\le\epsilon R} J(s)\int_{|s|\le\epsilon R}\int_{|t|\le\epsilon R}\mu_\epsilon(s)\mu_\epsilon(t) = (\int\psi)^2 Y \sup_{|s|\le\epsilon R} \to 0.\end{aligned} \tag{4.228}$$

□

After all this preliminary work we are in a position to to prove the remaining part of 'weak=strong'.

LEMMA 4.15. *If $u \in L^2(\mathbb{R})$ has w as a weak L^2-derivative then w is also the Sobolev derivative of u.*

PROOF. Let's assume first that u has compact support, so we can use the discussion above. Then set $\phi_n = \mu_{1/n} * u$ where $\mu \in \mathcal{C}^1_c(\mathbb{R})$ is chosen to be non-negative and have integral $\int \mu = 0$; μ_ϵ is defined in (4.220). Now from Lemma 4.14 it follows that $\phi_n \to u$ in $L^2(\mathbb{R})$. Also, from Lemma 4.13, $\phi_n \in \mathcal{C}^1_c(\mathbb{R})$ has derivative given by (4.216). This formula can be written as a pairing in L^2 :

$$(\mu_{1/n})' * u(x) = (u(s), -\frac{d\mu_{1/n}(x-s)}{ds})^2_L = (w(s), \frac{d\mu_{1/n}(x-s)}{ds})_{L^2} \tag{4.229}$$

using the definition of the weak derivative of u. It therefore follows from Lemma 4.14 applied again that

$$\phi'_n = \mu_{/m1/n} * w \to w \text{ in } L^2(\mathbb{R}). \tag{4.230}$$

Thus indeed, ϕ_n is an approximating sequence showing that w is the Sobolev derivative of u.

In the general case that $u \in L^2(\mathbb{R})$ has a weak derivative but is not necessarily compactly supported, consider a function $\gamma \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R})$ with $\gamma(0) = 1$ and consider the sequence $v_m = \gamma(x)u(x)$ in $L^2(\mathbb{R})$ each element of which has compact support. Moreover, $\gamma(x/m) \to 1$ for each x so by Lebesgue dominated convergence, $v_m \to u$ in $L^2(\mathbb{R})$ as $m \to \infty$. As shown above, v_m has as weak derivative

$$\frac{d\gamma(x/m)}{dx}u + \gamma(x/m)w = \frac{1}{m}\gamma'(x/m)u + \gamma(x/m)w \to w$$

as $m \to \infty$ by the same argument applied to the second term and the fact that the first converges to 0 in $L^2(\mathbb{R})$. Now, use the approximating sequence $\mu_{1/n} * v_m$ discussed converges to v_m with its derivative converging to the weak derivative of v_m. Taking $n = N(m)$ sufficiently large for each m ensures that $\phi_m = \mu_{1/N(m)} * v_m$ converges to u and its sequence of derivatives converges to w in L^2. Thus the weak derivative is again a Sobolev derivative. □

Finally then we see that the three definitions are equivalent and we will freely denote the Sobolev/strong/weak derivative as du/dx or u'.

11. Fourier transform and L^2

Recall that one reason for proving the completeness of the Hermite basis was to apply it to prove some of the important facts about the Fourier transform, which we already know is a linear operator

$$L^1(\mathbb{R}) \longrightarrow \mathcal{C}_\infty(\mathbb{R}),\ \hat{u}(\xi) = \int e^{ix\xi}u(x)dx. \tag{4.231}$$

Namely we have already shown the effect of the Fourier transform on the 'ground state':

$$\mathcal{F}(u_0)(\xi) = \sqrt{2\pi}e_0(\xi). \tag{4.232}$$

By a similar argument we can check that

$$\mathcal{F}(u_j)(\xi) = \sqrt{2\pi}i^j u_j(\xi)\ \forall\ j \in \mathbb{N}. \tag{4.233}$$

As usual we can proceed by induction using the fact that $u_j = \operatorname{Cr} u_{j-1}$. The integrals involved here are very rapidly convergent at infinity, so there is no problem with the integration by parts in

(4.234)

$$\begin{gathered}
\mathcal{F}(\frac{d}{dx}u_{j-1}) = \lim_{T\to\infty}\int_{-T}^{T} e^{-ix\xi}\frac{du_{j-1}}{dx}dx \\
= \lim_{T\to\infty}\left(\int_{-T}^{T}(i\xi)e^{-ix\xi}u_{j-1}dx + \left[e^{-ix\xi}u_{j-1}(x)\right]_{-T}^{T}\right) = (i\xi)\mathcal{F}(u_{j-1}), \\
\mathcal{F}(xu_{j-1}) = i\int \frac{de^{-ix\xi}}{d\xi}u_{j-1}dx = i\frac{d}{d\xi}\mathcal{F}(u_{j-1}).
\end{gathered}$$

Taken together these identities imply the validity of the inductive step:

$$\mathcal{F}(u_j) = \mathcal{F}((-\frac{d}{dx} + x)u_{j-1}) = (i(-\frac{d}{d\xi} + \xi)\mathcal{F}(u_{j-1}) = i\,\mathrm{Cr}(\sqrt{2\pi}i^{j-1}u_{j-1}) \tag{4.235}$$

so proving (4.233).

So, we have found an orthonormal basis for $L^2(\mathbb{R})$ with elements which are all in $L^1(\mathbb{R})$ and which are also eigenfunctions for $\mathcal{F}$.

Theorem 4.7. *The Fourier transform maps $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ into $L^2(\mathbb{R})$ and extends by continuity to an isomorphism of $L^2(\mathbb{R})$ such that $\frac{1}{\sqrt{2\pi}}\mathcal{F}$ is unitary with the inverse of $\mathcal{F}$ the continuous extension from $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ of*

$$\mathcal{F}(f)(x) = \frac{1}{2\pi}\int e^{ix\xi}f(\xi). \tag{4.236}$$

Proof. This really is what we have already proved. The elements of the Hermite basis e_j are all in both $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ so if $u \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ its image under $\mathcal{F}$ is in $L^2(\mathbb{R})$ because we can compute the L^2 inner products and see that

$$(\mathcal{F}(u), e_j) = \int_{\mathbb{R}^2} e_j(\xi)e^{ix\xi}u(x)dxd\xi = \int \mathcal{F}(e_j)(x)u(x) = \sqrt{2\pi}i^j(u, e_j). \tag{4.237}$$

Now Bessel's inequality shows that $\mathcal{F}(u) \in L^2(\mathbb{R})$ (it is of course locally integrable since it is continuous).

Everything else now follows easily. $\square$

Notice in particular that we have also proved Parseval's and Plancherel's identities for the Fourier transform:-

$$\|\mathcal{F}(u)\|_{L^2} = \sqrt{2\pi}\|u\|_{L^2},\ (\mathcal{F}(u), \mathcal{F}(v)) = 2\pi(u, v),\ \forall\ u, v \in L^2(\mathbb{R}). \tag{4.238}$$

Now there are lots of applications of the Fourier transform which we do not have the time to get into. However, let me just indicate the definitions of Sobolev spaces and Schwartz space and how they are related to the Fourier transform.

First Sobolev spaces. We now see that $\mathcal{F}$ maps $L^2(\mathbb{R})$ isomorphically onto $L^2(\mathbb{R})$ and we can see from (4.234) for instance that it 'turns differentiations by x into multiplication by ξ'. Of course we do not know how to differentiate L^2 functions so we have some problems making sense of this. One way, the usual mathematicians trick, is to turn what we want into a definition.

Definition 4.6. The Sobolev spaces of order s, for any $s \in (0, \infty)$, are defined as subspaces of $L^2(\mathbb{R})$:

$$H^s(\mathbb{R}) = \{u \in L^2(\mathbb{R}); (1 + |\xi|^2)^s\hat{u} \in L^2(\mathbb{R})\}. \tag{4.239}$$

It is natural to identify $H^0(\mathbb{R}) = L^2(\mathbb{R})$.

These Sobolev spaces, for each positive order s, are Hilbert spaces with the inner product and norm

$$(u, v)_{H^s} = \int (1 + |\xi|^2)^s\hat{u}(\xi)\overline{\hat{v}(\xi)},\ \|u\|_s = \|(1 + |\xi|^2)^{\frac{s}{2}}\hat{u}\|_{L^2}. \tag{4.240}$$

That they are pre-Hilbert spaces is clear enough. Completeness is also easy, given that we know the completeness of $L^2(\mathbb{R})$. Namely, if u_n is Cauchy in $H^s(\mathbb{R})$ then it follows from the fact that

$$\|v\|_{L^2} \le C\|v\|_s\ \forall\ v \in H^s(\mathbb{R}) \tag{4.241}$$

that u_n is Cauchy in L^2 and also that $(1 + |\xi|^2)^{\frac{s}{2}}\hat{u}_n(\xi)$ is Cauchy in L^2. Both therefore converge to a limit u in L^2 and the continuity of the Fourier transform shows that $u \in H^s(\mathbb{R})$ and that $u_n \to u$ in H^s.

These spaces are examples of what is discussed above where we have a dense inclusion of one Hilbert space in another, $H^s(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$. In this case the inclusion in *not* compact but it does give rise to a bounded self-adjoint operator on $L^2(\mathbb{R})$, $E_s : L^2(\mathbb{R}) \longrightarrow H^s(\mathbb{R}) \subset L^2(\mathbb{R})$ such that

$$(u, v)_{L^2} = (E_s u, E_s v)_{H^s}. \tag{4.242}$$

It is reasonable to denote this as $E_s = (1 + |D_x|^2)^{-\frac{s}{2}}$ since

$$u \in L^2(\mathbb{R}^n) \Longrightarrow \widehat{E_s u}(\xi) = (1 + |\xi|^2)^{-\frac{s}{2}}\hat{u}(\xi). \tag{4.243}$$

It is a form of 'fractional integration' which turns any $u \in L^2(\mathbb{R})$ into $E_s u \in H^s(\mathbb{R})$.

Having defined these spaces, which get smaller as s increases it can be shown for instance that if $n \geq s$ is an integer then the set of n times continuously differentiable functions on $\mathbb{R}$ which vanish outside a compact set are dense in H^s. This allows us to justify, by continuity, the following statement:-

Proposition 4.9. *The bounded linear map*

$$\frac{d}{dx} : H^s(\mathbb{R}) \longrightarrow H^{s-1}(\mathbb{R}),\ s \geq 1,\ v(x) = \frac{du}{dx} \Longleftrightarrow \hat{v}(\xi) = i\xi\hat{u}(\xi) \tag{4.244}$$

is consistent with differentiation on n times continuously differentiable functions of compact support, for any integer $n \geq s$.

In fact one can even get a 'strong form' of differentiation. The condition that $u \in H^1(\mathbb{R})$, that $u \in L^2$ 'has one derivative in L^2' is actually equivalent, for $u \in L^2(\mathbb{R})$ to the existence of the limit

$$\lim_{t\to 0} \frac{u(x+t)u(x)}{t} = v, \text{ in } L^2(\mathbb{R}) \tag{4.245}$$

and then $\hat{v} = i\xi\hat{u}$. Another way of looking at this is

$$\begin{gathered} u \in H^1(\mathbb{R}) \Longrightarrow u : \mathbb{R} \longrightarrow \mathbb{C} \text{ is continuous and} \\ u(x) - u(y) = \int_y^x v(t)dt,\ v \in L^2. \end{gathered} \tag{4.246}$$

If such a $v \in L^2(\mathbb{R})$ exists then it is unique – since the difference of two such functions would have to have integral zero over any finite interval and we know (from one of the exercises) that this implies that the function vansishes a.e.

One of the more important results about Sobolev spaces – of which there are many – is the relationship between these 'L^2 derivatives' and 'true derivatives'.

Theorem 4.8 (Sobolev embedding). *If n is an integer and $s > n + \frac{1}{2}$ then*

$$H^s(\mathbb{R}) \subset \mathcal{C}^n_\infty(\mathbb{R}) \tag{4.247}$$

consists of n times continuosly differentiable functions with bounded derivatives to order n (which also vanish at infinity).

This is actually not so hard to prove, there are some hints in the exercises below.

These are not the only sort of spaces with 'more regularity' one can define and use. For instance one can try to treat x and ξ more symmetrically and define smaller spaces than the H^s above by setting

$$H^s_{\text{iso}}(\mathbb{R}) = \{u \in L^2(\mathbb{R}); (1+|\xi|^2)^{\frac{s}{2}}\hat{u} \in L^2(\mathbb{R}),\ (1+|x|^2)^{\frac{s}{2}}u \in L^2(\mathbb{R})\}. \tag{4.248}$$

The 'obvious' inner product with respect to which these 'isotropic' Sobolev spaces $H^s_{\text{iso}}(\mathbb{R})$ are indeed Hilbert spaces is

$$(u,v)_{s,\text{iso}} = \int_{\mathbb{R}} u\overline{v} + \int_{\mathbb{R}} |x|^{2s}u\overline{v} + \int_{\mathbb{R}} |\xi|^{2s}\hat{u}\overline{\hat{v}} \tag{4.249}$$

which makes them look rather symmetric between u and $\hat{u}$ and indeed

$$\mathcal{F} : H^s_{\text{iso}}(\mathbb{R}) \longrightarrow H^s_{\text{iso}}(\mathbb{R}) \text{ is an isomorphism } \forall\ s \ge 0. \tag{4.250}$$

At this point, by dint of a little, only moderately hard, work, it is possible to show that the harmonic oscillator extends by continuity to an isomorphism

$$H : H^{s+2}_{\text{iso}}(\mathbb{R}) \longrightarrow H^s_{\text{iso}}(\mathbb{R})\ \forall\ s \ge 2. \tag{4.251}$$

Finally in this general vein, I wanted to point out that Hilbert, and even Banach, spaces are not the end of the road! One very important space in relation to a direct treatment of the Fourier transform, is the Schwartz space. The definition is reasonably simple. Namely we denote Schwartz space by $\mathcal{S}(\mathbb{R})$ and say

$$\begin{gathered} u \in \mathcal{S}(\mathbb{R}) \Longleftrightarrow u : \mathbb{R} \longrightarrow \mathbb{C} \\ \text{is continuously differentiable of all orders and for every } n, \\ \|u\|_n = \sum_{k+p\le n} \sup_{x\in\mathbb{R}} (1+|x|)^k |\frac{d^p u}{dx^p}| < \infty. \end{gathered} \tag{4.252}$$

All these inequalities just mean that all the derivatives of u are 'rapidly decreasing at ∞' in the sense that they stay bounded when multiplied by any polynomial.

So in fact we know already that $\mathcal{S}(\mathbb{R})$ is not empty since the elements of the Hermite basis, $e_j \in \mathcal{S}(\mathbb{R})$ for all j. In fact it follows immediately from this that

$$\mathcal{S}(\mathbb{R}) \longrightarrow L^2(\mathbb{R}) \text{ is dense.} \tag{4.253}$$

If you want to try your hand at something a little challenging, see if you can check that

$$\mathcal{S}(\mathbb{R}) = \bigcap_{s>0} H^s_{\text{iso}}(\mathbb{R}) \tag{4.254}$$

which uses the Sobolev embedding theorem above.

As you can see from the definition in (4.252), $\mathcal{S}(\mathbb{R})$ is not likely to be a Banach space. Each of the $\|\cdot\|_n$ is a norm. However, $\mathcal{S}(\mathbb{R})$ is pretty clearly not going to be complete with respect to any one of these. However it is complete with respect to all, countably many, norms. What does this mean? In fact $\mathcal{S}(\mathbb{R})$ is a *metric space* with the metric

$$d(u,v) = \sum_n 2^{-n} \frac{\|u-v\|_n}{1+\|u-v\|_n} \tag{4.255}$$

as you can check. So the claim is that $\mathcal{S}(\mathbb{R})$ *is* complete as a metric space – such a thing is called a Fréchet space.

What has this got to do with the Fourier transform? The point is that
(4.256)
$$\mathcal{F}:\mathcal{S}(\mathbb{R})\longrightarrow\mathcal{S}(\mathbb{R}) \text{ is an isomorphism and } \mathcal{F}(\frac{du}{dx})=i\xi\mathcal{F}(u),\ \mathcal{F}(xu)=-i\frac{d\mathcal{F}(u)}{d\xi}$$

where this now makes sense. The dual space of $\mathcal{S}(\mathbb{R})$ – the space of continuous linear functionals on it, is the space, denoted $\mathcal{S}'(\mathbb{R})$, of tempered distributions on $\mathbb{R}$.

12. Schwartz distributions

We do not have time in this course to really discuss distributions. Still, it is a good idea for you to know what they are and why they are useful. Of course to really appreciate their utility you need to read a bit more than I have here. First think a little about the Schwartz space $\mathcal{S}(\mathbb{R})$ introduced above. The metric in (4.255) might seem rather mysterious but it has the important property that *each* of the norms $\|\cdot\|_n$ defines a continuous function $\mathcal{S}(\mathbb{R})\longrightarrow\mathbb{R}$ with respect to this metric topology. In fact a linear map
(4.257)
$$T:\mathcal{S}(\mathbb{R})\longrightarrow\mathbb{C} \text{ linear is continuous iff } \exists\ N,C \text{ s.t. } \|T\phi\|\le C\|\phi\|_N\ \forall\ \phi\in\mathcal{S}(\mathbb{R}).$$

So, the continuous linear functionals on $\mathcal{S}(\mathbb{R})$ are just those which are continous with respect to one of the norms.

These functionals are exactly the space of *tempered distributions*

$$\mathcal{S}'(\mathbb{R})=\{T:\mathcal{S}(\mathbb{R})\longrightarrow\mathbb{C} \text{ linear and continuous}\}. \tag{4.258}$$

The relationship to functions is that each $f\in L^2(\mathbb{R})$ (or more generally such that $(1+|x|)^{-N}\in L^1(\mathbb{R})$ for some N) defines an element of $\mathcal{S}'(\mathbb{R})$ by integration:

$$T_f:\mathcal{S}(\mathbb{R})\ni\phi\longmapsto\int f(x)\phi(x)\in\mathbb{C}\Longrightarrow T_f\in\mathcal{S}'(\mathbb{R}). \tag{4.259}$$

Indeed, this amounts to showing that $\|\phi\|_{L^2}$ is a continuous norm on $\mathcal{S}(\mathbb{R})$ (so it must be bounded by a multiple of one of the $\|\phi\|_N$, which one?)

It is relatively straightforward to show that $L^2(\mathbb{R})\ni f\longmapsto T_f\in\mathcal{S}'(\mathbb{R})$ is injective – nothing is 'lost'. So after a little more experience with distributions one comes to identify f and T_f. Notice that this is just an extension of the behaviour of $L^2(\mathbb{R})$ where (because we can drop the complex conjugate in the inner product) by Riesz' Theorem we can identify (linearly) $L^2(\mathbb{R})$ with it dual, exactly by the map $f\longmapsto T_f$.

Other elements of $\mathcal{S}'(\mathbb{R})$ include the delta 'function' at the origin and even its 'derivatives' for each j

$$\delta^j:\mathcal{S}(\mathbb{R})\ni\phi\longmapsto(-1)^j\frac{d^j\phi}{dx^j}(0)\in\mathbb{C}. \tag{4.260}$$

In fact one of the main points about the space $\mathcal{S}'(\mathbb{R})$ is that differentiation and multiplication by polynomials is well defined

$$\frac{d}{dx}:\mathcal{S}'(\mathbb{R})\longrightarrow\mathcal{S}'(\mathbb{R}),\ \times x:\mathcal{S}'(\mathbb{R})\longrightarrow\mathcal{S}'(\mathbb{R}) \tag{4.261}$$

in a way that is consistent with their actions under the identification $\mathcal{S}(\mathbb{R}):\phi\longmapsto T_\phi\in\mathcal{S}'(\mathbb{R})$. This property is enjoyed by other spaces of distributions but the fundamental fact that the Fourier transform extends to

$$\mathcal{F} : \mathcal{S}'(\mathbb{R}) \longrightarrow \mathcal{S}'(\mathbb{R}) \text{ as an isomorphism} \tag{4.262}$$

is more characteristic of $\mathcal{S}'(\mathbb{R})$.

13. Poisson summation formula

We have talked both about Fourier series and the Fourier transform. It is natural to ask: What is the connection between these? The Fourier series of a function in $L^2(0, 2\pi)$ we thought of as given by the Fourier-Bessel series with respect to the orthonormal basis

$$\frac{\exp(ikx)}{\sqrt{2\pi}},\ k \in \mathbb{Z}. \tag{4.263}$$

The interval here is just a particular choice – if the upper limit is changed to T then the corresponding orthonormal basis of $L^2(0, T)$ is

$$\frac{\exp(i2\pi kx/T)}{\sqrt{T}},\ k \in \mathbb{Z}. \tag{4.264}$$

Sometimes the Fourier transform is thought of as the limit of the Fourier series expansion when $T \to \infty$. This is actually not such a nice limit, so unless you *have* (or *want*) to do this I recommend against it!

A more fundamental relationship between the two comes about as follows. We can think of $L^2(0, 2\pi)$ as 'really' being the 2π-periodic functions restricted to this interval. Since the values at the end-points don't matter this does give a bijection – between 2π-periodic, locally square-integrable functions on the line and $L^2(0, 2\pi)$. On the other hand we can also think of the periodic functions as being defined on the circle, $|z| = 1$ in $\mathbb{C}$ or identified with the values of $\theta \in \mathbb{R}$ modulo repeats:

$$\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z} \ni \theta \longmapsto e^{i\theta} \in \mathbb{C}. \tag{4.265}$$

Let us denote by $\mathcal{C}^\infty(\mathbb{T})$ the space of infinitely differentiable, 2π-periodic functions on the line; this is also the space of smooth functions on the circle, thought of as a manifold.

How can one construct such functions. There are plenty of examples, for instance the $\exp(ikx)$. Another way to construct examples is to sum over translations:-

Lemma 4.16. *The map*

$$A : \mathcal{S}(\mathbb{R}) \ni f \longrightarrow \sum_{k\in\mathbb{Z}} f(\cdot - 2\pi k) \in \mathcal{C}^\infty(\mathbb{T}) \tag{4.266}$$

is surjective.

Proof. That the series in (4.266) converges uniformly on $[0, 2\pi]$ (or any bounded interval) is easy enought to see, since the rapid decay of elements of $\mathcal{S}(\mathbb{R})$ shows that

$$|f(x)| \le C(1+|x|)^{-2},\ x \in \mathbb{R} \Longrightarrow |f(x - 2\pi k)| \le C'(1+|k|)^{-2},\ x \in [0, 2\pi] \tag{4.267}$$

since if $k > 2$ $|x-2\pi k| \ge k$ if $x \in [0, 2\pi]$. Clearly (4.267) implies uniform convergence of the series. Since the derivatives of f are also in $\mathcal{S}(\mathbb{R})$ the series obtained by term-by-term differentiation also converges uniformly and by standard arguments the limit Ag is therefore infinitely differentiable, with

$$\frac{d^jAf}{dx^j} = A\frac{d^jf}{dx^j}. \tag{4.268}$$

This shows that the map A, clearly linear, is well-defined. Now, how to see that it is surjective? Let's first prove a special case. Indeed, look for a function $\psi \in \mathcal{C}_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$ which is non-negative and such that $A\psi = 1$. We know that we can find $\phi \in \mathcal{C}_c^\infty(\mathbb{R})$, $\phi \geq 0$ with $\phi > 0$ on $[0, 2\pi]$. Then consider $A\phi \in \mathcal{C}^\infty(\mathbb{T})$. It must be stricly positive, $A\phi \geq \epsilon > 0$ since it is larger that ϕ. So consider instead the function

$$\psi = \frac{\phi}{A\phi} \in \mathcal{C}_c^\infty(\mathbb{R}) \tag{4.269}$$

where we think of $A\phi$ as 2π-periodic on $\mathbb{R}$. In fact using this periodicity we see that

$$A\psi \equiv 1. \tag{4.270}$$

So this shows that the constant function 1 is in the range of A. In general, just take $g \in \mathcal{C}^\infty(\mathbb{T})$, thought of as 2π-periodic on the line, and it follows that

$$f = Bg = \psi g \in \mathcal{C}_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R}) \text{ satsifies } Af = g. \tag{4.271}$$

Indeed,

$$Ag = \sum_k \psi(x - 2\pi k) g(x - 2\pi k) = g(x) \sum_k \psi(x - 2\pi k) = g \tag{4.272}$$

using the periodicity of g. In fact B is a right inverse for A,

$$AB = \mathrm{Id} \text{ on } \mathcal{C}^\infty(\mathbb{T}). \tag{4.273}$$

□

Question 2. What is the null space of A?

Since $f \in \mathcal{S}(\mathbb{R})$ and $Af \in \mathcal{C}^\infty(\mathbb{T}) \subset L^2(0, 2\pi)$ with our identifications above, the question arises as to the relationship between the Fourier transform of f and the Fourier series of Af.

Proposition 4.10 (Poisson summation formula). *If $g = Af$, $g \in \mathcal{C}^\infty(\mathbb{T})$ and $f \in \mathcal{S}(\mathbb{R})$ then the Fourier coefficients of g are*

$$c_k = \int_{[0,2\pi]} g e^{-ikx} = \hat{f}(k). \tag{4.274}$$

Proof. Just substitute in the formula for g and, using uniform convergenc, check that the sum of the integrals gives after translation the Fourier transform of f. □

If we think of recovering g from its Fourier series,

$$g(x) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} c_k e^{ikx} = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \hat{f}(k) e^{ikx} \tag{4.275}$$

then in terms of the Fourier transform on $\mathcal{S}'(\mathbb{R})$ alluded to above, this takes the rather elegant form

$$\frac{1}{2\pi} \mathcal{F}\left(\sum_{k \in \mathbb{Z}} \delta(\cdot - k)\right)(x) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} e^{ikx} = \sum_{k \in \mathbb{Z}} \delta(x - 2\pi k). \tag{4.276}$$

The sums of translated Dirac deltas and oscillating exponentails all make sense in $\mathcal{S}'(\mathbb{R})$.