Class notes

1 Tychonoff's theorem

If $\{X_i\}_{i\in I}$ are topological spaces then the product space is defined as $X = \prod_{i\in I} X_i = \left\{ f : I \to \bigcup_{i\in I} X_i \;\middle|\; (\forall i)(f(i) \in X_i) \right\}$; the product topology is defined to be the coarsest topology (i.e. the topology with the fewest open sets) for which all the projections π_i ($\pi_i(x) = x_i$) are continuous. That is, the sets $\bigcap_{j=1}^{n} \pi_{i_j}^{-1}(N_{i_j})$ where N_i are open sets in X_i form a base for the topology on X. In this topology a net $\langle f_\alpha \rangle$ converges iff $\forall i \in I, f_\alpha(i)$ converges, that is, the topology is that of pointwise convergence of functions. The fact that X is nonempty for general nonempty X_i is equivalent to the axiom of choice, AC.

Theorem 1 (Tychonoff). [1] *Assume $\{X_i\}_{i\in I}$ are compact for all $i \in I$. Then $X = \prod_{i\in I} X_i$ is compact in the product topology.*

Proof. Let $\mathcal{F}$ be a family of closed sets in X with the finite intersection property (f.i.p.). We want to show $\bigcap_{F\in\mathcal{F}} F \neq \emptyset$. Clearly this is the case if the same holds for any larger family $\mathcal{F}'$. A subtle point in the proof is to take the largest such set. Note that any chain of families (not necessarily of closed sets) with the f.i.p. $\mathcal{F}_\alpha \subset \mathcal{F}_{\alpha'} \subset \cdots$ has an upper bound, with the f.i.p, namely their union. By Zorn's lemma, there is a maximal family with the f.i.p., $\mathcal{M} \supset \mathcal{F}$. In the following "construct", "choose" etc. are just ways of speaking, as we rely on the AC.

We now construct a point in X which should be in all $F \in \mathcal{F}$ (and, in fact, all $M \in \mathcal{M}$). For any i, $\{\overline{\pi_i(M)} | M \in \mathcal{M}\}$ is a family of closed sets in X_i with the f.i.p. Then, for each i there is an $m_i \in \bigcap_{M\in\mathcal{M}} \overline{\pi_i(M)}$. Choose an m_i for each i and let $m = (m_i)_{i\in I}$.

If we show that $\bigcap_{j=1}^{n} \pi_{i_j}^{-1}(O_{i_j})$ (O_i open nbd of m_i) intersect nontrivially each F, this will imply that $m \in F$ for all our F. This is because each F is closed and for each F it follows that any open nbd of m intersects nontrivially F, implying, by elementary topology, $m \in F$.

The property above is implied by the following: for any O_i as above, $\pi_i^{-1}(O_i) \in \mathcal{M}$.

[1] Adaptation of a Bourbaki proof, see also Loomis, see p.11

Now, for any $M \in \mathcal{M}$ we have, by construction, $\overline{\pi_i(M)} \cap O_i \neq \emptyset$. Thus $\overline{\pi_i(M) \cap O_i} \neq \emptyset$ implying $\pi_i(M) \cap O_i \neq \emptyset$ which in turn means $M \cap \pi_i^{-1}(O_i) \neq \emptyset$. Then, adjoining any single set $\pi_i^{-1}(O_i)$ to $\mathcal{M}$, the f.i.p. is preserved. But, then by the maximality of $\mathcal{M}$, $\pi_i^{-1}(O_i) \in \mathcal{M}$, and this holds for any i ending the proof . ☐

2 Further discussions

Assume $\{X_i\}_{i\in\mathbb{N}}$ are second countable. Check that the product space X is also second countable. (The proof, in general, relies on a weak version of the AC (to select a countable open base for each X_i), but does not require it if the X_i are linearly ordered which is the case if $X_i = \mathbb{R}$, or, after simple modifications, $X_i = \mathbb{R}^n$ or $\mathbb{C}$.)

Recall that a sequence is a function $f : \mathbb{N} \to X$ and a subsequence $f \circ g$ is defined by a $g : \mathbb{N} \to \mathbb{N}$ s.t. $\lim_{n\to\infty} g(n) = \infty$.

Without further use of the AC we can prove compactness of the product X of compact, second countable spaces.

Proof. Let $f : \mathbb{N} \to X$ be a sequence. Since X_1 is compact, there is a subsequence defined by an f_1 s.t. $(f \circ f_1)_1 : \mathbb{N} \to X_1$ is convergent. Inductively, there is a subsequence defined by an f_n s.t. all $(f \circ f_n)_i, i = 1, 2, ...n$ are convergent. Define g by $g(k) = f_k(k)$. Then, as you can easily check, $(f \circ g)_i, i \in \mathbb{N}$ are all convergent, implying that $(f \circ g) : \mathbb{N} \to X$ is convergent. ☐

2.1 Arzelá-Ascoli's theorem

Let X be a separable metric space with metric d; let E be a countable dense subset of X. Let $\{f_n\}_{n\in\mathbb{N}} : X \to \mathbb{C}$ be a sequence of equicontinuous and pointwise equibounded functions. A pointwise equibounded family $\mathcal{F}$ is one s.t. $\forall x \in X \sup_{F\in\mathcal{F}} |F(x)| = M(x) < \infty$ and equicontinuous means that $\forall\epsilon\exists\delta$ s.t. $\forall x, y \in X$ and $F \in \mathcal{F}$, $d(x, y) < \delta \Rightarrow |F(x) - F(y)| < \epsilon$. For the purpose of Arzelá-Ascoli's theorem below, equicontinuity can be replaced with the weaker condition $\forall e \in E$ there is an r s.t. for all y with $d(y, e) < r$ and all F we have $|F(e) - F(y)| < \epsilon$, which can be seen using the compact cover formulation of compactness.

Theorem 2. *Every sequence $\{F_n\}_{n\in\mathbb{N}} : X \to \mathbb{C}$ of equicontinuous and pointwise equibounded function has a subsequence which converges uniformly on compact sets.*

Proof. Let K be a compact set in X. Let $M(x)$ be as above. Then the space $Y = \prod_{e\in E}\{z \in \mathbb{C} : |z| \leqslant M(e)\}$ is compact. (By the remarks at the beginning of this section, the AC is not needed in this setting.) This means there is a subsequence defined by a $g : \mathbb{N} \to \mathbb{N}$ s.t. $\{F_{g(n)}\}_{n\in\mathbb{N}} : X \to \mathbb{C}$ restricted to E converges. This subsequence converges on X. Indeed, for any $\epsilon > 0$ there is an $e \in E$ close enough to x, $d(x, e) < \delta$, s.t. $(\forall n)$ $(|F_n(x) - F_n(e)| < \epsilon/3)$. Now by the triangle inequality, $|F_n(x) - F_m(x)| < \epsilon$ for all n, m large enough.

If K is compact, there is a *finite* set $E_n = \{e_1, ...e_n\}$ s.t. for all $x \in K$, $d(x, E_n) < \delta$, δ as above. Check that this implies uniform convergence in K. ☐

Uniform convergence implies that the limit F of the subsequence is also continuous, and in fact adjoining F to the sequence, the new sequence is also equicontinuous and pointwise equibounded.

An important example of an equibounded, equicontinuous family is the following. Consider the ball B_1 of radius one in $L^1((a,b))$ and the linear map $K : B_1 \to B_{|b-a|}$ given by $KF = \int_a^x F$. Check that $K(B_1)$ is an equibounded, equicontinuous family.

Such a linear map is called *compact operator*.

2.2 Linear operators, bounded operators, compact operators

Let X, Y be Banach spaces and $L \in \mathcal{L}(X,Y)$. L is called a *bounded operator* if $\|L\| := \sup_{\|x\|\leqslant 1} \|Lx\| < \infty$. Check $\|\|$ so defined on $L \in \mathcal{L}(X,Y)$ is a norm and the space $B(X,Y)$ of bounded operators is a Banach space. Check that $\|Lx\| \leqslant \|L\|\|x\|$. In $B(X) := B(X,X)$, multiplication is continuous: $\|L_1L_2\| \leqslant \|L_1\|\|L_2\|$. Bounded operators thus form an algebra. A Banach space which is also an algebra, and in which multiplication is continuous, as above, is called a *Banach algebra*.

An operator $K : X \to Y$ is called compact if it maps every bounded set in X into a *precompact* set in Y (that is, a set whose closure is compact). Compact operators form a linear subspace of $B(X,Y)$

Show that the compact operators from X to X, $K(X)$ form a two sided ideal in $B(X)$. That is, the left or right product of a bounded operator by a compact one is compact.

3 Hilbert spaces

The Euclidian norm in $\mathbb{R}^n$, $\|x\|^2 = \sum_{i=1}^n x_i^2$ arises from a scalar (or inner) product $\langle x,y\rangle = \sum_{i=1}^n x_iy_i$, which in turn has an important geometric significance: for two nonzero vectors x, y the angle between them satisfies $\cos\phi = \dfrac{\langle x,y\rangle}{\|x\|\|y\|}$. Since $\cos\phi \in [-1,1]$ the *Cauchy-Schwarz inequality* $|\langle x,y\rangle| \leqslant \|x\|\|y\|$ holds.

The inner product extends to $\mathbb{C}^n$, where it also defines the norm, by $\langle x,y\rangle = \sum_{i=1}^n x_i\overline{y_i}$. Cauchy-Schwarz holds in $\mathbb{C}^n$ as well, as you can easily check. Furthermore, as metric spaces, $\mathbb{R}^n, \mathbb{C}^n$ are complete: any Cauchy sequence converges (understood: to a point in $\mathbb{R}^n$ or $\mathbb{C}^n$).

$\mathbb{R}^n$ and $\mathbb{C}^n$ can be seen as the space of functions defined on $\{1, ..., n\}$ with values in $\mathbb{R}$ and $\mathbb{C}$ resp. In this interpretation, we can easily dispense with the condition that vectors have finitely many components: after all, the domain $\{1, ..., n\}$ can be replaced by any set. But can we extend the notions of norm and inner product too?

The first candidate is to take $\mathbb{N}$ instead of $\{1, ..., n\}$, and naturally define $\langle x,y\rangle$ by $\sum_{i=1}^\infty x_i\overline{y_i}$ and $\|x\| = \sqrt{\langle x,x\rangle}$. Importantly, the Cauchy-Schwarz inequality survives this extension. This space is denoted by ℓ^2.

Theorem 3 (Cauchy-Schwarz). *For any $x, y \in \ell^2$ we have $|\langle x,y\rangle| \leqslant \|x\|\|y\|$.*

Proof. There is nothing to prove if $x = 0$ or $y = 0$, so we assume this is not the case. Note now that for any z, $\|z\| \geqslant 0$. In particular, for any $a \in \mathbb{C}$ we have

$$0 \leqslant \|x - ay\|^2 = \langle x,x\rangle + |a|^2\langle y,y\rangle - 2\Re(a\langle x,y\rangle) \tag{CS 1}$$

We write $\langle x,y\rangle = |\langle x,y\rangle|e^{i\alpha}$ (if $\langle x,y\rangle = 0$ any α works). By replacing a by $|a|e^{-i\alpha}$ we see that $f(|a|) = \langle x,x\rangle + |a|^2\langle y,y\rangle - 2|a||\langle x,y\rangle| \geqslant 0$. The trick is now to note that $f(|a|)$ is a quadratic polynomial in $|a|$ which is nonnegative, and thus it has nonpositive discriminant: $4|\langle x,y\rangle|^2 - 4\langle x,x\rangle\langle y,y\rangle \leqslant 0$, which is what we intended to prove. □

Hilbert spaces are defined by abstracting the structure needed for the properties above to hold.

The following definition is standard (I copy-pasted it from Wiki).

A Hilbert space $\mathcal{H}$ is a real or complex inner product space that is also a complete metric space with respect to the distance function induced by the inner product. To say that $\mathcal{H}$ is a complex inner product space means that $\mathcal{H}$ is a complex vector space on which there is an inner product $\langle x,y\rangle$ associating a complex number to each pair of elements x,y of $\mathcal{H}$ that satisfies the following properties:

The inner product of a pair of elements is equal to the complex conjugate of the inner product of the swapped elements:

$$\langle y,x\rangle = \overline{\langle x,y\rangle}\,. \tag{1}$$

The inner product is linear in its first argument. For all complex numbers a and b,

$$\langle ax_1 + bx_2, y\rangle = a\langle x_1,y\rangle + b\langle x_2,y\rangle\,. \tag{2}$$

The inner product of an element with itself is positive definite:

$$\langle x,x\rangle \geq 0 \tag{3}$$

where the case of equality holds precisely when $x = 0$. It follows from properties (1) and (2) that a complex inner product is antilinear in its second argument, meaning that

$$\langle x, ay_1 + by_2\rangle = \bar{a}\langle x,y_1\rangle + \bar{b}\langle x,y_2\rangle\,. \tag{1+2}$$

A real inner product space is defined in the same way, except that $\mathcal{H}$ is a real vector space and the inner product takes real values. Such an inner product will be bilinear: that is, linear in each argument.

The norm is the real-valued function

$$\|x\| = \sqrt{\langle x,x\rangle}\,,$$

and the distance d between two points x,y in $\mathcal{H}$ is defined in terms of the norm by

$$d(x,y) = \|x-y\| = \sqrt{\langle x-y, x-y\rangle}\,.$$

Theorem 4 (Cauchy-Schwarz)**.** *For any $x,y \in \ell^2$ we have $|\langle x,y\rangle| \leqslant \|x\|\|y\|$.*

Proof. Check that the proof we gave for ℓ^2 goes through. □

Proposition 5. *The function $x \to \|x\|$ is a norm on $\mathcal{H}$.*

Proof. First of all, by definition, $\|x\| = 0$ iff $x = 0$ and $\|\lambda x\| = |\lambda|\|x\|$. To prove the triangle inequality, we note that

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re\langle x, y\rangle \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

$\square$

Proposition 6. *ℓ^2 is complete.*

Proof. If $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in ℓ^2, then for every $i \in \mathbb{N}$ the number sequence $\{(x_n)_i\}_{n\in\mathbb{N}}$ is Cauchy (indeed $|(x_n)_i - (x_m)_i|^2 \leqslant \|x_n - x_m\|^2$). Let $y_i = \lim_n (x_n)_i$. Let n_0 be s.t. $(\forall n, m \geqslant n_0), (\|x_n - x_m\| < 1)$. The triangle inequality implies that $\forall n \geqslant n_0, \|x_n\| \leqslant C$ where $C = 1 + \|x_{n_0}\|$. It follows that, for all n, $\sum_{i=1}^{n} |y_i|^2 = \lim_{k\to\infty} \sum_{i=1}^{n} |(x_k)_i|^2 \leqslant C$ and since $|y_i|$ are positive and the sums are bounded, the sum converges to $\|y\|^2 \leqslant C$, that is $y \in \ell^2$. Similarly, since $\lim_{k\to\infty} \sum_{i=0}^{n} |(x_k)_i - y_i|^2 = 0$ for any n, by the above we can use dominated convergence to show that $\|x_k - y\| \to 0$.

$\square$

Proposition 7. *The inner product is a continuous function from $\mathcal{H} \times \mathcal{H}$ to $\mathbb{C}$. In particular, if $\{(x_n, y_n)\}_{n\in\mathbb{N}}$ converges to (x, y) then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Proof. By Cauchy-Schwarz

$$|\langle x - x', y - y'\rangle| \leqslant \|x - x'\|\|y - y'\|$$

and the result follows easily (write down the details!). $\square$

Proposition 8 (The parallelogram law)**.**

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \tag{P.L.}$$

Proof. A straightforward calculation, see (CS 1) above. $\square$

We see that a Hilbert space is a complete normed space where the norm comes from an inner product. A natural and important question arises: given a norm, can we always define an inner product that induces the norm? The answer is no and, remarkably, (P.L.) is the necessary and sufficient condition for the norm to come from an inner product.

Proposition 9. *Let $\mathcal{S}$ be a complete normed space, with norm $\|\cdot\|$. Then the norm comes from an inner product iff it satisfies the parallelogram law.*

Proof. We have already shown that an inner-product-induced norm satisfies the parallelogram law. In the opposite direction, a calculation assuming the existence of an inner product leads the following explicit formula for the inner product, called *the polarization identity*:

$$\langle x,\ y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right)\ \forall\, x, y \in \mathcal{H}$$

It remains to check that *assuming the parallelogram law* the formula above defines an inner product (meaning: with properties (1)...(3) above). This is elementary, but by no means trivial! See

Internet-proof, 27., p. 157-160. Interestingly, the proof is an adaptation of Euclid's proof of the three line theorem; see N. Falkner, MAA 100,3, (1993). $\square$

Corollary 10. *The inner product is continuous.*

3.1 Orthogonality

The notion of orthogonality, $x \perp y$ if by definition $\langle x, y\rangle = 0$ obviously extends to general Hilbert spaces. So does the following

Proposition 11. *If $x_1, ..., x_n$ are pairwise orthogonal, then*

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

Proof. Since pairwise orthogonality implies $x_1 + \cdots + x_{n-1} \perp x_n$, the proof follows by induction from the case $n = 2$, which is immediate. $\square$

3.2 The Gram-Schmidt process

We recall that, given a family $\{x_i\}_{i\in\mathbb{N}}$ of linearly independent vectors (no finite linear combination with nonzero coefficients vanishes), we can construct, from them, an orthonormal family $\{e_i\}_{i\in\mathbb{N}}$, inductively: start with $v_1 = x_1$; let c be s.t. $v_2 = cv_1 + x_2 \perp v_1$ (calculate c!). Having constructed $v_1, ..., v_n$, choose $c_{n1}, ..., c_{nn}$ s.t. $v_{n+1} = c_{n1}x_1 + ... + c_{nn}x_n + x_{n+1}$ is orthogonal on $v_1, ..., v_n$ (check that there is such a choice!). Then $\{v_i\}_{i\in\mathbb{N}}$ is an orthogonal family with the property that Span($\{x_1, ..., x_n\}$=Span($\{v_1, ..., v_n\}$ for all n. All these v_i are nonzero vectors, and an orthonormal family is simply given by $e_i = v_i/\|v_i\|$.

Proposition 12. *Let $\{v_i\}_{i\in\mathbb{N}}$ be a set of vectors in $\mathcal{H}$ and let $\mathcal{V} = span(\{v_i\}_{i\in\mathbb{N}})$. We assume (without loss of generality, as we will see) that$\mathcal{V}$ is infinite dimensional. Then there exists an orthonormal set $\{e_i\}_{i\in\mathbb{N}}$ such that $\mathcal{V} = span(\{e_i\}_{i\in\mathbb{N}})$.*

Proof. We can assume w.l.o.g. that $\{v_i\}_{i\in\mathbb{N}}$ are linearly independent, otherwise it is easy to replace them with such a set (check!). Then the Gram-Schmidt procedure above precisely provides such a set $\{e_i\}_{i\in\mathbb{N}}$. $\square$

Concretely, with $\pi_v x = \dfrac{\langle v, x\rangle}{\langle v, v\rangle} v$ the process is

$$\begin{aligned}
v_1 &= x_1, & e_1 &= \frac{v_1}{\|v_1\|} \\
v_2 &= x_2 - \pi_{v_1} x_2, & e_2 &= \frac{v_2}{\|v_2\|} \\
&\vdots & &\vdots \\
v_k &= x_k - \sum_{j=1}^{k-1} \pi_{v_j} x_k, & e_k &= \frac{v_k}{\|v_k\|}. \\
\cdots
\end{aligned}$$

3.3 Another simple proof of Cauchy-Schwarz

Proof. In case x, y are linearly dependent the result is immediate. Otherwise wlog, we may assume $\|x\| = \|y\| = 1$. Define $e_1 = x$ and let e_2 be obtained from e_1 and y by Gram-Schmidt. Then $y = y_1 e_1 + y_2 e_2$ and

$$|\langle x, y\rangle|^2 = y_1^2 \leqslant y_1^2 + y_2^2 = 1 \text{ (draw a picture too!)}$$

□

3.4 Orthogonal projections

Definition 13 (The orthogonal complement of a space)**.** *If $\mathcal{S}$ is a subspace of $\mathcal{H}$, then its orthogonal complement, $\mathcal{S}^\perp$ is the **closed linear subspace** (check these properties!) of $\mathcal{H}$ defined by*

$$\mathcal{S}^\perp = \{x \in \mathcal{H} : (\forall y \in \mathcal{S})(\langle x, y\rangle = 0\}$$

The direct sum of two subspaces $\mathcal{V}_1, \mathcal{V}_2$ is defined, as usual, by

$$\mathcal{V}_1 \oplus \mathcal{V}_2 = \{v_1 + v_2 : v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2\}$$

Proposition 14. *If $\mathcal{M} \subset \mathcal{H}$ is a closed subspace, then $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. Any $x \in \mathcal{H}$ can be uniquely written as $m + m^\perp$ with $m \in \mathcal{M}$ and $m^\perp \in \mathcal{M}^\perp$.*

We have $(\mathcal{M}^\perp)^\perp = \mathcal{M}$. Let $\pi_{\mathcal{M}}$ be given by $\pi_{\mathcal{M}}(x) = m$ with x, m as above. We have $\forall m \in \mathcal{M} \neq \pi_{\mathcal{M}}(x), \|x - m\| > \|x - \pi_{\mathcal{M}}(x)\|$, that is $\pi_{\mathcal{M}}(x)$ is the closest point to x in $\mathcal{M}$, and $x - \pi_{\mathcal{M}} x \in \mathcal{M}^\perp$. Conversely, if $m \in \mathcal{M}$ is s.t. $x - m \in \mathcal{M}^\perp$, then $m = \pi_{\mathcal{M}} x$.

If $t \in \mathcal{M}$, then $\pi_{\mathcal{M}} t =$. In particular, $\pi_{\mathcal{M}}^2 x = \pi_{\mathcal{M}} x$. If $y \in \mathcal{M}^\perp$ then $\pi_{\mathcal{M}} y = 0$.

Proof. • We first show that there is a $y = \pi_{\mathcal{M}}(x) \in \mathcal{M}$ s.t. $\forall m \in \mathcal{M} \neq y,\ \|x - m\| > \|x - y\|$.

Let $d = \inf_{y \in \mathcal{M}} \|x - y\|$. Obviously $d \leqslant \|x\|$, since $0 \in \mathcal{M}$. Thus there is a sequence $\{y_m\}_{m \in \mathbb{N}}$ in $\mathcal{M}$ s.t. $d - \|x - y_m\| \to 0$. We show that this sequence is convergent in $\mathcal{M}$. Since $\mathcal{M}$ is a closed subspace of the complete Hilbert space $\mathcal{H}$, it suffices to show that $\{y_m\}_{m \in \mathbb{N}}$ is Cauchy. Here we use the parallelogram law:

$$\|y_m - y_n\|^2 = \|(x - y_n) - (x - y_m)\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|2x - y_n - y_m\|^2$$
$$2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\|x - \tfrac{1}{2}(y_n + y_m)\|^2 \leqslant 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4d^2 \to 0 \text{ as } n, m \to \infty$$

Thus $y_n \to y =: \pi_{\mathcal{M}} x$, and clearly $\|x - \pi_{\mathcal{M}} x\| = d$.

• Next we show that $\langle x - \pi_{\mathcal{M}} x, m\rangle = 0$ for any $m \in \mathcal{M}$. Take $m = \pi_{\mathcal{M}} x - \alpha t, t \in \mathcal{M}$. Then

$$\|x - m\|^2 = \|x - \pi_{\mathcal{M}} x\|^2 + |\alpha|^2 \|t\|^2 + 2\Re(\overline{\alpha}\langle x - \pi_{\mathcal{M}} x, t\rangle$$

Assume there were a t s.t. $\langle x - \pi_{\mathcal{M}} x, t\rangle \neq 0$. Write $\langle x - \pi_{\mathcal{M}} x, t\rangle = |\langle x - \pi_{\mathcal{M}} x, t\rangle| e^{i\phi}$ and choose α s.t. $\alpha = -|\alpha| e^{i\phi}$. Let $m = \pi_{\mathcal{M}} x + \alpha t$. We then get

$$\|x - m\|^2 = \|x - \pi_{\mathcal{M}} x\|^2 + |\alpha|^2 \|t\|^2 - 2|\alpha||\langle x - \pi_{\mathcal{M}} x, t\rangle|$$

Since $|\alpha|\|t\|^2 - 2|\langle x - \pi_{\mathcal{M}} x, t\rangle| \to -2|\langle x - \pi_{\mathcal{M}} x, t\rangle| < 0$ as $\alpha \to 0$, check that if $|\alpha|$ is small enough then $|\alpha|^2 \|t\|^2 - 2|\alpha||\langle x - \pi_{\mathcal{M}} x, t\rangle| < 0$ contradicting the minimality of d.

Since $\langle x - \pi_{\mathcal{M}}x, m\rangle = 0$ for any $m \in \mathcal{M}$, we have $\|x - m\|^2 = \|x - \pi_{\mathcal{M}}x + (\pi_{\mathcal{M}}x - m)\|^2 = d^2 + \|\pi_{\mathcal{M}}x - m\|^2 > d^2$ unless $m = \pi_{\mathcal{M}}x$.

Also as a consequence of the fact that $\langle x - \pi_{\mathcal{M}}x, m\rangle = 0$ for any $m \in \mathcal{M}$, by definition

$$x - \pi_{\mathcal{M}}x \in \mathcal{M}^{\perp}$$

Now, if $t \in \mathcal{M}$, then $d = \|t - t\| = 0$ and thus $\pi_{\mathcal{M}}t = t$. If $\pi_{\mathcal{M}}t = t$, then by definition $t \in \mathcal{M}$.

• The space $\mathcal{M}^{\perp}$ is also linear and closed, because of the continuity of the scalar product. Therefore there is a unique $\pi_{\mathcal{M}^{\perp}}x \in \mathcal{M}^{\perp}$ s.t. $\inf_{y\in\mathcal{M}^{\perp}} \|x - y\| = \|x - \pi_{\mathcal{M}^{\perp}}x\|$ and $x - \pi_{\mathcal{M}^{\perp}}x$ is orthogonal on all the vectors in $\mathcal{M}^{\perp}$. Now,

$$z = x - \pi_{\mathcal{M}^{\perp}}x - \pi_{\mathcal{M}}x \Rightarrow \pi_{\mathcal{M}}z = \pi_{\mathcal{M}^{\perp}}z = 0$$

Thus

$$z \in \mathcal{M}^{\perp} \cap (\mathcal{M}^{\perp})^{\perp} = \{0\} \Rightarrow z = 0$$

Clearly any vector in $\mathcal{M}$ is in $(\mathcal{M}^{\perp})^{\perp}$. Conversely, $x \in (\mathcal{M}^{\perp})^{\perp} \Rightarrow \pi_{\mathcal{M}^{\perp}}x = 0 \Rightarrow x = \pi_{\mathcal{M}}x \in \mathcal{M}$.

□

It also follows that

$$\pi_{\mathcal{M}} + \pi_{\mathcal{M}^{\perp}} = I$$

the identity function on $\mathcal{H}$.

3.5 The Riesz-Fréchet theorem (Riesz representation theorem)

Proposition 15 (Riesz representation theorem; or the Riesz-Fréchet theorem)**.** *Let $\Lambda \in \mathcal{H}^*$. Then there is a unique $y \in \mathcal{H}$ s.t.*

$$\forall x \in \mathcal{H},\ \ \Lambda x = \langle x, y\rangle \tag{1}$$

Furthermore, $\|\Lambda\|_{\mathcal{H}^*} = \|y\|_{\mathcal{H}}$. In particular $\mathcal{H} = \mathcal{H}^*$.

Proof. Let $\mathcal{M} = \{x \in \mathcal{H} : \Lambda x = 0\}$. Clearly $\mathcal{M}$ is a closed linear subspace of $\mathcal{H}$ (why?). Now if $\mathcal{M} = \mathcal{H}$ then $y = 0$ is unique s.t. (1) holds, and we are done. Otherwise, we claim that $\mathcal{M}^{\perp}$ is one dimensional. Indeed, let $0 \neq e \in \mathcal{M}^{\perp}$; note that this implies $\Lambda e \neq 0$. We rescale e so that $\Lambda e = 1$. Let $0 \neq x \in \mathcal{M}^{\perp}$ and let $\Lambda x = b$ (again, necessarily $b \neq 0$). Then

$$x - be \in \mathcal{M}^{\perp} \text{ and } x - be \in \mathcal{M}\ \ (\text{since } \Lambda(x - be) = 0)\ \Rightarrow x - be = 0$$

This means x is linearly dependent on e, as stated. For $x \in \mathcal{H}$ we have, for some c,

$$x = \pi_{\mathcal{M}}x + \pi_e x = \pi_{\mathcal{M}}x + \frac{\langle x, e\rangle}{\|e\|^2}e \Rightarrow \Lambda x = \frac{\langle x, e\rangle}{\|e\|^2} =: \langle x, y\rangle \Rightarrow |\Lambda x| \leqslant \frac{\|x\|}{\|e\|} \Rightarrow \|\Lambda\| \leqslant \frac{1}{\|e\|}$$

Uniqueness follows from this explicit calculation. For the norm, we note the inequality above and the fact that by definition

$$\frac{|\Lambda e|}{\|e\|} = \frac{1}{\|e\|} \Rightarrow \|\Lambda\| = \frac{1}{\|e\|} = \|y\|$$

□

3.6 Bessel's inequality, Parseval's equality, Hilbert space bases

Theorem 16 (Bessel's inequality). *Let $\{e_i\}_{i\in\mathbb{N}}$ be an orthonormal sequence in $\mathcal{H}$. Then*

$$\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 \leqslant \|x\|^2$$

Proof. Let $\mathcal{H}_n = \mathrm{Span}(\{e_1, ..., e_n\}) := \{c_1e_1 + \cdots c_ne_n : c_i \in \mathbb{C}\}$. Clearly, $\mathcal{H}_n$ is a closed subspace of $\mathcal{H}$. We can then write

$$x = \pi_{\mathcal{H}_n}x + x^{\perp} = \sum_{i=1}^{n}\langle x, e_i\rangle e_i + x^{\perp}$$

and, by the Pythagorean equality,

$$\|x\|^2 = \|\pi_{\mathcal{H}_n}x\|^2 + \|x^{\perp}\|^2 \geqslant \|\pi_{\mathcal{H}_n}x\|^2 = \sum_{i=1}^{n}|\langle x, e_i\rangle|^2$$

Since this holds for any n, the result follows. □

Direct proof.

$$0 \leqslant \left\| x - \sum_{k=1}^{n}\langle x, e_k\rangle e_k \right\|^2 = \|x\|^2 - 2\sum_{k=1}^{n}|\langle x, e_k\rangle|^2 + \sum_{k=1}^{n}|\langle x, e_k\rangle|^2 = \|x\|^2 - \sum_{k=1}^{n}|\langle x, e_k\rangle|^2$$

□

Corollary 17. *Let $\{e_\alpha\}_{\alpha\in A}$ be an orthonormal family in $\mathcal{H}$. Then, for any $x \in \mathcal{H}$*

$$\sum_{\alpha\in A}|\langle x, e_\alpha\rangle|^2 =: \sup_{\substack{F\subset A\\ F \text{ finite}}}\sum_{\alpha\in F}|\langle x, e_\alpha\rangle|^2 \leqslant \|x\|^2$$

and the set $B = \{\alpha \in A : \langle x, e_\alpha\rangle \neq 0\}$ is countable. (Indeed each of the sets $\{\alpha \in A : |\langle x, e_\alpha\rangle| > 1/n\}$ must be finite.)

Proof. Only countability needs to be shown. It is well known however that an uncountable sum, defined as above, of strictly positive numbers is infinity. □

Definition 18. *An orthonormal set $\{e_\alpha\}_{\alpha\in A}$ is called a Hilbert space basis in the separable Hilbert space $\mathcal{H}$ if any $x \in \mathcal{H}$ can be written as a finite or countably infinite linear combination*

$$x = \sum_{\alpha\in A} c_\alpha e_\alpha$$

Note that this is not *the same as a vector space basis!*

Note that, because of the continuity of the inner product, we can calculate the coefficients c_k

by taking inner products with the e_k:

$$x = \sum_{k=1}^{\infty} \langle x, e_{\alpha_k} \rangle e_{\alpha_k} \tag{Fourier}$$

Proposition 19. *Any separable Hilbert space $\mathcal{H}$ has a (Hilbert space) basis.*

Proof. Let $\{v_i\}_{i\in\mathbb{N}}$ be a countable dense set in $\mathcal{H}$. The closure of the span of $\{v_i\}_{i\in\mathbb{N}}$ is, of course, $\mathcal{H}$, and so is the span of $\{e_i\}_{i\in\mathbb{N}}$, constructed by Gram-Schmidt. Note that, by Bessel's inequality,

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leqslant \|x\|^2 \Rightarrow \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k \in \mathcal{H}$$

But then $x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ since their difference is orthogonal to all e_k. □

Corollary 20. *$\mathcal{H}$ is separable iff it has a countable base.*

Possibly nonseparable $\mathcal{H}$ have possibly uncountable bases; the proof, as you would expect, uses AC (Zorn's lemma is convenient). Try to phrase a statement and prove it.

Note 21. *in applications* Nonseparable Hilbert spaces are quite rare in applications. Furthermore, Corollary 17 shows that *even in non-separable Hilbert spaces* we only need a countable family *at a time*, so we will not worry about more general orthonormal families, as they would not add anything nontrivial to the results below.

Theorem 22. *If $\{e_i\}_{i\in\mathbb{N}}$ is an orthonormal set in a separable Hilbert space $\mathcal{H}$, then the following are equivalent:*

*a. (**Completeness**) If $\forall j$, $\langle x, e_j \rangle = 0$, then $x = 0$.*

*b. (**Parseval's identity**) $\forall x \in \mathcal{H}$, $\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$.*

c. $\{e_i\}_{i\in\mathbb{N}}$ is a basis for $\mathcal{H}$.

Proof. (b.$\Rightarrow$ a.) is clear.

(a.$\Rightarrow$ c.) Bessel's inequality implies that $\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$ converges. The series $\sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ converges as well, since $\left\| \sum_{k=m}^{n} \langle x, e_k \rangle e_k \right\|^2 = \sum_{k=m}^{n} |\langle x, e_k \rangle|^2$. Clearly $\left\langle e_j, x - \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k \right\rangle = 0 \forall j$ implying the result.

(c.$\Rightarrow$ b.) This is simply Pythagoras plus the continuity of the norm. □

An example of a Hilbert basis is the set $e_k = (0, .., 1, 0...)$, with 1 in the kth position, in ℓ^2.

Definition 23. *Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces and $U : \mathcal{H}_1 \to \mathcal{H}_2$ be one-to-one linear and norm preserving. Then U is called an isometry.*

Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces and $U : \mathcal{H}_1 \to \mathcal{H}_2$ be one-to-one onto, linear and inner product preserving, $\langle Ux, Uy \rangle = \langle x, y \rangle$. Then U is called unitary.

Proposition 24. *U is unitary iff it is an isometry.*

Proof. Preservation of the inner product immediately implies norm preservation. In the opposite direction, use the polarization identity! □

Unitary maps are isomorphisms of Hilbert spaces: they preserve all the relevant abstract features of a Hilbert space.

Proposition 25 (Any two separable Hilbert spaces are isomorphic)**.** *Any separable Hilbert space $\mathcal{H}$ is isomorphic to ℓ^2.*

Proof. Let $\{e_n\}_{n\in\mathbb{N}}$ be a Hilbert basis in $\mathcal{H}$. Define $U : \mathcal{H} \to \ell^2$ by $U(x) = (\langle x, e_1\rangle, ..., \langle x, e_n\rangle, ...)$. Check that this is an isometry! □

For a nonseparable Hilbert space with a Hilbert basis $\{e_\alpha\}_{\alpha\in A}$, a similar statement holds, except $\ell^2 = \ell^2(\mathbb{N})$ is replaced by the more general $\ell^2(A)$. See Folland for the definitions. Then prove the statement yourselves.

3.7 Adjoints

In analysis, elements of $L(X, Y)$ are often called **bounded operators**.

Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded operator. Its adjoint is defined as the operator A^* with the property

$$\forall x, y \in \mathcal{H}, \langle Ax, y\rangle = \langle x, A^* y\rangle$$

Exercise 1. *A. Use the Riesz representation theorem to show that A^* defined above exists and is unique. Check that $A^{**} = A$.*

B. For fixed y, $\langle x, A^* y\rangle$ is a linear functional, and by the Riesz representation theorem

$$\|A^* y\| = \sup_{\|u\|=1} |\langle u, A^* y\rangle| = \sup_{\|u\|=1} |\langle Au, y\rangle| \leqslant \|A\|\|y\|$$

This implies $\|A^*\| \leqslant \|A\|$. Then $\|A\| = \|A^{**}\| \leqslant \|A^*\|$, hence $\|A\| = \|A^*\|$. Thus A^* is a bounded operator with the same norm as A. Check that $\|A^* A\| = \|AA^*\| = \|A\|^2$.

4 Consequences of the Baire category theorem

In the following, we use the notations

$$B_a(x) = \{y \in X : \|y - x\| < a\}; \quad B_a(0) \equiv B_a$$

A Baire space is a topological space with the following property: for any countable family of open dense sets $\{U_n\}_{n=1}^\infty$, their intersection $\bigcap_{n=1}^\infty U_n$ is dense.

As a reminder, the Baire category theorem states

Theorem 26 (Baire category theorem)**.** *Every complete metric space is a Baire space. Equivalently, a non-empty complete metric space is* **not** *a countable union of nowhere-dense sets [equivalently, nowhere-dense closed sets].*

This theorem has a number of fundamental consequences in analysis.

Theorem 27 (Uniform boundedness principle)**.** *a. Assume X is a Banach space, Y is a normed space, and $A \subset L(X,Y)$.*

Then

$$\Big(\forall x \in X, \sup_{T\in A} \|Tx\| < \infty\Big) \Leftrightarrow \sup_{T\in A} \|T\| < \infty$$

b. (Generalization) If X and Y are normed spaces and there is a non-meager set $X_1 \subset X$ such that $(\forall x \in X_1, \sup_{T\in A} \|Tx\| < \infty)$, then $\sup_{T\in A} \|T\| < \infty$.

Proof. a. ($\Leftarrow$) is trivial. ($\Rightarrow$) For $n \in \mathbb{N}$ let

$$E_n = \{x \in X : \sup_{T\in A} \|Tx\| \leqslant n\} = \cap_{T\in A}\{x \in X : \forall T \in A, \|Tx\| \leqslant n\}$$

Clearly, E_n are closed and $X = \bigcup_{n\in\mathbb{N}} E_n$. Then, there is an m s.t. E_m has nonempty interior, $\overline{B_a}(x_0) \subset E_m$. Take any u with $\|u\| = 1$. Then both x_0 and $x_0 + au$ are in $\overline{B_a}(x_0)$ and

$$Tu = \tfrac{1}{a}T(au) = \tfrac{1}{a}T(x_0 + au - x_0) \Rightarrow \|Tu\| \leqslant \tfrac{1}{a}\|T(x_0 + au)\| + \tfrac{1}{a}\|Tx_0\| \leqslant \frac{2m}{a} \Rightarrow \sup_{\substack{\|u\|=1\\ T\in A}} \|Tu\| \leqslant \frac{2m}{a}$$

b. Copy the proof above, basically. □

Theorem 28 (The open mapping theorem)**.** *Let X, Y be Banach spaces and $T \in L(X,Y)$ be* **surjective***. Then if $\mathcal{O}$ is open in X, $T(\mathcal{O})$ is open in Y.*

Proof *(*an adaptation of Reed-Simon p. 82*).* We start with some straightforward preparatory steps to reduce the complexity of the more difficult part of the proof. It is enough to prove that for any x and $\mathcal{N}_x$ a neighborhood of it, $T(\mathcal{N}_x)$ is a neighborhood of $T(x)$. By linearity,

$$T(x + \mathcal{O}) = T(x) + T(\mathcal{O})$$

and it suffices to prove this for $x = 0$. Clearly the property holds if for all r there is an r' s.t.

$$T(B_r^X) \supset B_{r'}^Y; \tag{A}$$

By linearity

$$T(B_r) = rT(B_1)$$

and it is enough to prove (A) for *some* r. Again by linearity it is enough to show that for some r, $T(B_r)$ contains *some ball*, not necessarily centered at zero, that is, $T(B_r)$ has nonempty interior.

Now, $X = \bigcup_{n=1}^{\infty} B_n$ and T is onto, we must have

$$Y = \bigcup_{n=1}^{\infty} T(B_n) = \bigcup_{n=1}^{\infty} \overline{T(B_n)}$$

By the Baire category theorem, at least one of the $\overline{T(B_n)}$ has nonempty interior. By linearity, this happens for all n.

What we really need however is stronger, namely that, for some n, $T(B_n)$ has nonempty interior (thus all $T(B_j)$ do). To go from $\overline{T(B_1)}$ to $T(B_m)$ for some m is nontrivial. (Note that $T(\overline{B_1}) \subset \overline{T(B_1)}$, but the inclusion can be strict as Exercise 2 below shows.)

Now, since $\overline{T(B_n)}$ contains some ball, by dilations $\overline{T(B_{n+k})}$ contains a ball around zero, therefore there is an ϵ s.t. $B_\epsilon \subset \overline{T(B_1)}$. The following lemma implies the result of the theorem.

Lemma 29. *If $\overline{T(B_1)}$ contains a ball, B_ϵ, then $\overline{T(B_1)} \subset T(B_2)$ (in fact $\overline{T(B_1)} \subset T(B_{1+\delta})$ for any $\delta > 0$.)*

Proof. Let $y \in \overline{T(B_1)}$. Then, there are points x in B_1 s.t. $T(x)$ is as close as we want to y. Let x_1 be s.t. $y - T(x_1) \in B_{\epsilon/2} \subset \overline{T(B_{1/2})}$ (by scaling). We continue in the same way, inductively: let $x_2 \in T(B_{1/2})$ be s.t. $(y - T(x_1)) - T(x_2) \in B_{\epsilon/4}$ ... let $x_{n+1} \in T(B_{1/2^n})$ be s.t. $y - T(x_1) - \cdots - T(x_{n+1}) \in B_{\epsilon/2^n}$. But you see that $x = \sum_n x_n$ converges to an element in $\overline{B_{1/2+1/4+\ldots}} \subset B_2$, and by continuity $y = Tx$, thus $y \in B_2$. (By modifying the selection of $\{x_n\}$, you can prove the result above with $1 + \delta$ instead of 2.) □

□

Theorem 30 (Inverse mapping theorem)**.** *If $T \in L(X, Y)$ is one to one onto, then T^{-1} is also continuous, $T^{-1} \in L(Y, X)$.*

Proof. T is one-to-one, thus onto, thus open, implying by definition continuity of T^{-1}. □

Definition 31. *If $T \in L(X, Y)$, its* **graph** *is*

$$\Gamma(T) = \{(x, y) \in X \times Y : y = Tx\} = \{(x, Tx) : x \in X\}$$

Theorem 32 (Closed graph theorem)**.** *Let X, Y be Banach spaces and $T \in L(X, Y)$. Then T is bounded* **iff** *$\Gamma(T)$ is a closed subset of $X \times Y$ (note that we do not assume that $T(X) = Y$.)*

Proof. Assume T is continuous. If $\{(x_n, T(x_n))\}_{n \in \mathbb{N}}$ converges to (x, y) then in particular $x_n \to x$. But then, by continuity $T(x_n) \to y$, hence $(x, y) \in \Gamma(T)$ (are sequences sufficient in this context?)

In the opposite direction, note that $\Gamma(T)$ is a linear closed subspace of $X \times Y$, thus itself a Banach space. Now the projections $\pi_1 : \Gamma(T) \to X, \pi_2 : \Gamma(X) \to Y$ given y $\pi_1(x, Tx) = x$ and $\pi_2(x, Tx) = Tx$ are manifestly continuous. Note that $\pi_1(x, Tx) = x$ is a linear bijection between X and $\Gamma(T)$, thus, by the inverse mapping theorem, its inverse is continuous too. But $T = \pi_2(\pi_1^{-1})$ is a composition of continuous functions. □

Note 33. *We see that a bounded operator between Banach spaces can fail to have a bounded inverse only for the trivial reason that it does not have an inverse* **at all** *(that is, if it is not surjective or not injective).*

Definition 34. *An A operator is closed if $\Gamma(A)$ is closed.*

Exercise 2. *a. Let $\mathcal{H} = L^2[0, 1]$. Then the operator A defined on $\mathcal{H}$ by*

$$(Af)(x) = \int_0^x f(s)ds$$

is bounded (check).

a. Show that $M =$ran$(A) \neq \mathcal{H}$.

b. What is $M^\perp$?

c. Show that (the linear space) ran(A) is **not** *closed. (This implies A is not invertible from $\mathcal{H}$ to $\mathcal{H}$.) Note that $\Gamma(A)$ is* **closed** *nevertheless. ($\Gamma(A)$ is a closed subspace of $\mathcal{H} \times \mathcal{H}$ which does not mean that the* direct images *of the projections $\pi_{1,2}$ of $\Gamma(A)$ are closed!)*

d. show that $A(\mathcal{H})^\circ (= A(\overline{\mathcal{H}})^\circ) = \emptyset$ while $\overline{A(\mathcal{H})} = \mathcal{H}$.

Corollary 35. *Corollary 1: Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms defined on X. Assume X is a Banach space in both norms, and that furthermore, for some $C_1 > 0$ and all $x \in X$ we have $\|x\|_1 \leqslant C_1 \|x\|_2$. Then the two norms are equivalent, that is, there is a $C_2 > 0$ s.t. for all $x \in X$, $\|x\|_2 \leqslant C_2 \|x\|_1$.*

Proof. Exercise. (Hint: take $T = I$, that is, $Tx = x$ for all x.) □

5 Semicontinuouos functions

Definition 36. *Let f be a real-valued (or extended-real valued [2]) function on X, a topological space. Then f is called* **lower semicontinuous** *if for any $\alpha \in \mathbb{R}$ the set*

$$\{x : f(x) > \alpha\}$$

is open, and **upper semicontinuous** *if for any $\alpha \in \mathbb{R}$ the set*

$$\{x : f(x) < \alpha\}$$

is open.

Check that a function $f : X \to \mathbb{R}$ is continuous iff it is both upper and lower semicontinuous. Examples of functions that are only semi-continuous are:

a. Characteristic functions of **open sets**: these are lower semicontinuous.

b. Characteristic functions of **closed sets**: these are upper semicontinuous.

c. The sup of any collection of lower semicontinuous functions is lower semicontinuous. The inf of any collection of upper semicontinuous functions is upper semicontinuous.

Though it's straightforward, it's useful to go through the arguments and check all this.

5.1 Support of a function

Definition 37. *If f is a complex-valued function on X, then the support of f is defined as supp$(f) = \overline{\{x : f(x) \neq 0\}}$.*

We say f is supported in $\mathcal{O}$ if supp$f \subset \mathcal{O}$*, and we write $f \prec \mathcal{O}$. If $f \in C(X, [0,1])$, C is closed and $f(C) = \{1\}$, then we write $K \prec f$.*

5.2 Urysohn's lemma

In a normal space, closed sets are separated by open sets. It means, if C_1, C_2 are closed, then there *are disjoint open sets* $\mathcal{O}_1, \mathcal{O}_2$ containing C_1, C_2, respectively. This property is, interestingly, equivalent to an apparently stronger property, that there is a continuous function f which is zero on C_1 and one on C_2. That is, indicator functions can be smoothened in a way that does not alter their functionality.

[2] In the sense of the one point compactification of $\mathbb{R}$.

Note 38. *In a normal space, for any closed set C and open set $\mathcal{O}$ there is an open sets $\mathcal{O}_1$ s.t.*

$$C \subset \mathcal{O}_1 \subset \overline{\mathcal{O}_1} \subset \mathcal{O}$$

(check this: $C \cap \overline{\mathcal{O}^c} = \emptyset$; thus, we can separate C from $\overline{\mathcal{O}^c}$ by open sets...)

Theorem 39 (Urysohn's lemma)**.** *Let X be normal. For any two nonempty closed disjoint subsets A, B of X, there is an $f \in C(X, [0,1])$ such that $f(A) = \{0\}$ and $f(B) = \{1\}$.*

Equivalently,

"For any $C \subset \mathcal{O}$, C closed and $\mathcal{O}$ open, there is an $f \in C(X, [0,1])$ such that $f(C) = \{1\}$ *and* $f(\mathcal{O}) = \{0\}$."

Note that this does not say that f can only be zero on A, or 1 on B, a property which is *stronger*. This theorem is quite deep. The idea is to squeeze a countably infinite family of (distinct) open sets between A and B, order them using the rationals in $[0,1]$, $\{\mathcal{O}_r\}_{r\in\mathbb{Q}}$ in such a way that the order of the rationals is preserved

$$s > r \Rightarrow \overline{\mathcal{O}_s} \subset \mathcal{O}_r \tag{(*)}$$

(meaning also that the sets are densely ordered.) Define $f(x) = r$ if $x \in \mathcal{O}_r$ and extend f by continuity. Basically.

It's not obvious that such a construction is possible and that it yields the right answer; we need more work.

Proof (following Rudin). Let $r_0 = 0$, $r_1 = 1$, and let $r_3, r_4, ...$ be an enumeration of the rationals in $(0,1)$. Let $\mathcal{O}_0, \mathcal{O}_1$ be open sets such that

$$C \subset \mathcal{O}_1 \subset \overline{\mathcal{O}_1} \subset \mathcal{O}_0 \subset \overline{\mathcal{O}_0} \subset \mathcal{O}$$

Inductively, suppose that for $n \geqslant 2$ we have constructed $\mathcal{O}_{r_1}, ..., \mathcal{O}_{r_n}$ so that for all $i, j \leqslant n$ we have

$$r_j > r_i \Rightarrow \overline{\mathcal{O}_{r_j}} \subset \mathcal{O}_{r_i}$$

Order the $r_i, i \leqslant n : 0 < r'_1 < ... < r'_n < 1$. Take the next rational in the list, r_{n+1}, and and place it between r'_i and r'_{i+1} so that

$$0 < r'_1 < r'_2 < ... < r'_i < r'_{n+1} < r'_{i+1} < ... < r'_n < 1$$

Now choose a $\mathcal{O}_{r_{n+1}}$ so that

$$\overline{\mathcal{O}_{r'_{i+1}}} \subset \mathcal{O}_{r_{n+1}} \subset \overline{\mathcal{O}_{r_{n+1}}} \subset \mathcal{O}_{r'_i}$$

In this way, we get a family $\{\mathcal{O}_r\}_{r\in\mathbb{Q}\cap(0,1)}$ with the property (*) above.

Let now

$$f_r(x) = \begin{cases} r \text{ if } x \in \mathcal{O}_r \\ 0 \text{ } otherwise \end{cases} \; ; \; f = \sup_r f_r; \quad g_s(x) = \begin{cases} 1 \text{ if } x \in \overline{\mathcal{O}_s} \\ s \text{ } otherwise \end{cases} \; ; \; g = \inf_s g_s \tag{(**)}$$

$\square$

Clearly, f is lower semicontinuous, g is upper semicontinuous, $f(X) \subset [0,1]$, $f(C) = \{1\}$, $f(\overline{\mathcal{O}_0}) = \{0\}$.

We show that $f = g$, which implies continuity. Note that, by the ordering of the sets, $f_r(x) = r < 1$ for some r and $g_s(x) = s$, then $x \notin \overline{\mathcal{O}_s}$, thus $\overline{\mathcal{O}_r} \not\subset \mathcal{O}_s \Rightarrow r \not> s$, hence $g_s(x) \geqslant f_r(x)$ for all $\forall r, s,\ f_r \leqslant g_s$, hence $f \leqslant g$.

Assume $f(x) < g(x)$ for some $x \in [0,1]$. Then $f(x) < r < s < g(x)$ for some rationals r, s. Since $f(x) < r$ we have that $f_r(x) = 0$ implying $x \notin \mathcal{O}_r$. Similarly, since $g(x) > s$ we must have $x \in \overline{\mathcal{O}_s}$. This contradicts (**).

5.3 A few more facts about LCH

In the following, $\mathbb{X}$ will be a locally compact space (LCH). A set is said to be precompact if its closure is compact. A space is $\sigma-$compact if it is the countable union of compact sets.

Lemma 40. *$E \subset \mathbb{X}$ is closed* **iff** *$E \cap K$ is closed for any compact K.*

Proof. Exercise. □

Proposition 41. *For any x and any open set $\mathcal{O}$ containing x there is a precompact open set $\mathcal{O}' \ni x$ with $\overline{\mathcal{O}'} \subset \mathcal{O}$.*

Proof. Let $\mathcal{O}''$ be any precompact neighborhood of x. We can replace $\mathcal{O}$ with $\mathcal{O} \cap \mathcal{O}''$; thus, wlog, we assume $\mathcal{O}$ is precompact. Then $\partial\mathcal{O}$ and x are closed and Note 38 completes the proof. □

Proposition 42. *Let K be compact and $\mathcal{O} \supset K$ open in $\mathbb{X}$. Then there exists a precompact $\mathcal{O}'$ s.t. $K \subset \mathcal{O}' \subset \overline{\mathcal{O}'} \subset \mathcal{O}$.*

Proof. By Proposition 41 K can be covered with precompact open sets $\{\mathcal{O}_\alpha\}$ with closure in $\mathcal{O}$ and thus by a finite subset of them $\{\mathcal{O}_i\}_{i \leqslant n}$. □

Theorem 43 (Urysohn's Lemma, LCH version)**.** *Let $K \subset \mathcal{O}$ as in Proposition 42. Then there is an $f \in C([0,1], \mathbb{X})$ and a precompact $\mathcal{O}', \overline{\mathcal{O}'} \subset \mathcal{O}$ s.t. $f(K) = \{1\}$ and $f(\overline{\mathcal{O}'}^c) = \{0\}$.*

Proof. Straightforward application of Urysohn and of the previous results. □

Also with a similar proof we have the following

Theorem 44 (Tietze Extension Theorem)**.** *Let K be compact and $f \in C(K)$. Then there exists $g \in C(\mathbb{X})$ s.t. $g|_K = f$.*

Proposition 45. *If $\mathbb{X}$ is second countable, then $\mathbb{X}$ is $\sigma-$compact.*

Proof. Let $\mathcal{T} = \{\mathcal{O}_i\}_{i \in \mathbb{N}}$ be a countable base. Each $x \in \mathbb{X}$ has, by assumption, a precompact neighborhood $\mathcal{O}'_x$. Since $\mathcal{T}$ is a base, there is an $i(x)$ and an $\mathcal{O}_{i(x)} \subset \mathcal{O}'_x$ s.t. $x \in \mathcal{O}_{i(x)} \subset \mathcal{O}'_x$. Then, $\overline{\mathcal{O}_{i(x)}} \subset \overline{\mathcal{O}'_x}$ is compact and $\mathbb{X} = \bigcup_{i(x), x \in \mathbb{X}} \overline{\mathcal{O}_{i(x)}}$, a countable union since it is a subfamily of $\mathcal{T}$. □

Proposition 46. *If $\mathbb{X}$ is $\sigma-$ compact, then there is a countable family of precompact open sets $\{\mathcal{O}_n\}_{n \in \mathbb{N}}$ such that $\overline{\mathcal{O}}_n \subset \mathcal{O}_{n+1}$ for all n and $\mathbb{X} = \cup_{n \in \mathbb{N}} \mathcal{O}_n$.*

Proof. This is an easy inductive construction. Let $\mathbb{X} = \cup_{n \in \mathbb{N}} K_n$. Let $\mathcal{O}_1$ be a precompact neighborhood of K_1, and for general n, let $\mathcal{O}_n$ be a precompact neighborhood of $K_n \cup \overline{\mathcal{O}}_{n-1}$. □

Here $\mathbb{X}$ is always an LCH space, $C, K, \mathcal{O}$ are a closed, compact and open resp. sets in $\mathbb{X}$.

5.4 Partitions of unity

If $E \subset X$ is any topological space, a partition of unity on E is a collection of continuous functions $\{\rho_\alpha\}_{\alpha \in A}$ with values in $[0,1]$ with the property that

• for any x there is a neighborhood of x where only finitely many ρ_α are nonzero.

• $\sum_\alpha \rho_\alpha = 1$.

A partition is subordinate to an open cover $\mathcal{O}_\alpha$ if every $\rho_\alpha \prec \mathcal{O}_\alpha$ for any α.

Partitions have many uses in mathematics. An interesting application is defining integrals on manifolds (with respect to some form). One relies on coordinates to define the integral on a coordinate patch and then uses a partition of unity subordinate to a coordinate patch covering to extend the integral to the whole manifold.

Theorem 47. *Let K be compact. For any open cover $\{\mathcal{O}_j\}_{j\leqslant n}$ of K there exists a partition of unity on K, $\{\rho_j\}_{j\leqslant n}$ with $\rho_j \prec \mathcal{O}_j, j \leqslant n$.*

Proof. (adapted from Rudin) We first find $\mathcal{O}'_j \subset \mathcal{O}_j, j = 1, ..., n$ which are precompact and which still cover K. For each $x \in K$ x is in some $\mathcal{O}_j$, and there is an $\mathcal{O}_x \subset \mathcal{O}_j$ precompact containing it. Now K must be contained in a finite union of these sets, $\cup_{ji}\mathcal{O}_{ij}$. For each j we simply define $\mathcal{O}'_j = \cup_i \mathcal{O}_{ij}$.

Now, by Urysohn, construct for each j a $\overline{\mathcal{O}'_j} \prec g_j \prec \mathcal{O}_j$. Let

$$\rho_1 = g_1; \quad \rho_2 = g_2(1-g_1); \;\; ...; \quad \rho_n = g_n(1-g_{n-1})\cdots(1-g_1)$$

Clearly $\rho_j \prec \mathcal{O}_j$. By induction we check that

$$\rho_1 + \cdots + \rho_n = 1 - (1-g_1)\cdots(1-g_n)$$

Now, for $x \in K$ at least one g_i is 1, and thus the sum above is 1 on K. □

5.5 Continuous functions

We denote by $\mathbb{C}^{\mathbb{X}}$ the space of all functions from $\mathbb{X}$ to $\mathbb{C}$.

Definition 48. • *The space $C_c(\mathbb{X})$ of functions with compact support is $\{f \in C(\mathbb{X}) : supp(f) \text{ is compact}\}$.*

• *f vanishes at infinity if $\left\{x : f(x) \geqslant n^{-1}\right\}_{n\in\mathbb{N}}$ are compact. The space of such functions is denoted by $C_0(\mathbb{X})$. Clearly $C_c(\mathbb{X}) \subset C_0(\mathbb{X})$. Show that $C_0(\mathbb{X}) \subset BC(\mathbb{X})$.*

The topology of uniform convergence is given by $f_n \to f$ if $\|f_n - f\|_u \to 0$ where $\|\cdot\|_u$ is the usual sup norm on $\mathbb{X}$.

Note however that it is **not** assumed that $\|f\|_u < \infty$ or $\|f_n\|_u < \infty$!, so this is not saying that we have a normed space!

In this section we will write $\|\cdot\|$ for the sup norm $\|\cdot\|_u$.

The topology of uniform convergence *on compact sets*, $f_n \to f$ if $\|f_n - f\|_K \to 0$ for all K compact where $\{\|\cdot\|_K\}_{K \text{ compact}}$ is the *family of seminorms* in which the supremum is taken over K. Obviously, $\|f\|_K < \infty$ for a continuous function.

Proposition 49. *$C(\mathbb{X})$ is a closed subspace of $\mathbb{C}^{\mathbb{X}}$ in both topologies above.*

Proof. Straightforward. □

Here $\mathbb{X}$ is always an LCH space, $C, K, \mathcal{O}$ are a closed, compact and open resp. sets in $\mathbb{X}$.

Note 50 (and exercises)**.** Show that the one-point compactification of $\mathbb{X}^*$ is Hausdorff compact.

The functions f in $C_0(\mathbb{X})$ extended by $f(\infty) = 0$ are the continuous functions on $\mathbb{X}^*$ which vanish at infinity. Thus we can identify the continuous functions on an LCH that vanish at infinity with the continuous functions on a CH that vanish at a point. This vanishing condition has interesting consequences.

Sup-norm convergence on $\mathbb{X}$ is stronger than uniform convergence on compact sets. The closure of $C_c(\mathbb{X})$ w.r.t the sup norm on on $\mathbb{X}$ is $C_0(\mathbb{X})$.

6 Radon measures

In order to better understand properties of various mathematical objects it is often very useful to analyze the natural functions (ones compatible with the structure) defined on them. These would be linear functionals on topological vector spaces, representations in the case abstract algebraic structures and in the case of topological spaces, the space of *continuous* functions defined on them (in fact specifying the continuous functions determines the topology). We can go one step further, look at continuous functions as a topological space (in the topologies mentioned in the previous section) and analyze its dual.

Two important subspaces of continuous functions on $\mathbb{X}$ are $C_c(\mathbb{X})$ and $C_0(\mathbb{X})$; we start with $C_c(\mathbb{X})$. As it turns out, *positive* linear functionals on $C_c(\mathbb{X})$ carry substantial information. In particular they generate measures on $\mathbb{X}$, which are then, by construction, compatible with the topology.

Definition 51. *Λ is a positive linear functional on if $\Lambda f \geqslant 0$ for any $f \geqslant 0$.*

Of course, this is the same as requiring

$$f \leqslant g \Rightarrow \Lambda f \leqslant \Lambda g$$

The following continuity property is automatic from positivity.

Proposition 52. *Let Λ be a positive linear functional on $C_c(\mathbb{X})$. There is a $C_K \geqslant 0$ such that*

$$|\Lambda f| \leqslant C_K \|f\|_K$$

where $\|f\|_K$ is the sup norm on K.

Proof. Since the positive and negative parts of a function are continuous, by linearity we can assume that f is real and nonnegative. Fix a ϕ, $K \prec \phi$. Then

$$f = \phi f \leqslant \|f\|_K \phi \Rightarrow 0 \leqslant (\Lambda \phi) \Lambda f$$

□

Obvious candidates for positive functionals on $C_c(\mathbb{X})$ are integrals with respect to positive Borel measures,

$$\Lambda f = \int_{\mathbb{X}} f d\mu \tag{2}$$

A relatively straightforward analysis shows that the measure needs to have additional properties, for instance

Proposition 53. *If (2) is a positive functional on $C_c(\mathbb{X})$, then $\mu(K) < \infty$ for any K.*

Proof. Formula (2) extends Λ to a positive functional on $L^1(\mu, \mathbb{X})$. Take any compact K and f s.t. $K \prec f$. Then $\chi_K \leqslant f$ and

$$0 \leqslant \mu(K) = \int_{\mathbb{X}} \chi_K d\mu \leqslant \int_{\mathbb{X}} f d\mu < \infty$$

□

We will see that, if $\mathbb{X}$ is second countable, the positive functionals are exactly these: (2) for some positive measure with μ finite on compact sets. If $\mathbb{X}$ is any LCH, the functionals are still given by (2) where the measure must be more regular, a *Radon measure*. (Regularity is automatic in second countable spaces.)

In fact, what a positive linear functional naturally generates is an outer Radon measure μ^*.

Definition 54. *An outer measure is Radon if*

1. *For any K, $\mu^*(K) < \infty$[3] (μ^* is locally finite).*
2. *any $\mathcal{O}$ is μ^*-measurable. Thus Borel sets are μ^*-measurable.*
3. *$\forall E \subset \mathbb{X}, \mu^*(E) = \inf\{\mu^*(\mathcal{O}) : \mathcal{O} \supset E\}$ (outer regularity)*
4. *$\forall \mathcal{O}$, $\mu^*(\mathcal{O}) = \sup\{\mu^*(K) : K \subset \mathcal{O}\}$ (inner regularity on open sets).*

By the Caratheodory theorem, μ^* restricted to the σ-algebra $\mathfrak{M}$ of μ^*-measurable sets

$$E \in \mathfrak{M} \Leftrightarrow \forall A \subset \mathbb{X},\ \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) \tag{3}$$

is a **measure** on $\mathfrak{M}$, $\mu = \mu^*$ on $\mathfrak{M}$ which we naturally call **Radon measure**.

A Radon measure is a measure which is outer regular, locally finite, and inner regular on open sets.

Check that the measure obtained from a Radon outer measure does indeed have the properties above. It also follows that

- μ is complete, i.e. $(E \subset E' \wedge \mu(E) = 0) \Rightarrow \mu(E') = 0$.

Lemma 55. *A Radon measure is inner regular on all measurable sets of finite measure, and more generally on all measurable σ-finite sets.*

Proof. Indeed, 1) assume first $m = \mu(E) < \infty$ and let $\mathcal{O} \supset E, \mu(\mathcal{O} \setminus E) < \epsilon/2$, $\mathcal{O}' \supset \mathcal{O} \setminus E$, $\mu(\mathcal{O}') < \epsilon$. Let $K \subset \mathcal{O}$, $\mu(K) > m - \epsilon$. Then $K' = K \cap (\mathcal{O}')^c \subset E$ is compact and $\mu(K') > m - 2\epsilon$.

2. Take now an E with $\mu(E) = \infty$. By assumption $E = \cup_{j \in \mathbb{N}} E_j$ where $\mu(\cup_{j \leqslant n} E_j) \to \infty$. By 1) above, there is a family $K_j \subset \cup_{j \in \mathbb{N}} E_j$ with $\mu(K_j) \to \infty$. □

Exercise 3 (When are sets outer-Radon measurable?). *If μ^* is a Radon outer measure, show that $E \subset \mathbb{X}$ is μ^*-measurable iff $E \cap K$ is measurable for every K.*

[3] In more general spaces one requires that every point has a neighborhood of finite measure; for LCH this is equivalent to the given condition.

Hint: Reduce the problem to measurability of all $E \cap \mathcal{O}$, $\mu(\mathcal{O}) < \infty$

Definition 56. *A measure μ is regular if it is inner and outer regular on all measurable sets.*

Note 57. *We see that a Radon measure can fail to be regular only when $\mathbb{X}$ is very large (e.g., not not second countable) and there only on sets of infinite measure which are not σ-finite.*

Theorem 58 (Riesz representation theorem)**.** *Let Λ be a positive linear functional on $\mathbb{X}$. Then, there exists a unique Radon measure on a σ-algebra $\mathfrak{M} \supset \mathcal{B}(\mathbb{X})$ s.t. (2) holds.*

Furthermore, for all $\mathcal{O}$

$$\mu(\mathcal{O}) = \sup_{f \prec \mathcal{O}} \Lambda f \tag{4}$$

and for all K

$$\mu(K) = \inf_{K \prec f} \Lambda f \tag{5}$$

Proof. **1. Uniqueness**

Assume we have two measures $\mu_{1,2}$ with the properties above. Using outer regularity and inner regularity on open sets, it is enough to show they coincide on compact sets. Let K be arbitrary and $\mathcal{O} \supset K$ be s.t. $\mu_2(\mathcal{O}) < \mu_2(K) + \epsilon$. Let $K \prec f \prec \mathcal{O}$; reasoning as in Proposition 53, we have

$$\mu_1(K) = \int_{\mathbb{X}} \chi_K d\mu_1 \leqslant \int_{\mathbb{X}} f d\mu_1 = \Lambda f = \int_{\mathbb{X}} f d\mu_2 \leqslant \int_{\mathbb{X}} \chi_{\mathcal{O}} d\mu_2 = \mu_2(\mathcal{O}) \leqslant \mu_2 + \epsilon$$

and interchanging $1 \leftrightarrow 2$ we have $|\mu_1(K) - \mu_2(K)| < \epsilon$. □

Construction of μ and $\mathfrak{M}$

It is then natural to define

$$\mu(\mathcal{O}) = \sup\{\Lambda f : f \prec \mathcal{O}\} \tag{6}$$

Now we note that

$$\mu(\mathcal{O}) \leqslant \sum_{j \in \mathbb{N}} \mu(\mathcal{O}_j) \text{ if } \mathcal{O} \subset \cup_{j \in \mathbb{N}} O_j \tag{7}$$

Indeed, for any $f \prec \mathcal{O}$, $K = \operatorname{supp} f \subset \cup_1^n \mathcal{O}_i$ for some n. With $\rho_i \prec \mathcal{O}_i$ a partition of unity, we see that $f = \sum_1^n f\rho_i$ and

$$\Lambda f = \sum_{i=1}^n \Lambda(f\rho_i) \leqslant \sum_{i=1}^n \mu(\mathcal{O}_i) \text{ since } f\rho_i \prec \mathcal{O}_i$$

The natural candidate for an outer measure is: for $E \subset \mathbb{X}$,

$$\mu^*(E) = \inf\{\mu(\mathcal{O}) : E \subset \mathcal{O}\} = \inf\{\sum_j \mu(\mathcal{O}_j) : E \subset \bigcup_j \mathcal{O}_j\} \text{ (by (7))}. \tag{8}$$

▶ By Prop. 1.10 (Folland), μ^* **is an outer measure** on $\mathbb{X}$.

◀

▶ Now we show that **open sets are μ^*-measurable**. For this we need to check (3) when $\mu^*(A) < \infty$.

▷ Furthermore, (3) holds if it holds when $A = \mathcal{O}'$, an arbitrary open set. Indeed, given A and $\mathcal{O}$, $\forall\epsilon$ there is an $\mathcal{O}' \supset A$ with $\mu^*(A) > \mu(\mathcal{O}') - \epsilon$, hence,

$$\mu^*(A) > \mu(\mathcal{O}') - \epsilon = \mu^*(\mathcal{O}' \cap \mathcal{O}) + \mu^*(\mathcal{O}' \cap \mathcal{O}^c) - \epsilon \geqslant \mu^*(A \cap \mathcal{O}) + \mu^*(A \cap \mathcal{O}^c) - \epsilon \tag{9}$$

◁

▷ When $A = \mathcal{O}'$ is an arbitrary open set of finite measure, we take $f \prec \mathcal{O}' \cap \mathcal{O}$ s.t. $\Lambda f > \mu(\mathcal{O}' \cap \mathcal{O}) - \epsilon$. Next, choose $g \prec \mathcal{O}' \setminus \operatorname{supp}(f)$ s.t. $\Lambda g > \mu(\mathcal{O}' \setminus \operatorname{supp}(f)) - \epsilon$. Noting that $f + g \prec \mathcal{O}'$, it follows that

$$\Lambda f + \Lambda g = \Lambda(f + g) \leqslant \mu(\mathcal{O}')$$

Hence,

$$\mu(\mathcal{O}') \geqslant \mu(\mathcal{O}' \cap \mathcal{O}) + \mu^*(\mathcal{O}' \setminus \mathcal{O}) - 2\epsilon \tag{10}$$

◀

▶ **μ satisfies (5)** Let $K \prec f$, $0 < \alpha < 1$ and $\mathcal{O}_\alpha = f^{-1}(\alpha, \infty)$. If $g \prec \mathcal{O}_\alpha$, then of course $g \leqslant \alpha^{-1} f$, implying

$$\forall\alpha \in (0,1),\ \mu(K) \leqslant \mu(\mathcal{O}_\alpha) = \sup_{g \prec \mathcal{O}_\alpha} \Lambda g \leqslant \alpha^{-1}\Lambda f \Rightarrow \mu(K) \leqslant \Lambda f$$

(in particular $\mu(K) < \infty$). In the opposite direction, we want to find an f, $K \prec f$ s.t. $\mu(K) > \Lambda f - \epsilon$. Let $\mathcal{O} \supset K$ be s.t. $\mu(\mathcal{O}) < \mu(K) + \epsilon$ and take $K \prec f \prec \mathcal{O}$. Then $\Lambda f \leqslant \mu(\mathcal{O}) < \mu(K) + \epsilon$ as desired.

◀

▶ **μ is inner regular on open sets**. Let $m < \mu(\mathcal{O})$; choose $f \prec \mathcal{O}$ s.t. $\Lambda f > m$, and let $K = \operatorname{supp}(f)$. For any set $\mathcal{O}' \supset K$ we have $f \prec \mathcal{O}'$, hence $\mu(\mathcal{O}') \geqslant \Lambda f$, and thus $\mu(K) \geqslant \Lambda f > m$, implying regularity.

In the process, we obtained an estimate for the constant in Proposition 53:

$$\operatorname{supp}(f) = K \Rightarrow \Lambda f \leqslant \mu(K) \tag{11}$$

◀

▶ $\boxed{\textbf{For any } f \prec \mathbb{X}, \Lambda f = \int_{\mathbb{X}} f d\mu}$ We can assume $f \in C_c(\mathbb{X}, [0,1])$. Let $K = \operatorname{supp}(f)$ and

$$y_1 < 0 < y_2 < \cdots < 1 < y_n \text{ be s.t } \forall i \max_i\{y_i - y_{i-1}\} < \epsilon \text{ and } \mu(f^{-1}(\{y_i\})) = 0$$

Let $\mathcal{O} \supset K$, $\mu(\mathcal{O}) < \infty$. Then $f^{-1}((y_{i-1} - y_i)) \cap \mathcal{O} := \mathcal{O}_i$ are open, mutually disjoint and $\mu(\mathcal{O} \setminus \cup_i \mathcal{O}_i) = 0$. For $i = 1, ..., n$ choose $K_i \subset \mathcal{O}_i$ so that $\mu(\mathcal{O}_i \setminus K_i) < \frac{\epsilon}{n}$ and g_i s.t. $K_i \prec g_i \prec \mathcal{O}_i$. If $\frac{\epsilon'}{n} := \mu(O_i) - \Lambda g_i$, then $\epsilon' < \epsilon$. By the mean value theorem, $\forall i\ \exists v_i \in [y_{i-1}, y_i]$ s.t

$$\int_{\mathbb{X}} f d\mu = \sum_i \int_{\mathcal{O}_i} f d\mu = \sum_{i=1}^n v_i \mu(\mathcal{O}_i) = \Lambda\left(\sum_{i=1}^n v_i g_i\right) - \epsilon' \Rightarrow \left|\int_{\mathbb{X}} f d\mu - \Lambda \sum_{i=1}^n g_i v_i\right| < \epsilon \tag{12}$$

Write $f - \sum_i v_i g_i = f_1 + f_2$ with $f_1 = \sum_i (f - v_i) g_i$ and $f_2 = f - f \sum_i g_i$. By (12), $\|f_1\|_u < \epsilon$, hence $|\Lambda f_1| < \epsilon$. Now $\|f_2\|_u \leqslant 1$ and $f_2 \prec \cup_i(\mathcal{O}_i \setminus K_i)$; hence, by (11), $|\Lambda f_2| < \epsilon$. The triangle inequality and (12) now give

$$\left|\int_{\mathbb{X}} f d\mu - \Lambda f\right| < 3\epsilon \qquad (\square \text{ of Thm. } 58)$$

A measure is regular if, by definition, it is inner and outer regular on all measurable sets (we only showed *inner regularity* on open sets).

Proposition 59. *Assume $\mathbb{X}$ is σ-compact. Let μ be a Radon measure and $\mathfrak{M}$ be the σ-algebra of μ-measurable sets. Then (a) For any $E \in \mathfrak{M}$ and $\epsilon > 0$ there is a pair $C \subset E \subset \mathcal{O}$ s.t. $\mu(\mathcal{O} \setminus F) < \epsilon$.*

(b) μ is a regular Borel measure.

(c) If $E \in \mathfrak{M}$, then there is a pair (F, G) of F_σ, G_δ sets s.t. $F \subset E \subset \mathcal{O}$ and $\mu(\mathcal{O} \setminus F) = 0$.

Proof. Let $X = \cup_n K_n$ where K_n are compact. CLet $E \in \mathfrak{M}$. Clearly, $\mu(E \cap K_n) < \infty$ and thus, by outer regularity, for any $\epsilon > 0$, there are $\mathcal{O}_n \supset E \cap K_n$ with $\mu(\mathcal{O}_n \setminus [E \cap K_n]) < \epsilon 2^{-n-1}$. With $\mathcal{O} = \cup_n \mathcal{O}_n$, we have $\mathcal{O} \setminus E \subset \cup_n(\mathcal{O}_n \setminus [E \cap K_n])$ and thus

$$\mu(\mathcal{O} \setminus E) < \epsilon/2$$

The same is true for E^c, and thus there is an open set $\mathcal{O}' \supset E$ s.t. $\mu(\mathcal{O}' \setminus E) < \epsilon/2$. If $C = (\mathcal{O}')^c$, then C is closed and $E \setminus C = E \cap \mathcal{O}' = \mathcal{O}' \setminus E^c$ implying the result.

Note that every closed set C is σ-compact, since $C \cap K_n$ is compact for every n and $C = \cup(C \cap K_n)$. But then, by the fact that μ is a measure, $\mu(C) = \lim_n \mu(\cup_{j=1}^n [C \cap K_j])$ proving inner regularity of closed sets, thus by (a), of all sets.

(c) Apply (a) with $\epsilon = j^{-1}, j \in \mathbb{N}$: there exist $C_j \subset E \subset \mathcal{O}_j$ s.t. $\mu(\mathcal{O}_j \setminus C_j) < \epsilon$. Now $F = \cup F_j \subset E \subset G = \cap \mathcal{O}_j$ and the result follows. $\square$

6.1 The Baire σ-algebra

Another natural σ-algebra when studying $C_c(\mathbb{X})$ is the smallest σ-algebra with respect to which the functions in C_c are measurable, $\mathcal{B}_0(\mathbb{X})$, whose elements are called Baire sets. Clearly $\mathcal{B}_0(\mathbb{X}) \subset \mathcal{B}(\mathbb{X})$; the two coincide if $\mathbb{X}$ is second countable (see Exercise 5/p. 216 in Folland).

6.2 Regularity of Borel measures

In this section we assume that $\mathbb{X}$ has the additional property that

$$\text{every } \mathcal{O} \subset \mathbb{X} \text{ is } \sigma\text{-compact} \tag{13}$$

This is the case if $\mathbb{X}$ is second countable.

Theorem 60. *Assume $\mathbb{X}$ satisfies (13). Then, every locally finite Borel measure λ on $\mathbb{X}$ is regular (and thus Radon).*

Proof. The functional $\Lambda f = \int_{\mathbb{X}} f d\lambda$ is well-defined on $\mathbb{C}(\mathbb{X})$ (since continuous functions are measurable, and $f = 0$ outside K implies $|f| \leqslant \|f\| \chi(K) \Rightarrow \Lambda|f| \leqslant \|f\|\lambda(K)$). Then, there is a regular Radon measure μ s.t.

$$\int_{\mathbb{X}} f d\lambda = \int_{\mathbb{X}} f d\mu$$

We now show that $\lambda = \mu$.

Take an $\mathcal{O}$, and let $\mathcal{O} = \cup_j K_j$, where K_j can be arranged to be increasing, as in Proposition 46. For each i, let $K_i \prec f_i \prec \mathcal{O}$. Clearly, $K_i \nearrow \mathcal{O}$. Now, since $\chi_{K_i} \leqslant f_i \leqslant \chi_O$ we have $f_i \to \chi(\mathcal{O})$; defining $g_i = \max_{j \leqslant i} f_i$, g_i are increasing to $\chi_{\mathcal{O}}$ and by the monotone convergence theorem,

$$\lambda(\mathcal{O}) = \lim_{i\to\infty} \int_{\mathbb{X}} g_i d\lambda = \lim_{i\to\infty} \int_{\mathbb{X}} g_i d\mu = \mu(\mathcal{O}) \tag{14}$$

Now, with $E \in \mathcal{B}(\mathbb{X})$ arbitrary, by the regularity of the measure μ, there is a pair $C \subset E \subset \mathcal{O}$ with $\epsilon > \mu(\mathcal{O} \setminus C) = \lambda(\mathcal{O} \setminus C)$ (the last equality by (14) and the fact that $\mathcal{O} \setminus C$ is open). If $\mu(\mathcal{O}) = \infty$ then $\mu(E) = \lambda(E) = \infty$. Otherwise,

$$\mu(\mathcal{O}) = \lambda(\mathcal{O}), \ \ |\mu(\mathcal{O}) - \mu(E)| < \epsilon \text{ and } |\lambda(\mathcal{O}) - \lambda(E)| \Rightarrow |\mu(E) - \lambda(E)| < 2\epsilon$$

□

Corollary 61. *Locally finite Borel measures on $\mathbb{R}^n$ are regular.*

Proposition 62. *If μ is a Radon measure on $\mathbb{X}$, then $C_c(\mathbb{X})$ is dense in $L^p, 1 \leqslant p < \infty$.*

Proof. Given the density of simple functions, it suffices to show that χ_E can be approached arbitrarily in p norm, when $\mu(E) < \infty$. Take then $K \subset E \subset \mathcal{O}$ with $\mu(\mathcal{O} \setminus K) < \epsilon$ and let $K \prec f \prec \mathcal{O}$. Then, f and χ_E can differ only on $\mathcal{O} \setminus K$, and the difference is at most one. This means that $\|f - \chi_E\|_p < \epsilon^{1/p}$. □

Exercise 4. *In preparation for the next section, show that $C_0(\mathbb{X})$ is a closed subspace of $BC(\mathbb{X})$ in the sup norm $\|\cdot\|$, and thus it is a Banach space, and that $C_c(\mathbb{X})$ is a dense subset in it.*

7 The dual of $C_0(\mathbb{X})$

Let's first determine what are the positive, continuous linear functionals on $C_0(\mathbb{X})$. Let Λ be such a functional; clearly its restriction to $C_c(\mathbb{X})$ is a positive linear functional and thus

$$\Lambda f = \int_{\mathbb{X}} f d\mu \tag{15}$$

where μ is a Radon measure, so the question is which of these extend to $C_0(\mathbb{X})$. Note first that if $f \prec \mathbb{X}$, then $\|f\| = 1$ and also that, of course, $\mathbb{X}$ itself is open. Thus, by (4), we have

$$\mu(\mathbb{X}) = \sup_{f \prec \mathbb{X}} \Lambda f \leqslant \|\Lambda\|\|f\| = \|\Lambda\| < \infty \tag{16}$$

Such a measure is called a finite Radon measure. Conversely, $\mu(\mathbb{X}) < \infty$ implies that Λ in (15) has norm at most $\mu(\mathbb{X})$. We find that

Proposition 63. *Λ is a positive linear functional on $C_0(\mathbb{X})$ (automatically continuous by the arguments above)* **iff** *it is given by (15) for a positive Radon measure μ, where $\mu(\mathbb{X}) < \infty$.*

We now turn to general, complex, continuous linear functionals on $C_0(\mathbb{X})$, that is, we want to find $C_0^*(\mathbb{X})$. Since the real and imaginary part of a continuous linear functional are real-valued continuous linear functionals, it suffices to determine these. We will see that, by an appropriate decomposition of the functional, the real-valued continuous linear functionals are still of the form (4), for a signed measure μ s.t. $|\mu|(\mathbb{X}) < \infty$.

Definition 64. *A subset C of a vector space V is a reproducing positive cone if*

1. $x, y \in C$ and $a, b \geqslant 0$ imply $ax + by \in C$
2. $C \cap (-C) = 0$
3. $\forall z \in V \exists x, y \in C$ s.t. $z = x - y$

Check that $C_0^+ = C_0(\mathbb{X}, [0, \infty))$ is a reproducing positive cone in $C_0(\mathbb{X})$. For $f, g \in C_0^+$, $\max\{f, g\} \in C_0^+$, $\min\{f, g\} \in C_0^+$.

Lemma 65. *Let C be a reproducing positive cone. Any $L : C \to [0, \infty)$ s.t. $a, b \geqslant 0$ and $f, g \in C$ imply $L(ax + by) = aLx + bLy$ extends as a linear functional on V.*

Proof. Note that if x, x', y, y' are in C and $x - y = x' - y'$ then $Lx - Ly = Lx' - Ly'$ (apply L to $x + y' = x' + y$). If, for $z \in V$, we set $Lz = Lx - Ly$ where $z = x - y$, then L is well defined and linear, and extends L from C to V as it is easy to check. $\square$

Lemma 66. *If $\Lambda \in C_0^*(\mathbb{X})$ there exist positive functionals $\Lambda^\pm \in C_0^*(\mathbb{X})$ s.t. $\Lambda = \Lambda^+ - \Lambda^-$.*

Proof. As mentioned, we expect that Λ is given by an expression of the type (15) for some measure. Recall that in the Jordan decomposition of a measure μ we have $\mu^+(E) = \sup_{E' \subset E} \mu(E')$. This motivates the following construction.

Define

$$\Lambda^+ f = \sup_{g \in C_0^+, g \leqslant f} \Lambda g, \qquad \text{for } f \in C_0^+ \tag{17}$$

Check that $f \in C_0^+ \Rightarrow \Lambda^+ f \geqslant 0$. We show that

$$f, g \in C_0^+ \text{ and } a, b \geqslant 0 \Rightarrow \Lambda^+(af + bg) = a\Lambda^+ f + b\Lambda^+ g$$

The fact that $\Lambda^+(|a|f) = |a|\Lambda^+ f$ follows from (17). It remains to check that $\Lambda^+(f_1 + f_2) = \Lambda^+ f_1 + \Lambda^+ f_2$ on C_0^+. This is straightforward noting that $g \leqslant f_1 + f_2$ in $C_0^+ \iff \exists g_1$ and g_2, in C_0^+, $g = g_1 + g_2$ and $g_i \leqslant f_i$. (In the $\Rightarrow$ direction, let $g_1 = \min\{g, f_1\}$ and $g_2 = g - g_1$.) Extend Λ^+ as Lemma 65. Now, $\Lambda^- := \Lambda^+ - \Lambda$ is evidently linear and positive, and thus Λ is the difference of two positive functionals.

Since $0 \leqslant g \leqslant f \Rightarrow \|g\| \leqslant \|f\|$, it follows from the definition (17) that

$$|\Lambda^+ f| \leqslant \sup_{0 \leqslant g \leqslant f} |\Lambda g| \leqslant \|\Lambda\| \|f\| \Rightarrow \|\Lambda^+\| \leqslant \|\Lambda\|$$

Exercise 5. *Let μ be a finite signed Radon measure and $\Lambda f = \int_{\mathbb{X}} f d\mu$. Let $\mu = \mu^+ - \mu^-$ be the Hahn-Jordan decomposition of μ. Show that the linear functional Λ^+ obtained in Lemma 66 is given by $\Lambda^+ f = \int_{\mathbb{X}} f d\mu^+$.*

$\square$

Definition 67. *μ is a signed Radon measure if $\mu = \mu_1 - \mu_2$ and μ_1, μ_2 are Radon measures. μ is a complex Radon measure if μ is a complex measure (finite, in particular) and $\mu = \mu_1 + i\mu_2$ where μ_1, μ_2 are signed Radon measures.*

Corollary 68. $\Lambda \in C_0^*(\mathbb{X})$ **iff** $\Lambda f = \int_{\mathbb{X}} f d\mu$ *where μ is a complex Radon measure.*

7.1 More facts about complex measures

A complex measure on a σ-algebra $\mathfrak{M}$ is a function on $\mathfrak{M}$ s.t.

$$\mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \;\text{ if }\; E = \bigoplus_{i=1}^{\infty} E_i, \quad E_i \in \mathfrak{M}\ \forall i \tag{18}$$

where $\oplus$ stands for disjoint union. Since we are not dealing with positive numbers, and $\oplus$ is invariant under permutations, the series above must converge absolutely.

It is natural to seek an upper bound for such a μ, which would simply mean a positive measure λ on $\mathfrak{M}$ such that $\lambda(E) \geqslant |\mu(E)|$ for all measurable E. Since $\lambda(\oplus E_i) = \sum_i \lambda(E_i)$, we see that we must have

$$\lambda(E) \geqslant \sum_{i=1}^{\infty} |\mu(E_i)| \;\text{ if }\; E = \bigoplus_{i=1}^{\infty} E_i, \quad E_i \in \mathfrak{M}\ \forall i$$

If we define $|\mu|$ on $\mathfrak{M}$

$$|\mu| = \sup_{\oplus_i E_i = E} \sum_{i=1}^{\infty} |\mu(E_i)| \tag{19}$$

it turns out that $|\mu|$, *the total variation of* μ is a positive measure on $\mathfrak{M}$. See Rudin, Chapter 6, and also the supplementary material. Clearly, $\mu \ll |\mu|$.

Theorem 69 (Rudin, Thm 6.4)**.** *If μ is a complex measure on X, then*

$$|\mu|(X) < \infty$$

Theorem 70 (Consequence of Rudin, Thm 6.4)**.** *Let μ be a complex measure on a σ-algebra $\mathfrak{M}$ is X. Then*

$$d\mu = e^{i\theta(x)} d|\mu| \text{ for some measurable } \theta : X \to [-\pi, \pi)$$

Proposition 71. *$\|\mu\| = |\mu|(X)$ is a norm on the linear space $M(X)$ of complex Radon measures.*

Proof.

$$\|\mu + \nu\| = |\mu + \nu|(X) = \sup_{\oplus_i E_i = X} \sum_i |\mu(E_i) + \nu(E_i)| \leqslant \sup_{\oplus_i E_i = X} \sum_i |\mu(E_i)| + \sup_{\oplus_i E_i = X} |\nu(E_i)| = \|\mu\| + \|\nu\|$$

and the rest is straightforward. □

Lemma 72. *If μ is a complex Radon measure and $\Lambda = f \mapsto \int_{\mathbb{X}} f d\mu$, $\Lambda : C_0(\mathbb{X}) \to \mathbb{C}$, then*

$$\|\mu\| = \|\Lambda\|$$

Proof. In one direction, with $\|f\| = 1$, $|\Lambda f| \leqslant \int_{\mathbb{X}} |f| d|\mu| \leqslant |\mu|(\mathbb{X}) = \|\mu\|$. In the opposite direction, by Theorem 69 $d|\mu| = v d\mu$ with $|v| = 1$. Let K be s.t $|\mu|(\mathbb{X} \setminus K) = \epsilon$ and $K \prec f$. Then,

$$\|\mu\| = \int_{\mathbb{X}} d|\mu| = \int_K d|\mu| + \epsilon = \left| \int_K f v d\mu \right| + \epsilon \leqslant \int_{\mathbb{X}} |f v| d|\mu| + \epsilon \leqslant \|\Lambda\| + \epsilon$$

□

Theorem 73 (The Riesz representation theorem)**.** *The map $\mu \to \Lambda_\mu$ is an isometric isomorphism of $M(\mathbb{X})$ to $C_0^*(\mathbb{X})$.*

Proof. The bijection was shown in Corollary , and Lemma 72 completes the proof. $\square$

Locally finite Borel measure in $\mathbb{R}^n$ are Radon, as discussed. By Theorem 1.16 (Folland) μ is a locally finite Borel measure on $\mathbb{R}^+$ iff it is a Lebesgue-Stieltjes measure, that is it is given by $\mu((a,b]) = F(b) - F(a)$ for some right-continuous, increasing, bounded F. Thus $\Lambda \in C_0^*(\mathbb{R})$ iff $\Lambda f = \int_{\mathbb{R}} f dF$ for some $F = F_1 - F_2 + i(F_3 - F_4)$ with F_i as above.

Note that $C_0([0,1]) \subset L^2([0,1])$, and the continuous functionals on $L^2[0,1])$ are given by the Riesz representation theorem, $f \mapsto \Lambda_\phi f = \int_{[0,1]} f\overline{\phi} dm$ where m is the Lebesgue measure and $\phi \in L^2$. Now $L^2([0,1]) \subset L^1([0,1])$ (by Cauchy-Schwarz) and thus the subclass of continuous functionals on $C_c([0,1])$ that extend to L^2 are generated by a subclass of measures μ s.t. $d\mu = \overline{\phi} dm, \phi \in L^2$. We may view then dF as some form of generalization of the Radon-Nykodim derivative. We'll make more sense of all this in distribution theory.

Definition 74. *The weak* topology on $M(\mathbb{X})$ is called* **the vague** *topology. It means $\mu_n \to \mu$ if $\int f d\mu_n \to \int f d\mu$ for all f.*

Exercise 6. *(a) Is $X = \mathbb{N}$ with the discrete topology a LCH space?*
(b) What is $C_0(X)^*$, if X is as in a)?

8 Fourier series

Fourier [4] introduced "trigonometric series" (now known as Fourier series) for the purpose of providing the general solution of the heat equation, for which only special solutions were known before his "Mémoire sur la propagation de la chaleur dans les corps solides" (1807) and "Théorie analytique de la chaleur" (Analytical theory of heat) (1822). This was of course a fundamental result, but the theory of Fourier series could only be placed on a rigorous basis later, by Dirichlet (1829) and Riemann, having to wait for a better theory of functions and of integration. In one dimension, Fourier series are of the form

$$\sum_{k \in \mathbb{Z}} c_k e^{ikx}, \; c_k \in \mathbb{C}$$

under suitable conditions on the coefficients to ensure convergence in L^2, or L^∞, or in more regular spaces, such as $C^n([-\pi, \pi])$.

Let's first start with L^2 convergence. It is easy to check that $\{e_k\}_{k\in\mathbb{Z}} = \{\frac{1}{\sqrt{2\pi}} e^{ikx}\}_{k\in\mathbb{Z}}$ form an orthonormal system in $\mathcal{H} = L^2([-\pi, \pi])$. Thus, by Bessel's inequality, for any $f \in \mathcal{H}$

$$\sum_{-\infty}^{\infty} \langle f, e_k \rangle e_k$$

converges in L^2. We now show that the closure of the span of $\{e_k\}_{k\in\mathbb{Z}}$ contains all indicator functions of intervals, implying, by the density of simple functions, that the closure of the span is $\mathcal{H}$.

[4] Jean-Baptiste Joseph-Fourier (1768-1830)

Theorem 75 (The Riemann-Lebesgue Lemma, first iteration...). *Assume $f \in L^1([-\pi,\pi])$. Then,*

$$\lim_{|n|\to\infty} \int_{-\pi}^{\pi} f(s) e^{ins} ds = 0 \tag{20}$$

Proof. Take first $f \in L^2([-\pi,\pi])$. The integral above equals $c_n = \sqrt{2\pi}\langle f, e_n\rangle$. By Bessel's inequality, $\sum_{n\in\mathbb{Z}} |c_n|^2 < \infty$, in particular $c_n \to 0$ as $|n| \to \infty$. Since L^2 is dense in L^1, the result follows by an $\epsilon/3$ argument. $\square$

8.1 The Dirichlet kernel

Let $f \in L^2(I)$. The symmetric partial Fourier sums of f are given

$$S_n(x) = \sum_{k=-n}^{n} \langle f, e_k\rangle e_k = \int_{-\pi}^{\pi} f(s) \sum_{k=-n}^{n} \overline{e_k(s)} e_k(x) ds = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) D_n(x-s) ds \tag{21}$$

where the **Dirichlet kernel** D_n is given by

$$D_n(x) = \sum_{k=-n}^{n} e^{ikx} = \frac{e^{-nix} - e^{nix}}{e^{-ix} - e^{ix}} = \frac{e^{-(n+1/2)ix} - e^{(n+1/2)ix}}{e^{-ix/2} - e^{ix/2}} = \frac{\sin((n+1/2)x)}{\sin(x/2)} \tag{22}$$

The Dirichlet kernel plays (as expected!) an important role in Fourier analysis. We know from

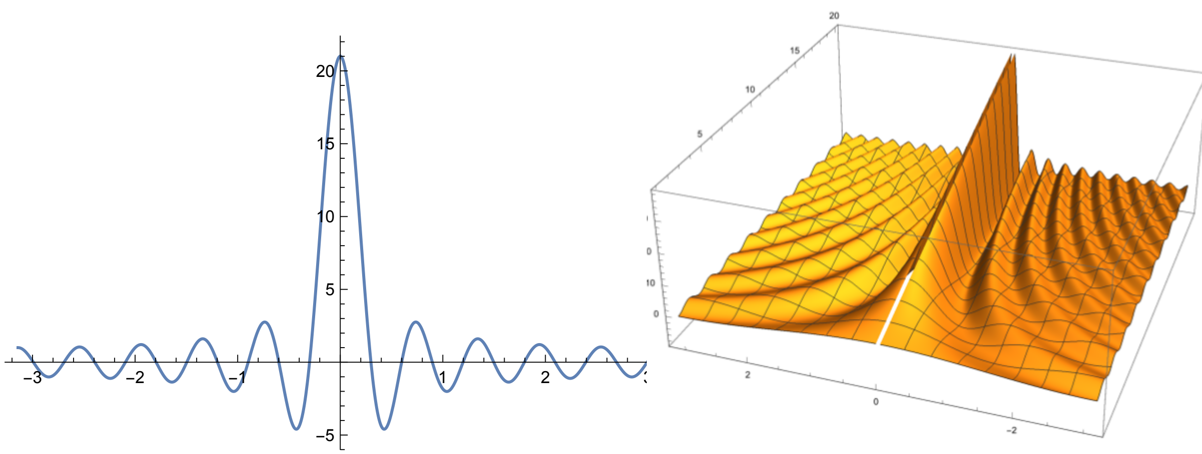

Figure 1: The Dirichlet kernel for $n = 10$, D_{10} (left) D_n on I for $n = 1, ..., 20$. The peak grows like n, with a width $1/n$ and oscillations of frequency n away from it (right)

Bessel's inequality that $S_n(x)$ converges **in the sense of** L^2, and as you saw already in a homework and will be reproved soon, converges to f.

Lemma 76. *Let $(a,b) \in [-\pi,\pi)$. Then*

$$\lim_{n\to\infty} \int_a^b D_n(x) dx = \begin{cases} 0 \text{ if } 0 \notin [a,b] \\ 2\pi \text{ if } 0 \in (a,b) \\ \pi \text{ if } 0 \in \{a,b\} \end{cases} \tag{23}$$

Proof. First, from the definition it follows that

$$\int_{-\pi}^{\pi} D_n(s)ds = 2\pi; \quad \text{and, since } D_n \text{ is even, } \int_0^{\pi} D_n(s)ds = \pi \tag{24}$$

Assume now $0 \notin [a,b]$. Then the function $g(x) = \chi_{(a,b)}(e^{-ix} - e^{ix})^{-1} \in L^2$ and (22) shows that $\int_a^b D_n(s)ds = 2\pi\langle g, e_{-n} - e_n\rangle \to 0$ as $n \to \infty$, by Theorem 75. If $0 \in (a,b)$, then

$$\int_a^b D_n(s)ds = \int_{-\pi}^{\pi} D_n(s)ds - \int_{-\pi}^{a} D_n(s)ds - \int_b^{\pi} D_n(s)ds = 2\pi + \epsilon_n$$

where $\epsilon_n \to 0$ as $n \to \infty$, since 0 is not in the last two intervals of integration. ☐

Proposition 77. *(a) If $(a,b) \in [-\pi,\pi]$, the Fourier series of $\chi_{(a,b)}$ converges in $\mathcal{H}$ to $\chi_{(a,b)}$. The Fourier series also converges pointwise to $\chi_{(a,b)}$ at any point of continuity and to the half sum of the right limit and left limit at the (at most two) points of discontinuity.*

(b) $\{e_k\}_{k\in\mathbb{Z}}$ form a Hilbert basis in L^2.

Proof. (a) Indeed,

$$\int_{-\pi}^{\pi} \chi_{(a,b)}(s)D_n(x-s)ds = \int_a^b D_n(x-s)ds = \int_{x-b}^{x-a} D_n(s)ds$$

and pointwise convergence follows from Lemma 76. Since the series converges in L^2 (by Bessel), the L^2 limit is the same as the pointwise limit, $\chi_{(a,b)}$ (justify this point).

(b) Staircase functions are in the closure of the span of $\{e_k\}_{k\in\mathbb{Z}}$ and are dense in L^2. ☐

We only proved convergence of symmetric sums to $\chi_{(a,b)}$. However, this combined with the fact that the Fourier series of $\chi_{(a,b)}$ is L^2 convergent implies that the Fourier series of $\chi_{(a,b)}$ converges in L^2 to $\chi_{(a,b)}$.

8.2 Pointwise convergence

Since trig polynomials are smooth functions, we might expect that the Fourier series of a continuous function converges pointwise. This is further suggested Proposition 77 (a). This however is not true. First, we note that trig polynomials are 2π-periodic, and we have to impose 2π-periodicity on the space of continuous functions. This is equivalent to identifying the endpoints of $[-\pi,\pi]$ upon which it becomes a circle, S^1. Continuous, periodic functions on $[-\pi,\pi]$ can be identified with $C(S^1)$.

Proposition 78. *For all $n \geqslant 1$, $\|D_n\|_1 \geqslant \frac{8}{\pi}\log n$.*

By modifying slightly the proof below, you can show that

$$\lim_{n\to\infty} \frac{\|D_n\|_1}{\log n} = \frac{8}{\pi}$$

Proof. Let $m = n + 1/2$. We have

$$\int_{-\pi}^{\pi} \left| \frac{\sin(mx)}{\sin(x/2)} \right| dx = 2 \int_0^{\pi} \frac{|\sin(mx)|}{\sin(\frac{x}{2})} \geqslant 4 \int_0^{\pi} \frac{|\sin(mx)|}{x} = 4 \int_0^{m\pi} \frac{|\sin x|}{x} dx$$
$$\geqslant 4 \sum_{k=0}^{m-1} \frac{(-1)^k}{k+1} \int_{k\pi}^{(k+1)\pi} \sin x dx = \frac{8}{\pi} \sum_{j=1} \frac{1}{j} \geqslant \frac{8}{\pi} (\log m + \gamma) \quad (25)$$

where γ is the Euler constant. □

What this shows is that, for any fixed a, the functionals $\Lambda_{a;N} = f \mapsto S_N(f;a)$ are not bounded in the Banach space $C(S^1)$. From this and the uniform boundedness principle we see that there is at least one continuous function for which the Fourier series diverges at some point. In fact, one can show that the family of continuous functions whose Fourier series converges at a given a is of first Baire category in $C(S^1)$.

Note 79. *It is a deep theorem (Carleson, 1966) that, for a fixed function in L^p, $p \in (1, \infty)$ (in particular, continuous), the set of points where the symmetric Fourier series converges pointwise is of full measure. In the opposite direction, for any set of zero measure there is a continuous function whose Fourier series diverges on that set.*

More regularity than continuity is needed for pointwise convergence. We start with a sufficient condition, definitely not optimal, but general enough, for now.

Proposition 80. *(a) If $f \in AC(S^1)$ and $f' \in L^2(S^1)$ (e.g. $f \in C^1(S^1)$), then*

$$\lim_{n\to\infty} \|S_n(f,x) - f(x)\|_u = 0$$

(b) The linear span of the $\{e_k\}_{k\in\mathbb{Z}}$ ("the trig polynomials") is dense in $C(S^1)$, in sup norm.

Proof. (a) Note first that, under these assumptions for f,

$$\int_{-\pi}^{\pi} f'(s) e^{-iks} ds = ik \int_{-\pi}^{\pi} f(s) e^{-iks} ds \Rightarrow S_n(f') = S_n(f)'$$

Uniform convergence follows from the fact that $\|f' - S(f_n)'\|_2 \to 0$ and $\|f - S_n\|_2 \to 0$ and Exercise 7 below.

(b) Periodic piecewise linear functions are dense in sup norm in the space of periodic continuous functions, and satisfy the assumptions in a). □

Note that density of trig polynomials in $C(S^1)$ does **not** imply that the Fourier **series** converge there. They generally don't, see discussion above.

Note 81. *Another way to prove (a) above is to use Parseval and Cauchy-Schwarz. Since $f' \in L^2$, we have $\sum_{k\in\mathbb{Z}} |kc_k|^2 < \infty$. This implies that*

$$\sum_{k\neq 0} |c_k| = \sum_{k\neq 0} |c_k| k(k^{-1}) \leqslant \left(\sum_{k\neq 0} |kc_k|^2 \right)^{1/2} \left(\sum_{k\neq 0} k^{-2} \right)^{1/2} < \infty$$

Thus the Fourier series converges absolutely, and then uniformly by the Weierstrass M test.

Exercise 7. *Assume $\{g_n\}_{n\in\mathbb{N}}$ is a sequence of functions in $AC([-\pi,\pi])$ with L^2 derivatives. Assume further that $\lim_{n\to\infty}\|g_n\|_2+\|g_n'\|_2=0$. Show that $\lim_{n\to\infty}\|g_n\|_\infty=0$.*

Exercise 8. *Show that when $f\in C^n([0,2\pi])$, $n\geqslant 1$, is 2π-periodic then the Fourier series of f, $\sum_{k=-\infty}^{\infty}c_ke^{ikx}$ converges to it uniformly together with the first $k-1$ derivatives, and the derivatives are given by $\sum_{k=-\infty}^{\infty}c_k(ik)^me^{ikx}$, $m\leqslant n-1$. In particular, the sequence $\{c_kk^m\}_{m\in\mathbb{Z}}$ is bounded.*

Show that if $f\in C^1$ except for a finite number of discontinuities where it has lateral limits, then the Fourier series of f converges pointwise everywhere to f except at the discontinuities, where it converges to the half sum of the lateral limits.

9 The heat equation

The heat equation is a *parabolic* partial differential equation that describes the variation in temperature in a given region Ω over time:

$$\frac{\partial u}{\partial t}=\Delta u,\quad u(0,x)=u_0(x),\ x\in\Omega;\ \ u(t,\cdot)_{|\partial\Omega}=f \tag{26}$$

where Δ is the Laplacian and the spacial variables run over some domain $\Omega\subset\mathbb{R}^n$. Here u_0 is the initial condition, the temperature distribution at $t=0$, and f is the boundary condition, the temperature distribution on $\partial\Omega$. The function u is assumed C^2 with continuous partial derivatives up to $\partial\Omega$.

Equilibrium distributions are time-independent solutions of (26), in the sense

$$\Delta u=0,\quad x\in\Omega;\ \ u_{|\partial\Omega}=f \tag{27}$$

Proposition 82 (Uniqueness)**.** *If u_1,u_2 solve (26) or (27), then $u_1=u_2$.*

Proof. If u_1,u_2 are solutions, then $u_1-u_2=v$ is a solution of the PDE with $u_0=0, f=0$. It suffices to show that the only such solution is zero. The proof is based on the *energy method*. Start with (26), $u_0=f=0$, multiply by v and integrate over Ω:

$$\int_\Omega v\frac{\partial v}{\partial t}dV=\frac{d}{dt}\int_\Omega v^2dV=\int_\Omega v\Delta v dV=\int_\Omega[\nabla\cdot(v\nabla v)-(\nabla v)^2]dV \tag{28}$$

where we used the identity $\nabla\cdot(v\nabla v)=(\nabla v)^2+v\Delta v$. Now, since $v=0$ on $\partial\Omega$ the divergence theorem implies

$$\int_\Omega\nabla\cdot(v\nabla v)dS=\int_{\partial\Omega}v\nabla v\cdot dS=0$$

and thus

$$\frac{d}{dt}\underbrace{\int_\Omega v^2dV}_{\geqslant 0}=-\int_\Omega(\nabla v)^2dV\leqslant 0 \tag{29}$$

Since $\int_\Omega v^2dV\geqslant 0$, is nonincreasing and vanishes at $t=0$, it means $\int_\Omega v^2dV=0$ and thus $v=0$ for all x,t. For (27), the left side of (29) is simply zero, giving $\nabla v=0\Rightarrow v=const=0$. □

If we find a solution to (26) or (30), we know it is the solution.

Let's analyze equilibrium distributions of (26) in two dimensions, with C^2 boundary condition in a disk. The equation becomes

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad u_{|_{S^1}} = f(\theta) \quad (f \in C^2) \tag{30}$$

This equation also describes the electric potential $u(x, y)$ in a disk where charges are placed on S^1 only, with a density f.

In polar coordinates we get

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0, \quad u_{|_{S^1}} = f(\theta) \tag{31}$$

A method of solving simple PDEs such as (31) is by separation of variables. Inserting $u(r, \theta) = R(r)T(\theta)$ in (31) and dividing by RT we get

$$r^2 \frac{R''}{R} + r\frac{R'}{R} = -\frac{T''}{T} \tag{32}$$

Now we note that the left side of the equation above does not depend on θ and the right side does dot depend on r, and thus they are independent of both variables, hence constant, say λ

$$r^2 \frac{R''}{R} + r\frac{R'}{R} = \lambda = -\frac{T''}{T} \tag{33}$$

The ODE $T'' = -\lambda T$ has the general solution $C_1 e^{i\sqrt{\lambda}t} + C_2 e^{-i\sqrt{\lambda}t}$. There are constraints on λ: T must be periodic of period 2π, and this means $\lambda = m^2, m \in \mathbb{Z}$, and then

$$T(\theta) = a_m e^{im\theta} + a_{-m} e^{-im\theta} \tag{34}$$

For the equation

$$r^2 R'' + rR' = \lambda R \tag{35}$$

we make the substitution $r = \ln x$, $R(r) = g(x)$ and get

$$g'' = m^2 g \Rightarrow g(x) = Ae^{mx} + Be^{-mx} \Rightarrow R(r) = Ar^m + Br^{-m} \tag{36}$$

if $m \neq 0$ and $R(r) = A + b \ln r$ for $m = 0$. We note that $\ln r$ and r^{-m} for $m > 0$ as well as r^m for $m < 0$ are not C^2. Retaining only the solutions that are C^2, we get the general separated-variables solutions

$$u_m(r, \theta) = r^{|m|} e^{im\theta}, m \in \mathbb{Z} \tag{37}$$

Now, (26) is linear, and thus if U and V are solutions, then so is $aU + bV$. The most general solution that we can obtain from (37) is the closure of the span of such solutions,

$$u(r, \theta) = \sum_{m \in \mathbb{Z}} a_m r^{|m|} e^{im\theta} \tag{38}$$

and with (38) we have at $r = 1$ (we'll check that the limit when $r \to 1$ exists),

$$\sum_{m \in \mathbb{Z}} a_m e^{im\theta} = f(\theta) \tag{39}$$

that is, the left side is the Fourier series of f. Since $f \in C^2$, $|a_m| \leqslant const/m^2$ for large m and, by the Weierstrass M test the series in (38) converges for all $r \leqslant 1$. The fact that u is C^2 for $r < 1$ follows from the general theory of power series. We have thus proved:

Theorem 83. *The heat equation in a disk (30) has a unique solution, (38).*

Exercise. *Separate variables in the time-dependent heat equation in a disk. The radial ODE has solutions as Bessel functions, $J_m(\lambda r)$; stop here if you are not familiar with them.*

9.1 Examples

(1) Take a disk where the temperature on the boundary is given by $f(\theta) = \sin\theta$. Then, the (unique) solution is simply $r\sin\theta = y$. (2) Similarly, for any trig polynomial, the series represent-

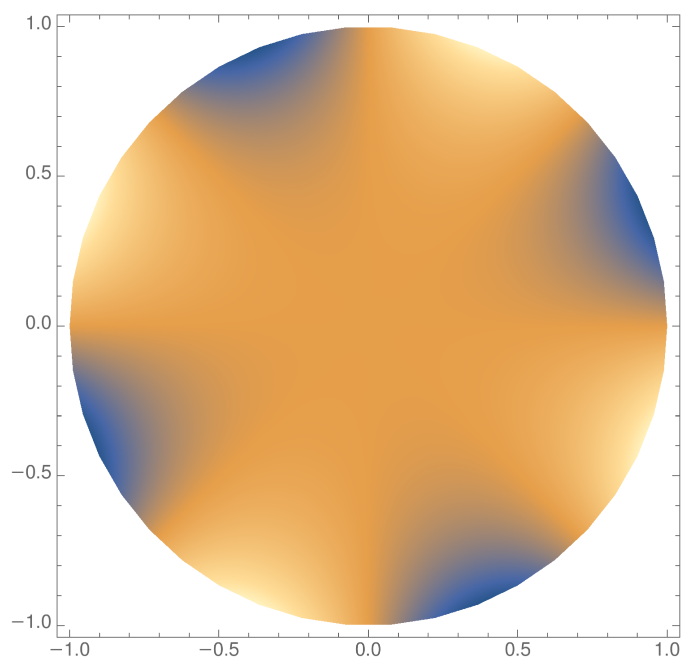

Figure 2: Solution of the heat equation in the disk with condition $\sin(4\theta)$ on S^1.

ing u is finite. It is interesting to see what happens if the temperature has many changes on the boundary, say $u = \sin(4\theta)$. Write the solution in closed form, as a function of x, y.

Exercise 9. *Show that the heat equation on S^1,*

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial \theta^2}; \quad u(0,x) = u_0(x) \in C^2(S^1) \tag{40}$$

has the unique solution

$$u(t,\theta) = \sum_{m\in\mathbb{Z}} a_m e^{-m^2t+imx} \text{ where } u_0(x) = \sum_{k\in\mathbb{Z}} a_m e^{imx} \tag{41}$$

In a few steps from here Fourier analysis intersects another major topic in analysis, complex function theory.

Lemma 84. *The Fourier coefficients of a real-valued function come in complex-conjugate pairs: $a_{-m} = \overline{a_m}$.*

Proof. Check this. □

Thus we can write

$$u(r,\theta) = 2\Re \sum_{m\geqslant 0} a_m r^m e^{im\theta} = 2\Re \sum_{m\geqslant 0} a_m (re^{i\theta})^m = 2\Re \sum_{m\geqslant 0} a_m z^m; \;\; z = x + i\sin y \tag{42}$$

where we wrote $re^{i\theta} = r\cos\theta + ir\sin\theta = x + iy = z$.

The series

$$S(z) = \sum_{m\geqslant 0} a_m z^m \tag{43}$$

converges absolutely and uniformly if $|z| < 1$ (check). A function $S(z)$, defined in an open connected region in $\mathbb{C}$ is said to be analytic if every point has an open disk around it where S has a convergent Taylor series. Our S is thus analytic in the open unit disk.

Under our assumption (30) the series converges absolutely and uniformly indeed, in the closed unit disk $|z| \leqslant 1$, and we have $f(\theta) = 2\Re S(e^{i\theta})$. Let's look again at the definition of the Fourier coefficients:

$$a_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-im\theta}d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} 2\Re S(e^{i\theta})e^{-im\theta}d\theta =^{(e^{i\theta}=\zeta)} \;\; = \frac{1}{2\pi i}\int_{S^1} 2\Re S(\zeta)\zeta^{-m-1}d\zeta \tag{44}$$

Substituting in (43) we get, for $|z| < 1$,

$$\Re S(z) = \sum_{m\geqslant 0} \frac{1}{2\pi i}\int_{S^1} \Re S(\zeta)\zeta^{-m-1}z^m d\zeta = \frac{1}{2\pi i}\int_{S^1} \Re S(z) \sum_{m\geqslant 0} z^m\zeta^{-m-1}d\zeta = \frac{1}{2\pi i}\int_{S^1} \frac{\Re S(\zeta)}{\zeta - z}d\zeta \tag{45}$$

A similar results holds with $\Re$ replaced by $\Im$. Indeed, $\Im S(re^{i\theta})$ satisfies the heat equation with boundary condition $\Im S(e^{i\theta})$. Adding up these two, we obtain the celebrated Cauchy formula

$$S(z) = \frac{1}{2\pi i}\int_{S^1} \frac{S(\zeta)}{\zeta - z}d\zeta, \;\; z \in \mathbb{D} \tag{46}$$

(for the unit disk, and under C^2 assumptions, which are too strong). This is simply meant to illustrate deeper links between various branches of analysis. It is not necessarily a particularly natural way to build complex analysis, nor is it the path that led Cauchy to it in the early nineteenth century.

Note 85. *(a) We did not* **not** *prove that the heat equation extended to $\mathbb{C}$ with a given complex boundary condition has a solution. It generally doesn't. See what the conditions are to have $\Re\sum_{m\geqslant 0} a_m z^m + i\Im\sum_{m\geqslant 0} b_m z^m = \sum_{m\geqslant 0} c_m z^m$.*

(b) Functions that are $\Re$ or $\Im$ of an analytic function are called harmonic *functions. They are the solutions in 2d of the steady-state heat equation.*

9.2 The vibrating string

The equation for a vibrating string is the one-dimensional *wave equation*

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0 \tag{47}$$

We change variables to place the *fixed* endpoints of the string at $-\pi, \pi$. Let the initial shape of the string be given by u_0. The problem becomes

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0; \;\; u(0,x) = u_0(x); \;\; u(t, \pm\pi) = 0 \tag{48}$$

Exercise 10. *Assume that $u_0(x)$ is C^2. Solve (48) by separation of variables and show that*

$$u(x,t) = \sum_{m=1}^{\infty} (a_m \cos(mt) + b_m \sin(mt)) \sin(mx);$$
$$u(0,x) = \sum_{m=1}^{\infty} a_m \sin(mx), \;\; u_t(0,x) = \sum_{m=1}^{\infty} m b_m \sin(mx) \tag{49}$$

Notice that the time dependence is a superposition of cosines of integer multiples of a fundamental frequency, generated by the *fundamental mode* $\sin x$. If we normalize again the units so that the fundamental mode is $440Hz$ (A 440) the next frequency is A 880, one octave up, and the third one is E 1320 *"a perfect fifth"*. The theory of harmony originates in the understanding of string vibrations, which goes back to ancient Greece (harmonikos = "skilled in music"). "Harmonic Analysis" takes its name from this.

9.3 The Poincaré-Wirtinger inequality

In full generality, this states

Proposition 86. *Let Ω be a bounded connected open subset of $\mathbb{R}^n$ with a Lipschitz boundary, and let $1 \leqslant p < \infty$. Then there exists a constant C, depending only on Ω and p such that for every function u in the Sobolev space $W^{1,p}\Omega$ we have*

$$\|u - u_\Omega\|_{L^p(\Omega)} \le C \|\nabla u\|_{L^p(\Omega)},$$

where

$$u_\Omega = \frac{1}{|\Omega|} \int_\Omega u(y)\, \mathrm{d}y$$

We will return to this form when we introduce Sobolev spaces. For now, let's look at this one dimensional version for functions in $C^1([a,b])$, $a < b \in \mathbb{R}$; $\|\cdot\|_2$ denotes the L^2 norm on $[a,b]$

Proposition 87. *If $f \in C^1([a,b])$ and $\int_a^b f = 0$, then*

$$\|f\|_2 \leqslant \frac{b-a}{\pi} \|f'\|_2$$

and the constant $(b-a)/\pi$ is optimal. Equality occurs iff $f(x) = A\sin(\pi(t-a)/(b-a))$.

Proof. The key element is proving the inequality (and optimality) when $f \in C^1(S^1)$. Let the Fourier coefficients of f be $\{c_n\}_{n\in\mathbb{Z}}$. We have

$$\|f\|_2 = \sum_{n\in\mathbb{Z}} |c_n|^2 \leqslant \sum_{n\in\mathbb{Z}} |nc_n|^2 = \|f'\|_2^2$$

The rest is left for the exercise below. □

Exercise 11. *(a) Fourier series can be defined of course for functions of more general period T. If we are interested in functions f periodic on $a, a+T$, then $f(\omega x + \beta)$ is periodic on $[-\pi,\pi]$, if $\omega = 2\pi/T$ and $\beta = -\pi - a\omega$. Carry out the changes of variables and write the Fourier series of f in terms of the exponentials $\{e^{ik\omega x}\}_{k\in\mathbb{Z}}$.*

(b) If f is as in the statement, extend it to a function on $[a-T, a+T]$ which is odd with respect to a, and then apply (a) and the result in the proof above.

9.4 The Riemann-Lebesgue lemma

Proposition 88. *If $f \in L^1(\mathbb{R})$, then $x \mapsto \int_{\mathbb{R}} f(s)e^{ixs}ds \in C_0(\mathbb{R})$*

Proof. First, $|e^{i(x+\epsilon)s} - e^{ixs}||f(s)| \leqslant 2|f(s)|$ and continuity follows by dominated convergence. For the second part, note that staircase functions are dense in $L^1(\mathbb{R})$, and for a staircase function, the integral goes to zero like *const*$/|x|$, by explicit calculation. Let $f \in L^1$, take $\epsilon > 0$, choose a staircase function g so that $\|f-g\|_1 < \epsilon/2$ and an R large enough so that for all $x, |x| \geqslant R$ we have $|\int_{\mathbb{R}} g(s)e^{ixs}ds| < \epsilon/2$. The rest is just the triangle inequality. □

Exercise 12. *Let $a > 0$ and consider the function f given by $f(x) = x^{-a}\chi_{[1,\infty)}(x)$. Show that $F(k) = \int_{\mathbb{R}} e^{ikx} f(x)dx \in C_0(\mathbb{R})$ if $a > 1$, and $F \in C_0(\mathbb{R}\setminus\{0\})$ if $a \leqslant 1$. Show furthermore that for $a \in (0,1)$, $k^{1-a}F(k)$ is bounded for small k, and, when $a = 1$, $F(k) + \ln k$ is bounded near $k = 0$. (Hint: integration by parts is one way, but it's probably simpler to change the variable to $u = kx$.)*

We have the following extension to Proposition 88:

Proposition 89. *If $a_1, a_2 > 0$, $C_1, C_2 \in \mathbb{C}$ and $f - C_1x^{-a_1}\chi[1,\infty) - C_2(-x)^{-a_2}\chi(-\infty,-1] \in L^1(\mathbb{R})$, then $x \mapsto \int_{\mathbb{R}} f(s)e^{ixs}ds \to 0$ as $k \to \infty$.*

Proof. This is immediate from Proposition 88 and Exercise 12. □

Exercise 13. *Extend this result to $\mathbb{R}^n$: if $f \in L^1(\mathbb{R}^n)$, then $\mathbf{x} \mapsto \int_{\mathbb{R}^n} f(\mathbf{s})e^{i\mathbf{x}\cdot\mathbf{s}}d^n\mathbf{s} \in C_0(\mathbb{R}^n)$.*

9.5 Hurwitz's proof of the isoperimetric inequality

A curve is rectifiable iff the supremum of the perimeters of polygons built by joining finitely many points on the curve is finite. With the intuition that the shortest distance between two points is along a straight line, we see that a curve is rectifiable iff its total length is finite. It is easy to show that the supremum is finite iff there is a parametrization $(x(t), y(t)), t \in [-\pi,\pi]$ with x, y in $\mathrm{BV}([-\pi,\pi])$.

Theorem 90. *Assume Γ is a rectifiable simple closed curve in $\mathbb{R}$ of length 2π. Then the area of the interior of the curve is $\leqslant \pi$ and it equals π iff the curve is a circle.*

Hurwitz gave the first rigorous proof of this theorem in 1902. He used Fourier series along the lines of the proof below, in which we assume a slightly stronger condition, that x, y are in $\mathrm{AC}([-\pi, \pi])$.

Proof. The area of a domain $\mathcal{D}$ is

$$A = \iint_{\mathcal{D}} dxdy \tag{50}$$

where $dxdy$ is the Lebesgue measure in $\mathbb{R}^2$. Recall Green's theorem,

$$\int_{\Gamma} Ldx + Mdy = \iint_{\mathrm{int}(\Gamma)} \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y} \right) dxdy$$

Now, for the vector field $L = -y, M = x$, Green's theorem and (75) give

$$A = \frac{1}{2} \int_{\Gamma} xdy - ydx = \left| \frac{1}{2} \int_{-\pi}^{\pi} \left[x(s)y'(s) - y(s)x'(s) \right] ds \right| \tag{51}$$

The arclength measure is given by $d\ell = \sqrt{(x')^2 + (y')^2} ds$. Parameterizing by arclength $\ell := t$ instead of s we have $\sqrt{\frac{dx}{dt}^2 + \frac{dy}{dt}^2} = 1$, and thus

$$\frac{1}{2\pi} \int_{\Gamma} (x')^2 + (y')^2 dt = 1 \tag{52}$$

As the curve is closed, x and y are in $AC(S^1)$ with derivative in L^2, and we have

$$x = \sum_{m \in \mathbb{Z}} a_m e^{imt};\ x' = \sum_{m \in \mathbb{Z}} ima_m e^{imt};\ y = \sum_{m \in \mathbb{Z}} b_m e^{imt};\ y' = \sum_{m \in \mathbb{Z}} imb_m e^{imt};\ a_m = \overline{a_{-m}};\ b_m = \overline{b_{-m}} \tag{53}$$

Parseval and (52) give

$$\sum_{m \in \mathbb{Z}} m^2 (|a_m|^2 + |b_m|^2) = 1 \tag{54}$$

Using (51) and again Parseval (how?) we get $A = \pi \left| \sum_{m \in \mathbb{Z}} m(a_m \overline{b_m} - b_m \overline{a_m}) \right|$ and thus

$$\pi^{-1} A \leqslant 2 \sum_{m \in \mathbb{Z}} |m||a_m||b_m| \leqslant \sum_{m \in \mathbb{Z}} |m|(|a_m|^2 + |b_m|^2) \leqslant \sum_{m \in \mathbb{Z}} |m|^2 (|a_m|^2 + |b_m|^2) \leqslant 1 \tag{55}$$

with equality iff $a_m = b_m = 0$ for all $|m| \geqslant 1$, which you can check is equivalent to Γ being a circle. □

10 Some conditions for pointwise convergence

$C^{\alpha}(S^1)$ is the class of functions on S^1 which are Hölder continuous of exponent α: $f \in C^{\alpha}(S^1)$ if

$$C_{\alpha}(f) = \sup_{x \neq y \in S^1} \frac{|f(x) - f(y)|}{|x - y|^{\alpha}} < \infty$$

Theorem 91. *If $f \in C^\alpha(S^1)$, $\alpha \in (0,1]$, then there is a constant $C > 0$ s.t. $\|S_n - f\|_u \leqslant C\ln(n)n^{-\alpha}$, where C depends on α only.*

Proof. This proof shows that convergence is linked to the rapid oscillation of D_n, through $\sin(nx + x/2)$, which triggers, in this class of functions, substantial cancellations.

We can assume

$$\|f\|_\infty \leqslant 1$$

Let $m = n + 1/2$. By changes of variables, (21) can also be written as $S_n(x) = (2\pi)^{-1}\int_{-\pi}^{\pi} D_n(s)f(x-s)ds$, and thus

$$2\pi(S_n(x) - f(x)) = \int_{-\pi}^{\pi} D_n(s)(f(x-s) - f(x))ds = \int_{|s|\leqslant\epsilon} D_n(s)(f(x-s) - f(x))ds + \int_{|s|>\epsilon} D_n(s)(f(x-s) - f(x))ds \quad (56)$$

where ϵ will be chosen suitably small. We start with an estimate of the $|s| \leqslant \epsilon$ integral. For small ϵ, $\sin(s/2) > \frac{2}{3}$ and

$$\int_{|s|\leqslant\epsilon} |D_n(s)(f(x-s) - f(x))|ds \leqslant 3\int_{|s|\leqslant\epsilon}\left|\frac{f(x-s)-f(x)}{s}\right|ds \leqslant 3C_\alpha(f)\int_{|s|\leqslant\epsilon}|s|^{\alpha-1}ds \leqslant \frac{6C_\alpha(f)\epsilon^\alpha}{\alpha}$$

Cancellations are responsible for decay in the remaining region; we identify the cancellations and rewrite the integral so that these are singled out: we have $\sin(ms) = -\sin(m(s+\frac{\pi}{m}))$. Let $I_k = \{x : |x| \in [\epsilon + k\frac{\pi}{m}, \epsilon + (k+1)\frac{\pi}{m}]\}$, $k_1 \in \mathbb{N}$ be the largest j so that $\epsilon + (2j-1)\frac{\pi}{m} < \pi$ and

$$h(s,x) = \frac{f(x-s) - f(x)}{\sin(s/2)}$$

We get

$$\int_{|s|>\epsilon} D_n(s)(f(x-s) - f(x))ds = \int_{|s|>\epsilon} h(s,x)\sin(ms)ds = \sum_{k=0}^{k_1}\int_{I_k} h(s,x)\sin(ms)ds + \epsilon_m \quad (57)$$

where ϵ_m is the contribution of the endpoint intervals:

$$|\epsilon_m| \leqslant \left|\int_{|s|\in[\epsilon+(2k_1-1)\frac{\pi}{m},\pi]} h(s,x)\sin(ms)ds\right| \leqslant \frac{C\pi}{m}, \;\; C < 3$$

We combine successive integrals by shifting the variable by $\mp\pi/m$ in all odd-index intervals:

$$\sum_{k=0}^{k_1}\int_{I_k} h(s,x)\sin(ms)ds = \sum_{j=0}^{k_1-1}\int_{I_{2j}\cup I_{2j+1}} h(s,x)\sin(ms)ds = \sum_{j=0}^{k_1-1}\int_{I_{2j}}(h(s,x) - h(s\mp\tfrac{\pi}{m}))\sin(ms)ds$$

Now, in each interval I_{2j}, $\sin ms$ is positive and the oscillations have been removed. At this stage, we can take absolute values without significant loss in the estimates.

$$\left|\sum_{k=0}^{k_1}\int_{I_k} h(s,x)\sin(ms)ds\right| \leqslant \sum_{j=0}^{k_1-1}\int_{I_{2j}} |h(s,x)-h(s\mp \tfrac{\pi}{m})|ds \tag{58}$$

we note that, if $s\delta > 0$ and $|s| > \epsilon$, then $|h(s+\delta,x)-h(s)| \leqslant |s|^{-2+\alpha}|\delta| + 2C_\alpha(f)|s|^{-1}|\delta|^\alpha$, and, for some $a>0, b>0$ only depending on $C_\alpha(f)$ the right side of (58) is, up to irrelevant constants $c_1,...,c_4$, bounded by

$$c_1 m^{-1}\int_\epsilon^\pi s^{-2+\alpha}ds + c_2 m^{-\alpha}\int_\epsilon^\pi s^{-1}ds \leqslant \frac{c_3}{m\epsilon^{1-\alpha}} + \frac{c_4}{m^\alpha}|\log\epsilon|$$

We now choose ϵ to obtain a best estimate (up to constants). Choose $\epsilon^\alpha = m^{-1}\epsilon^{-1+\alpha}$ or $\epsilon = m^{-1}$, to get for some other irrelevant constants $c_5,...,c_7$,

$$2\pi|S_n(x)-f(x)| \leqslant c_5 m^{-\alpha}\log m \tag{59}$$

□

Exercise 14. *Use the same approach to show that the Fourier coefficients of a function $f \in C^\alpha(S^1)$ decay at least as fast as const.$|n|^{-\alpha}$ as $n\to\infty$.*

Exercise 15 (Cesàro summation)**.** *(a) Let $\{a_n\}_{n=1}^\infty$ be a sequence, and denote by s_k its partial sums, $s_k = a_1+\cdots+a_k = \sum_{n=1}^{k} a_n$. The sequence $\{a_n\}_{n=1}^\infty$ is called* Cesàro summable, with Cesàro sum *A if the series of arithmetic means converges to A:*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} s_k = A$$

What is the Cesàro sum of $1-1+1-1\cdots$?

Exercise 16 (Abel means and Abel summability)**.** *If $\{a_n\}_{n\in\mathbb{Z}}$ is a sequence, then the* Abel mean *of the sequence is the* function

$$A(r,\theta) = \sum_{n=-\infty}^{\infty} r^{|n|}a_n e^{in\theta}$$

Note that, if a_n are the Fourier coefficients of a C^2 function f, then the Abel mean is the solution of the heat equation in the disk with f on the boundary! If the sequence is one-sided, that is indexed by $\mathbb{N}$, then one simply takes $a_m = 0$ for $m \leqslant 0$. The sequence is Abel summable *if*

$$\lim_{r\to 1} A(r,0) = A$$

exists. What is the Abel sum of $1-2+3-4\cdots$?

Show that (convergent to A)⇒ (Cesàro summable⇒ to A) ⇒(Abel summable to A).

We can think of these summation methods as extensions of convergent summation: extensions of the functional that associates to a convergent sequence its limit. These functionals have a number of expected properties, see. Both fail to commute with multiplication of sequences.

More powerful summation methods exist: Borel summation of series is an important summation method (it relies on a form of Fourier analysis!).

10.1 Approximation to the identity

The convolution of two functions on S^1 is defined as the *commutative and distributive* product

$$(f * g)(x) = \int_{-\pi}^{\pi} f(s)g(x-s)ds$$

Theorem 92 (Young's convolution inequality). *If $f \in L^1$, $g \in L^p$, and $1 \leqslant p \leqslant \infty$, then $\|f * g\|_p \leq \|f\|_1 \|g\|_p$.*

Proof. Use Minkowski's inequality for integrals. □

The difficulties in establishing convergence of Fourier series can be attributed to the divergence of the L^1 norm of the Dirichlet kernel. A *good kernel* is one which has most of the features of the Dirichlet kernel, but with finite L^1 norm.

Definition 93. *A family $\{K_n\}_{n\in\mathbb{N}} \subset L^1(S^1)$ is said to be an approximation to the identity (approximate identity) if*

(a) For all $n \geqslant 1$, with $\hat{K}_n := f \mapsto K_n * f$, we have

$$\int_{-\pi}^{\pi} K_n(s)ds = 1 \quad (\text{i.e. } \hat{K}_n 1 = 1) \tag{60}$$

(b)

$$\sup_{n\geqslant 1} \int_{-\pi}^{\pi} |K_n(s)|ds = M < \infty \quad ((\text{i.e. } \|\hat{K}_n\|_{L^\infty \to L^\infty} = 1) \tag{61}$$

(c) For any $\epsilon > 0$ we have

$$\lim_{n\to\infty} \int_{|x|\in[\epsilon,\pi]} |K_n(s)|ds = 0 \quad (\text{Approximate identity}) \tag{62}$$

For positive kernels, which are often encountered, (61) follows from (60).

Theorem 94. *(a) Let $\{K_n\}_{n\in\mathbb{N}}$ be an approximation to the identity family. Then, for any $f \in L^\infty(S^1)$ we have*

$$\lim_{n\to\infty} (K_n * f)(x) = f(x) \tag{63}$$

at any point where f is continuous. If $f \in C(S^1)$ then $\|K_n * f - f\|_\infty = 0$.

(b) If $f \in L^p(S^1), 1 \leqslant p < \infty$, then $\lim_{n\to\infty} \|\hat{K}_n f - f\|_p = 0$,

(c) For $1 \leqslant p \leqslant \infty$ we have $\sup_n \|\hat{K}_n\|_{p\to p} \leqslant M$. If $1 \leqslant p < \infty$, the sequence of operators $\{\hat{K}_n\}_{n\in\mathbb{N}}$ converges weakly to the identity.

Proof. The proof is similar to that of Theorem 91 (only simpler). Let x be a point of continuity of f. Given ϵ, let δ be s.t. $|f(x-s) - f(x)| \leqslant \epsilon$ if $|s| \leqslant \delta$. We decompose the integral $(K_n * f)(x) - f(x)$ as in (56),

$$\int_{-\pi}^{\pi} K_n(s)(f(x-s) - f(x))ds = \int_{|s|\leqslant\delta} K_n(s)(f(x-s) - f(x))ds + \int_{|s|>\delta} K_n(s)(f(x-s) - f(x))ds \tag{64}$$

We bound the first integral by using the sup norm for $f(x-s)-f(x)$ and the L^1 norm for K_n:

$$\left|\int_{|s|\leqslant\delta} K_n(s)(f(x-s)-f(x))ds\right| \leqslant \epsilon \int_{|s|\leqslant\delta} |K_n(s)|\,ds \leqslant \epsilon \tag{65}$$

and we use the assumptions on K_n in the second one

$$\left|\int_{|s|>\delta} K_n(s)(f(x-s)-f(x))ds\right| \leqslant \|f\|_\infty \int_{|s|>\delta} |K_n(s)|ds \to 0 \text{ as } n\to\infty \tag{66}$$

(b) Let $f\in L^p$ and let $g\in C(S^1)$ be s.t. $\|f-g\|_1<\epsilon$. Using Young's inequality for convolution, we see that $\sup_n \|K_n\|_{p\to p}\leqslant M$ and, for large enough n,

$$\|\hat{K}_n f - f\|_p \leqslant \|\hat{K}_n g - g\|_p + \|f-g\|_p + \|\hat{K}_n(g-f)\|_p \leqslant \|\hat{K}_n g - g\|_p + 2\|f-g\|_p \leqslant (2+M)\epsilon$$

(c) follows immediately from (b). □

Exercise 17. *Why isn't this proof working when $p=\infty$? Does the* **result** *extend to L^∞? (Leave this second part until the end of the next section.)*

11 The Fejér kernel

The Cesàro averages of S_n are

$$\sigma_n(f)(x) = \frac{1}{n}\sum_{k=0}^{n-1} S_n(f) = \frac{1}{n}\sum_{k=0}^{n-1} D_n * f = \left(\frac{1}{n}\sum_{k=0}^{n-1} D_n\right) * f =: F_n * f$$

Here F_n is the Fejér kernel,

$$F_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} \frac{\sin(nx+x/2)}{\sin(x/2)} = \frac{1}{n}\frac{\sin^2(nx/2)}{\sin^2(x/2)}$$

where the last equality is a simple exercise.

Lemma 95. *The Fejér kernel is an approximation to the identity.*

Proof. Since $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(s)ds = 1$ for all n, we have, for all n,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(s)ds = 1$$

and since F_n is a positive kernel, this proves (a) and (b) in the definition. It is clear that

$$\frac{1}{n}\int_{|s|>\epsilon} \frac{\sin^2(ns/2)}{\sin^2(s/2)}ds \leqslant \frac{1}{n}\int_{\pi\geqslant|s|>\epsilon} \frac{ds}{\sin^2(s/2)} \to 0 \text{ as } n\to\infty$$

□

We then find, as a corollary, the following.

Theorem 96. *If $f \in L^\infty(S^1)$, then $\{S_n(f)\}_{n\in\mathbb{N}}$ is Cesàro summable to f at any point of continuity of f. If $f \in C(S^1)$, then the Cesàro sums of $\{S_n(f)\}_{n\in\mathbb{N}}$ converge uniformly to f.*

Corollary 97. *If $f \in L^\infty(S^1)$ and its Fourier coefficients are all zero, then $f = 0$ at any point of continuity.*

Proof. This is immediate, since the Cesàro sums are zero. □

Corollary 98. *Trig polynomials are dense in $C(S^1)$ in $\|\cdot\|_\infty$.*

Proof. (We proved this already in Proposition 80, in a different way.) □

12 The Poisson kernel

Let $f \in L^\infty(S^1)$. Then, its Fourier coefficients $\{a_n\}_{n\in\mathbb{N}}$ are bounded, and thus the Abel means

$$A_r(f)(t) = \sum_{n\in\mathbb{Z}} a_n r^{|n|} e^{int} \tag{67}$$

converge absolutely and uniformly for $r < 1$, and we can interchange summation and integration to write

$$A_r(f)(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s) \sum_{n\in\mathbb{Z}} r^{|n|} e^{in(t-s)} ds = (P_r * f)(t) \tag{68}$$

where $P_r(t)$ is the Poisson kernel,

$$P_r(t) = \sum_{n\in\mathbb{Z}} r^{|n|} e^{int} = \frac{1-r^2}{1-2r\cos t + r^2} \tag{69}$$

as you can easily check.

Proposition 99. *P_r are an approximation to the identity* [5]

The fact that $\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r = 1$ (property (a)) follows from integrating the series term by term. For (b) we note that the kernels are positive. Property (c) follows from the fact that P_r are bounded and go to zero uniformly in any closed interval in $(0, \pi]$.

We have proved:

Theorem 100. *The Fourier series of an $L^\infty(S^1)$ function is Abel summable to f at any point of continuity of f. If $f \in C(S^1)$, then the series is uniformly Abel summable to f.*

Returning to the heat equation, we find that

Theorem 101. *The heat eq. (31) with f* **continuous**, *the uniform limit of $u(r,\theta)$ as $r \to 1$, has a unique solution (38) and (39).*

Corollary 102. *If $f \in L^1$ has zero Fourier coefficients then it is zero a.e. Thus the Fourier coefficients uniquely determine the function.*

Note again that this does not say that the Fourier series converges to the function in L^1! Kolmogorov showed in the 1920's that there exist L^1 functions for which the Fourier series diverges **everywhere.** See also Note 79 above.

[5] Indexed by the continuous variable $r \in [0, 1)$. The definition is virtually the same as in the discrete case.

Exercise 18. *Check that the map U, $U(f) = \{a_n\}_{n\in\mathbb{Z}}$, where $\{a_n\}_{n\in\mathbb{Z}}$ are the Fourier coefficients of f, is a unitary operator between $L^2([-\pi,\pi])$ and $\ell^2(\mathbb{Z})$. What is the image under U of the functions in $AC(S^1)$ with derivative in $L^2(S^1)$? (This is the domain of definition of the self-adjoint operator $i\frac{d}{dx}$ on S^1.)*

12.1 Several variables

Assume $f \in C^1((S^1)^2)$. Then,

$$f(x,\cdot) = \sum_{k\in\mathbb{Z}} c_k(x)e^{iky} \text{ where } c_k(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x,t)e^{-ikt}dt \tag{70}$$

Now, $c_k \in C^1(S^1)$ (why?), and hence

$$f(x,y) = \sum_{k\in\mathbb{Z}}\left(\sum_{m\in\mathbb{Z}} c_{k,m}e^{ikx}\right)e^{iky} \text{ where } c_{km} = \frac{1}{2\pi}\int_{-\pi}^{\pi} c_k(s)e^{-ims}ds = \iint_{[-\pi,\pi]^2} f(s,t)e^{-i(ms+kt)}dsdt \tag{71}$$

Uniform and absolute convergence (justify!) means that, we can write

$$f(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{Z}^2} c_{\mathbf{n}}e^{i\mathbf{n}\cdot\mathbf{x}} \tag{72}$$

Exercise 19. *(a) Under smoothness conditions as above, formulate and prove a theorem about Fourier series in n dimensions.*

(b) Write down a formula for the Fourier series of functions which are periodic, but have different periods in the different directions in $\mathbb{R}^n$.

The following exercise illustrated the duality between regularity (smoothness) and decay of the Fourier coefficients for functions that have point singularities. By the latter we mean that for each point at which the function is not smooth, there is an interval centered at that point in which there is no other point of non-smoothness.

Exercise 20. *Let*

$$f(x) = \sum_{n\in\mathbb{N}} \frac{\sin nx}{n^\alpha}; \quad (\alpha > 0, x \in [-\pi,\pi]) \tag{73}$$

(1) Show that (73) converges pointwise for all x.

(2) We now use a rudimentary form of Borel summation (see also more about this form of Borel summation) to determine the regularity of f. Using the definition of the Gamma function, show that

$$\frac{1}{n^\alpha} = \frac{1}{\Gamma(\alpha)}\int_0^\infty p^{\alpha-1}e^{-np}dp \tag{74}$$

Show that this implies that for $x \neq 0$ we have

$$f(x) = \frac{1}{\Gamma(\alpha)}\int_0^\infty p^{\alpha-1}\sum_{n\in\mathbb{N}}\sin(nx)e^{-np}dp = \frac{1}{\Gamma(\alpha)}\int_0^\infty p^{\alpha-1}\frac{\sin x}{2(\cosh p - \cos x)}dp \tag{75}$$

For $\alpha = 1$ the last integral is elementary,

$$2f(x) = \begin{cases} -x - \pi, \ x < 0 \\ 0, \ x = 0 \\ -x + \pi, \ x > 0 \end{cases}$$

Prove that $f(x)$ is C^∞ away from zero (actually, it is analytic).

(3) Take now $x > 0$ and small. Write (75) as

$$\frac{1}{x\Gamma(\alpha)} \int_0^\infty p^{\alpha-1} \frac{x^2 a(x)}{p^2 b(p) + x^2 c(x)} dp \tag{76}$$

and show that a, b and c are smooth in a neighborhood of zero, that $a(0) = b(0) = c(0) = 1$ and that $b(p) \geqslant 1$ for $p \geqslant 0$. With the change of variable $p = xq$ we get, for $x > 0$ small,

$$f(x) = \frac{1}{x^{\alpha-1}\Gamma(\alpha)} \int_0^\infty q^{\alpha-1} \frac{a(x)}{q^2 b(qx) + c(x)} dq \tag{77}$$

and that, as $x \to 0^+$ we have

$$\lim_{x\to 0} \int_0^\infty q^{\alpha-1} \frac{a(x)}{q^2 b(qx) + c(x)} dq = \int_0^\infty \frac{q^{\alpha-1}}{q^2+1} dq = \frac{\pi}{2\sin(\alpha\pi/2)} \tag{78}$$

(The last expression is most easily proved by the residue theorem, but you **don't need to justify it***; this explicit value is not terribly important here.) Use (78) to conclude that*

$$\lim_{x\to 0} \frac{f(x)}{\left(\frac{|x|^\alpha}{x}\right)} = \frac{\pi}{2\Gamma(\alpha)\sin(\alpha\pi/2)}; \ \ \alpha \in (0,1) \tag{79}$$

Thus f has precisely one point singularity, $x = 0$. Show that, for $\alpha \in (1,2)$, $f \in C^{\alpha-1}(S^1)$.

(d) Show that for $\alpha \in (0,1)$, $f \in L^p$ for any $p \in [1, 1/(1-\alpha))$. Are the Fourier coefficients of f those implied by the series?

Note 103. *It is useful to sketch this function for some $\alpha \in (0,1)$.*

This particular relation, $1/n^\alpha \mapsto (x - x_0)^{\alpha-1}$ between decay and regularity is generally true for point singularities. In the general class C^α, the (sharp) correspondence is $1/n^\alpha \leftrightarrow f \in C^\alpha$ with a proof similar to that of Theorem 91.

13 The Fourier transform

If f is not periodic, but compactly supported, we can extend it to a periodic function, with, say, the size of the support for period, and analyze it using Fourier series. Now if $f \in L^1(\mathbb{R}^n)$ is not periodic, it still makes sense to calculate the "coefficients",

$$(\mathcal{F}f)(k) = \hat{f}(k) = \int_{\mathbb{R}^n} e^{-2\pi i\langle x,k\rangle} dx, \ \ k \in \mathbb{R}^n \tag{80}$$

The function $\hat{f}$ is called the **Fourier transform** of f. The **inverse Fourier transform** (we'll shortly why "inverse") is

$$(\mathcal{F}^{-1}f)(k) = \check{f}(k) = \int_{\mathbb{R}^n} e^{2\pi i\langle x,k\rangle} dx, \quad k \in \mathbb{R}^n \tag{81}$$

Lemma 104. *The translation $\tau_a := f \mapsto f(x+a)$ is continuous in L^p, $1 \leqslant p < \infty$.*

Proof. Since $C_c(\mathbb{R}^n)$ is dense in $L^p, 1 \leqslant p < \infty$ it suffices to prove this for $C_c(\mathbb{R}^n)$. Let f be continuous and compactly supported in K. Translation is evidently linear, and thus it suffices to prove continuity at zero. We have

$$\lim_{a\to 0} \|\tau_a f - f\|_p \leqslant m(K)^{1/p} \| \lim_{a\to 0} \|\tau_a f - f\|_\infty = 0 \tag{82}$$

by uniform continuity. □

Finally, note that if f is 1-periodic, then the function $g = x \mapsto f((2\pi)^{-1}x)$ is 2π periodic, and thus we have

$$f((2\pi)^{-1}x) = \sum_{k\in\mathbb{Z}} e^{ikx} \frac{1}{2\pi} \int_0^{2\pi} f((2\pi)^{-1}s)e^{-iks} ds$$

which implies

$$f(x) = \sum_{k\in\mathbb{Z}} e^{2\pi ikx} \int_0^1 f(s)e^{-2\pi iks} ds \tag{83}$$

It is convenient to first analyze these transforms in a space of smooth, rapidly decreasing functions.

13.1 The Schwartz space $\mathcal{S}$

Let α, β be multiindices, that is tuples $(m_1, m_2, ..., m_n) \in \mathbb{N}_0^n$, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. We use the multidimensional conventions

$$|\alpha| = \sum_{i=1}^n \alpha_i, \quad x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}, \quad \partial^\alpha = \frac{\partial^{|\alpha|}}{(\partial x)^\alpha} \quad |x|^2 = \sum_{i=1}^n x_i^2; \text{ and } \binom{n}{\alpha} = \frac{n!}{\prod_{i=1}^n \alpha_i!}$$

The Schwartz space $\mathcal{S}$ of rapidly decreasing functions on $\mathbb{R}^n$ is defined as

$$S(\mathbb{R}^n) = \left\{ f \in C^\infty(\mathbb{R}^n) : \|f\|_{N,\beta} < \infty \quad \forall N \in \mathbb{N}_0, \beta \in \mathbb{N}_0^n \right\} \tag{84}$$

where

$$\|f\|_{N,\beta} = \sup_{x\in\mathbb{R}^n} \left| (1+|x|)^N \partial^\beta f(x) \right|. \tag{85}$$

These are smooth functions that decrease faster than any inverse power of x^2.

Proposition 105. *$\mathcal{S}$ is a Fréchet space.*

Proof. Only completeness needs to be checked. Since $C(\mathbb{R}^n)$ is complete, a Cauchy sequence $\{f_k\}_{k\in\mathbb{N}}$ in all $\|\cdot\|_{N,\beta}$ implies that $\{f_k\}_{k\in\mathbb{N}}$ convergences in all $\|\cdot\|_{N,\beta}$ to some functions $g_{N,\beta}$. It remains to identify these $g_{N,\beta}$, which is a simple exercise given the calculus theorem stating, in one dimension, that if $\{f_n'\}_{n\in\mathbb{N}}$ converge uniformly to some function h and $\{f_n(x_0)\}_{n\in\mathbb{N}}$ converges for some x_0, then $\{f_n\}_{n\in\mathbb{N}}$ converge uniformly on compact sets to some f and $h = f'$. □

Recall that a Fréchet space is a special case of a metrizable space, one in which the metric is translation invariant, $\rho(f,g) = \rho(f-g,0)$. If the family of semi-norms is $\|\cdot\|_n$, then a metric which induces the same topology is

$$\rho(f,0) = \sum_{n\in\mathbb{N}} 2^{-n}\frac{\|f\|_n}{1+\|f\|_n}$$

Conversely, a metric space is Fréchet if it is complete, *locally convex* (look up this notion) and the metric is translation-invariant.

Lemma 106. *The families of seminorms*

$$\left\{\||x|^\alpha\partial^\beta f\|_\infty\right\}_{N,\beta\in\mathbb{N}_0^n} \text{ and } \left\{\|(1+|x|)^N\partial^\beta f\|_\infty\right\}_{\alpha\in\mathbb{N}_0^n,N\in\mathbb{N}_0}; \tag{86}$$

induce the same topology on $\mathcal{S}$.

Proof. Indeed,

$$|x|^\alpha < (1+|x|)^{|\alpha|};\ \ (1+|x|)^N = \sum_{k=0}^N \binom{N}{k}|x|^k \leqslant \sum_{k=0}^N \binom{N}{k}\left(\sum_{i=1}^n |x_i|\right)^k = \sum_{\beta,|\beta|\leqslant N} a_\beta |x|^\beta$$

for some nonnegative coefficients a_β and thus the distance induced by the first family of seminorms goes to zero iff the distance induced by the second one does. □

Compactly supported smooth functions, $C_c^\infty(\mathbb{R}^n)$ are an important subset of $\mathcal{S}$. A prototypical such function is the function η below, compactly supported in the unit ball and smooth.

Proposition 107. *The function*

$$\eta(1-|x|^2) := \begin{cases} e^{-\frac{1}{1-|x|^2}};\ |x|<1 \\ 0;\ |x|\geqslant 1\end{cases} \tag{87}$$

is in $C_c^\infty(\mathbb{R}^n)$.

Proof. This follows from the chain rule and the fact that the function $t\mapsto e^{-1/t}\chi_{\mathbb{R}^+}(t)$ is in $C^\infty(\mathbb{R})$, see Exercise 3/p.239 in Folland. □

This function can be used as a building block to define other interesting compactly supported functions. For instance, the function

$$\phi(x) = \begin{cases} 1; |x|\leqslant 1 \\ \frac{\exp\left(\frac{1}{|x|^2-1}+\frac{1}{|x|^2-4}\right)}{1+\exp\left(\frac{1}{|x|^2-1}+\frac{1}{|x|^2-4}\right)};\ |x|\in(1,2) \\ 0; |x|\geqslant 2\end{cases} \tag{88}$$

is a smooth function, compactly supported in the ball of radius 2 and equals 1 in the closed ball of radius 1: $B_1(0)\prec\phi\prec B_2(0)^c$.

Proposition 108. *$C_c^\infty(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$*

Proof. Let $\phi_n = x \mapsto \phi(x/n)$ with ϕ as in (88). If $f \in \mathcal{S}(\mathbb{R}^n)$,then $\{f\phi_n\}_{n\in\mathbb{N}}$ is a sequence of compactly supported functions which, we claim, converges to f in the topology of $\mathcal{S}$. Indeed, we have

$$\partial^\alpha(f(x)\phi_n(x)) = \sum_{\beta\leqslant\alpha} n^{\beta-\alpha}\binom{\alpha}{\beta}\partial^\beta f(x)(\partial^{\alpha-\beta}\phi)(x/n)$$
$$= \partial^\alpha f(x) + \sum_{\beta\leqslant\alpha,\beta\neq\alpha} n^{\beta-\alpha}\binom{\alpha}{\beta}\partial^\beta f(x)(\partial^{\alpha-\beta}\phi)(x/n) - (1-\phi(x/n))\partial^\alpha f(x) \quad (89)$$

and note that, for any h, since $\sup_{x\in\mathbb{R}^n}|h(x)| = \sup_{x\in\mathbb{R}^n}|h(x/n)|$, we have

$$x^\gamma \sum_{\beta\leqslant\alpha,\beta\neq\alpha} n^{\beta-\alpha}\binom{\alpha}{\beta}\partial^\beta f(x)(\partial^{\alpha-\beta}\phi)(x/n) \to 0 \text{ as } n\to\infty$$

Finally, for any γ, if n is large enough, then $|\partial^\alpha f(x)| \leqslant C|x|^{-\gamma-1}$ if $|x| \geqslant n$, while $1-\phi(x/n) = 0$ if $|x| \leqslant n$ which means $x^\gamma(1-\phi(x/n))\partial^\alpha f(x) \to 0$ as well. ☐

Other important examples of functions in $\mathcal{S}$ are the Gaussians, or polynomials multiplying Gaussians,

$$x_i e^{-ax^2}, \quad (a>0)$$

Lemma 109. *The maps $\mathcal{F}$ and $\mathcal{F}^{-1}$ are continuous linear transformations from $\mathcal{S}$ into itself. Furthermore, $\mathcal{F}$ interchanges multiplication by the variable with differentiation, as follows:*

$$\mathcal{F}(\partial^\alpha x^\beta f) = (-1)^\beta(2\pi i)^{\alpha-\beta}k^\alpha\partial^\beta\mathcal{F}(f) \quad (90)$$

Proof. We have, by integration by parts,

$$\mathcal{F}(\partial^\alpha x^\beta f) = (-2\pi i k)^\alpha(-1)^\alpha\mathcal{F}(x^\beta f) = \frac{(-2\pi i k)^\alpha(-1)^\alpha}{(-2\pi i)^\beta}\partial^\beta\mathcal{F}(f) \quad (91)$$

Linearity is clear. Finally, we see that, up to a constant (that we can write down)

$$\|\hat{f}\|_{\alpha,\beta} = (2\pi)^{\alpha-\beta}\|\mathcal{F}(\partial^\alpha x^\beta f)\|_\infty \leqslant \sum_{\alpha'\leqslant\alpha,\beta'\leqslant|\beta|+n} c_{\alpha'\beta'}\|f\|_{\alpha',\beta'} \quad (92)$$

for some specific constants that can be determined by expanding out $\partial^\alpha x^\beta f$. ☐

Lemma 110 (Improper Riemann integrals and sums)**.** *Assume $f \in C(\mathbb{R}^n)$ and $\||x|^{n+1}f\|_u = M < \infty$. Then,*

$$\lim_{\epsilon\to 0}\epsilon^n\sum_{k\in\mathbb{Z}^n} f(\epsilon k) = \int_{\mathbb{R}^n} f(x)dx \quad (93)$$

Proof. Fix a $\delta > 0$ small enough and choose $R > \delta^{-1}$ so that for $|x| \geqslant R$ we have $|f(x)| < (M+1)/|x|^{n+1}$. Note that the number of points in the set $\{k \in \mathbb{Z}^n : m \leqslant \epsilon|k| < m+1\}$ cannot exceed the volume of the shell $\{x : \epsilon|x| \in [m-\epsilon, m+1+\epsilon]\}$, which is bounded by $C_n(m/\epsilon)^{n-1}$

for some positive C_n and thus, for small ϵ,

$$\left|\epsilon^n \sum_{k\in\mathbb{Z}^n} f(\epsilon k) - \epsilon^n \sum_{k\in\mathbb{Z}^n:\epsilon|k|<R} f(\epsilon k)\right| \leqslant C_n(M+1)\epsilon^n \sum_{m\geqslant R/\epsilon} \frac{\epsilon^2}{m^2} \leqslant C_n(M+1)\frac{\epsilon^{n+3}}{R} \leqslant \delta/2 \tag{94}$$

On the other hand, for some other constant C_n' we have

$$\left|\int_{|x|<R} f(x)dx - \int_{\mathbb{R}^n} f(x)dx\right| \leqslant \frac{(M+1)C_n'}{R} \leqslant const.\delta \tag{95}$$

Choosing ϵ small, we can arrange that

$$\left|\int_{\mathbb{R}^n} f(x)dx - \epsilon^n \sum_{k\in\mathbb{Z}^n} f(\epsilon k)\right| \leqslant \delta \tag{96}$$

since the sum above is simply a Riemann sum for the integral and f is smooth. □

Theorem 111 (Fourier inversion theorem in $\mathcal{S}$)**.** *(i) The Fourier transform is one to one from $\mathcal{S}(\mathbb{R}^n)$ onto itself and $\mathcal{S}(\mathbb{R}^n)$ $\mathcal{F}^{-1}\mathcal{F} = \mathcal{F}\mathcal{F}^{-1} = I$, the identity operator.*

(ii) (Plancherel) If $f \in \mathcal{S}(\mathbb{R}^n)$, then $\|f\|_2 = \|\mathcal{F}f\|_2$.

Proof. The fact that $\mathcal{F}$ is one-to-one onto follows from the inversion formula. Since $\mathcal{F}^{-1}\mathcal{F}$ is continuous, it suffices to show that $\mathcal{F}^{-1}\mathcal{F} = I$ on the dense set $C_c^\infty(\mathbb{R}^n)$. Take $f \in C_c^\infty(\mathbb{R}^n)$ and ϵ small enough so that f is supported in $K = [-\epsilon^{-1}/2, \epsilon^{-1}/2]^n$. Extending f as a periodic function on $\mathbb{R}^n$ with period ϵ^{-1} in any of the variables $x_1, ..., x_n$ we obtain

$$\begin{aligned} f(x) = \epsilon^n \sum_{k\in\mathbb{Z}^n} e^{2\pi i k\epsilon} \int_{-\frac{\epsilon^{-1}}{2}}^{\frac{\epsilon^{-1}}{2}} f(s)e^{-2\pi i k s\epsilon}ds = \epsilon^n \sum_{k\in\mathbb{Z}^n} e^{2\pi i k\epsilon} \int_{-\infty}^{\infty} f(s)e^{-2\pi i k s\epsilon}ds \\ = \epsilon^n \sum_{k\in\mathbb{Z}^n} e^{2\pi i k\epsilon}\hat{f}(k\epsilon) \to \mathcal{F}\mathcal{F}^{-1}f \end{aligned} \tag{97}$$

as $\epsilon \to 0$ by Lemma 110.

(ii) Similarly, it is enough to show this on $C_c^\infty(\mathbb{R}^n)$; if f is smooth and compactly supported, then,

$$\int_{\mathbb{R}^n} |f(s)|^2 ds = \int_K |f(s)|^2 ds = \epsilon^n \sum_{k\in\mathbb{Z}^n} |(\mathcal{F}f)(k\epsilon)|^2 \to \int_{\mathbb{R}^n} |(\mathcal{F}f)(k)|^2 dk \tag{98}$$

□

Corollary 112. *$\mathcal{F}$ extends to an isomorphism on $L^2(\mathbb{R}^n)$ with $\mathcal{F}^{-1}$ as its inverse.*

Proof. $\mathcal{S}^n$ is dense in $L^2(\mathbb{R}^n)$. □

Lemma 113. *$L^2(\mathbb{R}^n) \subset L^1(\mathbb{R}^n) \cap C_c^\infty(\mathbb{R}^n)$.*

Proof. Let $f \in L^1(\mathbb{R}^n) \cap C_c^\infty(\mathbb{R}^n)$ and K be the compact set outside which $|f| \leqslant 1$. Then,

$$\|f\|_2 = \int_K |f|^2 dm + \int_{K^c} |f|^2 dm \leqslant \int_K |f|^2 dm + \int_{K^c} |f| dm \leqslant \int_K |f|^2 dm + \|f\|_1 < \infty$$

□

Lemma 114 (A formula for $\mathcal{F}$ in L^2)**.** *If $f \in L^2$, then*

$$\lim_{R\to\infty} \left\| \hat{f} - \int_{|x|\leqslant R} e^{-2\pi iks} f(s)dm \right\|_2 = 0$$

If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then the extension of $\mathcal{F}$ to L^2 is the same as $\mathcal{F}$.

Proof. This is immediate since $f\chi_{|x|\leqslant R}$ converge to f both in L^1 and (by Parseval) in L^2. □

Note 115. *The result above is sometimes written*

$$\hat{f} = \operatorname*{l.i.m}_{R\to\infty} \int_{|x|\leqslant R} e^{-2\pi iks} f(s)dm$$

Note also that instead of balls we can take any sequence $\{K_n\}_{n\in\mathbb{N}}$ of compact sets s.t. $\lim_{n\to\infty} m(K^c) = 0$.

Theorem 116. *If $f \in L^1(\mathbb{R}^n)$ and $\mathcal{F}f \in L^1(\mathbb{R}^n)$, then $\mathcal{F}^{-1}\mathcal{F}f = f$.*

Proof. This follows immediately from Lemmas 113 and 114. □

Theorem 117 (Hausdorff-Young inequality)**.** *Assume $1 \leqslant p \leqslant 2$ and $p^{-1} + q^{-1} = 1$. Then, the Fourier transform is a bounded map from L^p to L^q with norm at most one. This is proved by interpolation, using Riesz-Thorin with $p_0 = q_0 = 2, p_1 = 1, q_1 = \infty$. We have $\|\hat{f}\|_2 = \|f\|_2$ and $\|\hat{f}\|_\infty \leqslant \|f\|_1$*

13.2 The Fourier inversion theorem, a direct approach

We show the inversion formula in $\mathbb{R}$. Let $f \in \mathcal{S}(\mathbb{R})$. Then, $\mathcal{F}^{-1}\mathcal{F}f$ equals

$$\lim_{R\to\infty} \int_{-R}^{R} e^{ikx} \int_{-\infty}^{\infty} e^{-iky} f(y)dydk = \lim_{R\to\infty} \int_{-\infty}^{\infty} f(x-u) \int_{-R}^{R} e^{iku}dk\,du = \lim_{R\to\infty} \left(\int_{-\infty}^{0} + \int_{0}^{\infty} \right) f(x-u)\frac{2\sin Ru}{u}du$$
$$= \lim_{R\to\infty} 2\int_0^\infty \frac{f(x+s)+f(x-s)-2f(x)}{s}\sin(Rs)ds + 4f(x)\int_0^\infty \frac{\sin s}{s}ds = 2\pi f(x)$$

In the last integral above we changed Ru to s, and the integral before it goes to zero by the Riemann-Lebesgue lemma Proposition 89 and the fact that the expression multiplying $\sin ku$ is smooth.

Note also the appearance in the process of the kernel $u^{-1}\sin(Ru)$, which has the same effect as the Dirichlet kernel, in concentrating the main contribution of the integral to a vanishing neighborhood of zero.

Proposition 118. *If $f(x) = e^{-\pi\alpha|x|^2}$ with $\Re(\alpha) > 0$, then $\hat{f}(k) = \alpha^{-n/2}e^{-\pi|k|^2/\alpha}$.*

Proof. In one dimension this follows from the fact that

$$\frac{d\hat{f}}{dk} = -\frac{2\pi}{\alpha}k\hat{f}$$

as it can be checked by integration by parts and that $\hat{f}(0) = \alpha^{-1/2}$. The extension to $\mathbb{R}^d$ is immediate, since the multiple integral is a product of one-dimensional integrals of the type above. □

14 Some applications of the Fourier transform

14.1 The Schrödinger equation for a free particle in $\mathbb{R}^d$

The wave function $\psi(x,t)$ of a particle has the following interpretation:$|\psi(x,t)|^2 dm$ is the probability density that, as a result of a measurement at time t, the particle will be found at position x. Then clearly we must have $\int_{\mathbb{R}^d} |\psi(x,t)|^2 dm(x) = 1$ for any t, in particular $\psi \in L^2(\mathbb{R}^d)$.

In the case of a single particle of mass m in an external potential $V(x,t)$, ψ satisfies the PDE

$$i\hbar \frac{\partial \psi}{\partial t} = -\tfrac{1}{2m}\Delta\psi + V(x,t)\psi$$

This is an evolution equation which requires an initial condition $\psi(x,t_0) = \psi_0(x)$. Here $E = -\frac{1}{2m}\Delta$ is the *kinetic energy operator* $E = \frac{p^2}{2m} =: \frac{1}{2m}\nabla^2$. In atomic units, $\hbar = 2m = 1$. A particle is free if the external potential is zero,

$$i\frac{\partial \psi}{\partial t} = -\Delta\psi$$

The Laplacian is a symmetric operator,

$$\underset{R\to\infty}{\text{l.i.m.}} \int_{|x|\leqslant R} (\psi\,\Delta\varphi - \varphi\,\Delta\psi)\; dV = \underset{R\to\infty}{\text{l.i.m.}} \oint_{|x|=R} \left(\psi\frac{\partial\varphi}{\partial\mathbf{n}} - \varphi\frac{\partial\psi}{\partial\mathbf{n}}\right) dS = 0$$

Lemma 119. *If $\|\psi(x,0)\|_2 = 1$, then $\|\psi(x,t)\|_2 = 1$ for all t.*

Such an evolution is called unitary, for obvious reasons.

Proof. By taking the complex conjugate of the Schrödinger equation,

$$-i\hbar\frac{\partial\overline{\psi}}{\partial t} = -\tfrac{1}{2m}\Delta\overline{\psi} + V(x,t)\overline{\psi}$$

Multiplying the first equation by $\overline{\psi}$, the second by ψ and subtracting, we get and subtracting the two equations, and integrating over $\mathbb{R}^d$ we get

$$i\frac{d}{dt}\int_{\mathbb{R}^d} |\psi|^2 dm = \underset{R\to\infty}{\text{l.i.m.}} \int_{|x|\leqslant R} \psi\,\Delta\overline{\psi} - \overline{\psi}\,\Delta\psi = 0$$

□

We now take the Fourier transform in $L^2(\mathbb{R}^d)$,

$$i\hat{\psi}' = 4\pi^2 k^2 \psi \Rightarrow \psi(x,t) = \widehat{\psi_0}(k) e^{-4\pi^2 i k^2 t}$$

The Fourier transform $\hat{\psi}$ is the probability amplitude of the momentum, k. We see that the probability distribuion in k is $|\widehat{\psi_0}|^2$, and it is independent of time. The momentum is conserved.

Now,

$$\psi(x,t) = \int_{\mathbb{R}^d} e^{-4\pi^2 i k^2 t + 2\pi i k x}\widehat{\psi_0}(k) dk$$

What happens when t becomes large? It is not difficult to see that the Riemann-Lebesgue lemma can be adapted to show that $\psi(x,t) \to 0$ as $t \to \infty$.

Concretely, let's assume that $\psi_0(x) = e^{-\pi\alpha x^2}$. Then, by Proposition 118, we have $\widehat{\psi_0}(k) = \alpha^{-d/2}e^{-\pi k^2/\alpha}$ and we get, using again Proposition 118,

$$\psi(x,t) = \alpha^{-d/2}\int_{\mathbb{R}^d} e^{2\pi ikx-\pi k^2(4\pi it+\alpha^{-1})}dk = (1+4i\pi\alpha t)^{-d/2}\exp\left(-\frac{\pi\alpha x^2}{16\pi^2\alpha^2t^2+1}+4i\frac{\pi^2\alpha^2 tx^2}{16\pi^2\alpha^2t^2+1}\right)$$

If $d=3$ we see that the probabilty of finding the particle in a ball of fixed radius decays roughly like t^{-3}, while the shape of the probability distribution is an ever widening Gaussian. The particle disperses out of any finite region.

14.2 The Airy equation

The Airy functions Ai and Bi satisfy the ODE

$$y'' = xy$$

The solutions are entire, since it is a linear ODE with entire coefficients. Taking the Fourier transform (with the normalization $\int_{\mathbb{R}} e^{-ikx}y(x)dx$) we get

$$-k^2\hat{y} = i\frac{d\hat{y}}{dk}$$

with the solution

$$\hat{y} = Ce^{ik^3/3}$$

meaning

$$y(x) = \int_{-\infty}^{\infty} e^{ik^3/3+ikx}dk$$

is (up to a multiplicative constant) one of the two linearly independent solutions of the ODE. With the normalization above, it is indeed, the Airy function $Ai(x)$. Or is it even a solution of the ODE? If we differentiate twice in x under the integral sign, we get an integral that does not converge, even conditionally.

But this does *not* mean that $y''(x)$ does not exist! It simply means that the representation is inadequate for this purpose. Instead, the contour of integration can be homotopically rotated:

$$y(x) = \int_{-\infty e^{-\pi i/6}}^{\infty e^{\pi i/6}} e^{ik^3/3+ikx}dk$$

In this way, whe $|k|$ is large, the integrand decreases roughly like $e^{-|k|^3/3}$, and $y(x)$ is now manifestly analytic in $\mathbb{C}$!

15 Convolution

Recall that the convolution of f and g is defined as

$$(f*g)(y) = \int_{\mathbb{R}^n} f(x)g(y-x)dx$$

Theorem 92 shows in particular that convolution is well defined on $L^1(\mathbb{R}^n) \times L^1(\mathbb{R}^n)$. The following theorem shows, in particular, that multiplication and convolution are Fourier-dual to each other.

Theorem 120. *Suppose $f, g \in L^1$. Then*

$$\widehat{f * g} = \hat{f}\hat{g};\ \widehat{fg} = \widehat{f * g}$$

and, if $a \in \mathbb{R}^n$, ten

$$\widehat{\tau_a f}(k) = e^{ika}\hat{f}(k)$$

Proof. This is a calculation, relying on Fubini:

$$(\widehat{f * g})(k) = \iint f(y-x)g(x)e^{-2\pi iky}dxdy = \iint f(y-x)g(x)e^{-2\pi ikx}e^{-2\pi ik(y-x)}dxdy$$
$$= \int f(y-x)e^{-2\pi ik(y-x)}d(y-x)\int g(x)e^{-2\pi ikx}dx = \hat{f}(k)\hat{g}(k) \quad (99)$$

The equality immediately following it is now obvious by the inversion formula. The last equality is clear from an immediate calculation. □

As a result, we should investigate further the properties of convolution.

Proposition 121. *Assuming that the integrals are well-defined (e.g., $f, g \in L^1$),*
a) $f * g = g * f$.
b) $(f * g) * h = f * (g * h)$.
c) For $a \in \mathbb{R}^n$, $\tau_a(f * g) = (\tau_a f) * g = f * (\tau_a g)$.
d) If $f, g \in \mathcal{S}$, then $f * g \in \mathcal{S}$.
e) If $A = \{x + y : x \in \operatorname{supp}(f), y \in \operatorname{supp}(g)\}$, then $\operatorname{supp}(f * g) \subset A$.

Proof. For e) note that for all x, if $z \notin A$ then $f(x)g(z-x) = 0$. The first three properties follow by density if they hold in L^1, where they are obvious from Theorem 120; d) is immediate from Theorem 120 and Lemma 109. □

Proposition 122. *Let p, q be conjugate exponents, $f \in L^p, g \in L^q$. Then $f * g$ exists pointwise everywhere, $f * g \in BC(\mathbb{R}^n)$ and $\|f * g\|_\infty \leqslant \|f\|_p\|g\|_q$. Furthermore, if $p \in (1, \infty)$, then $f * g \in C_0$.*

Proof. Pointwise existence and the uniform bound follow right away from Hölder's inequality. Noting that

$$\int f(x)g(y-x)dx = \int f(x)(Sg)(x)ds$$

where $S = \tau_y \circ J$, $(Jg)(x) = g(-x)$, continuity follows from Lemma 104. Finally, we note that $p \in (1, \infty)$ implies $q \in (1, \infty)$ and thus $C_c(\mathbb{R}^n)$ is dense in L^p and in L^q. By Proposition 121 e) C_c is preserved by convolution, and if $f_n \to f$ in L^p and $g_n \to g$ in L^q, then, by the first part of the Proposition, $f_n * g_n \to f * g$ uniformly. Since the uniform closure of C_c is C_0, the result follows. □

The following is a refinement of Proposition 121, d).

Proposition 123. *If $f \in L^1$ and $g \in C^k$ with $\partial^\alpha g \in BC$ for $|\alpha| \leqslant k$, then $f * g \in C^k$ and for all $\alpha, |\alpha \leqslant k$ we have $\partial^\alpha(f * g) = f * (\partial^\alpha g)$.*

Proof. This follows from the theorem of differentiation under the integral sign (T. 2.27, Folland). □

16 The Poisson summation formula

Theorem 124. *Assume $f \in C(\mathbb{R}^n)$, $\||x|^{n+\epsilon} f(x)\|_\infty < \infty$, and $\||k|^{n+\epsilon} \hat{f}(k)\|_\infty < \infty$ for some $\epsilon > 0$. Then,*

$$\sum_{k \in \mathbb{Z}^n} f(k) = \sum_{j \in \mathbb{Z}^n} \hat{f}(j)$$

and more generally,

$$\sum_{l \in \mathbb{Z}^n} f(x+l) = \sum_{j \in \mathbb{Z}^n} \hat{f}(j) e^{2\pi i j \cdot x} \tag{100}$$

The sum $(Pf)(x) := \sum_{k \in \mathbb{Z}^n} f(x+k)$ is called a *periodization* of f.

Proof. Note first that, under the given assumptions, the sums are uniformly and absolutely convergent. The function $\sum_{k \in \mathbb{Z}^n} f(x+k)$ is in $\mathbb{C}(\mathbb{T}^n) \subset L^2(\mathbb{T}^n)$. Note that

$$\begin{aligned}\hat{f}(j) = \int_{\mathbb{R}^n} e^{-2\pi i k \cdot x} f(x) dx &= \sum_{m \in \mathbb{Z}^n} \int_{\mathbb{T}+m} e^{-2\pi i k \cdot x} f(x) dx \\ &= \sum_{m \in \mathbb{Z}^n} \int_{\mathbb{T}} e^{-2\pi i k \cdot x} f(x+m) dx = \int_{\mathbb{T}} e^{-2\pi i k \cdot x} \sum_{m \in \mathbb{Z}^n} f(x+m) dx \end{aligned} \tag{101}$$

and (100) simply expresses convergence of the Fourier series of Pf to Pf. Uniform convergence implies pointwise convergence to Pf. The first equality, which is the most commonly used form of Poisson summation, follows from (100) by taking $x = 0$. □

This theorem has many applications, for instance in calculating sums in closed form, when the Fourier transform of a function is more easily summed than the function itself.

For instance, if $a \in \mathbb{R}^+$, we have

$$\widehat{\frac{1}{x^2 + a^2}} = a^{-1} e^{-a|k|}$$

which implies, using Poisson summation, that

$$\sum_{j \in \mathbb{Z}} \frac{1}{j^2 + a^2} = \pi a^{-1} \coth(a\pi) \tag{102}$$

(check!), which, by a limiting procedure (which?) contains the special case

$$\sum_{j \in \mathbb{N}} \frac{1}{j^2} = \frac{\pi^2}{6}$$

Eq. (102) is an instance of the Mittag-Leffler theorem, which expresses a meromorphic function by a "partial-fraction-like" expansion. In the same way we get

$$\sum_{j\in\mathbb{N}} \frac{1}{j^4+a^4} = \frac{\pi\left(\sinh\sqrt{2}\pi a + \sin\sqrt{2}\pi a\right)}{\sqrt{2}a^3\left(\cosh\sqrt{2}\pi a - \cos\sqrt{2}\pi a\right)}$$

implying (how?)

$$\sum_{j\in\mathbb{N}} \frac{1}{j^4} = \frac{\pi^4}{90}$$

Exercise 21. *Prove the duality formula*

$$\sum_{n\in\mathbb{Z}} e^{-\pi n^2 x} = x^{-1/2} \sum_{n\in\mathbb{Z}} e^{-\pi \frac{n^2}{x}}$$

Note: the function on the left side of the equation above is the Jacoby theta function, $\theta(x)$.

16.1 The Gibbs phenomenon

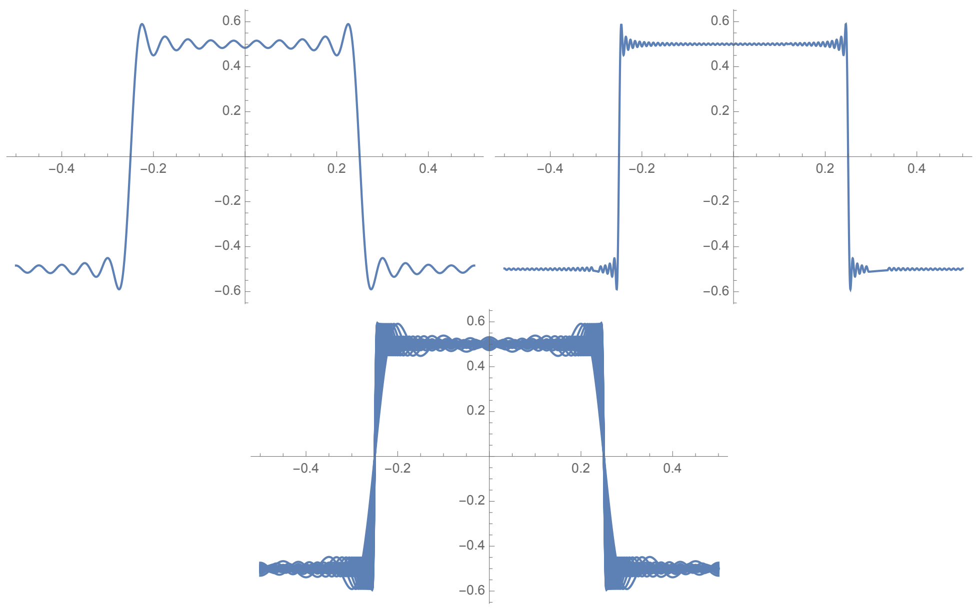

Figure 3: Characteristic function of $[-1/4, 1/4]$: the partial Fourier sum with 20 terms (left), 100 (right) and the graphical superposition of the Fourier sum with $20-100$ terms (below).

The Gibbs phenomenon is the remarkable way in which the Fourier series behaves at a jump discontinuity of a piecewise smooth function. The Gibbs phenomenon can be heard as "ringing" near transients, such as sounds from percussion instruments. It roughly results from the fact that we are trying to approximate a discontinuous function by smooth ones. Recalling the duality between smoothness and decay of the Fourier coefficients, a discontinuity will result in their slow decay. Therefore, the Fourier terms in the difference between a partial sum and the limit will have significant amplitude, resulting in fast oscillating defects. This "defect" only occurs in finite sums, since we know that in the limit the Fourier series converges everywhere to the average of the left and right limits of a piecewise-smooth function. This also means that the location of the maximum defect changes with thje number of terms, to allow for the limit to exist.

The Fourier sums of the function $f(x) = -1$ if $x \in (-1/2, 0)$ and 1 if $x \in (0, 1/2)$[6] is

$$S_N(x) = \sum_{k=0}^{N} \frac{4\sin(2\pi(2k+1)x)}{\pi(2k+1)}$$

The derivative of such a sum can be calculated explicitly,

$$S'_N(x) = 4\frac{\sin(4\pi(N+1)x)}{\sin(2\pi x)}$$

and an elementary argument shows that the first positive maximum of S_N occurs at $x_0 = \frac{1}{4(N+1)}$. We have

$$S_N(x_0) = \sum_{k=0}^{N} \frac{4\sin\left(\frac{2\pi(2k+1)}{4N+4}\right)}{\pi(2k+1)} \to \frac{1}{\pi}\int_0^{\pi} \frac{\sin x}{x} dx = \pi^{-1}\,\mathrm{Si}(\pi) = 1.1789797444\cdots \text{ as } N \to \infty$$

by recognizing that the sum is a Riemann sum for the integral (check!) We see that the sums converge nonuniformly to f, with an "overshot" of about 18% in uniform norm.

Exercise 22. *Show that the overshot by a factor of $\pi^{-1}\,\mathrm{Si}(\pi)$ of the Fourier sums occurs is the same at any jump discontinuiuty of a piecewise smooth function.*

17 Applications to PDEs

In this chapter we use ξ for the Fourier variable: this is the most frequent convention in PDEs.

17.1 The heat equation on the circle

This is given by

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad u(x,0) = f(x) \tag{103}$$

Assume first that f is smooth enough so that

$$f(x) = \sum_{j\in\mathbb{Z}} a_j e^{2\pi i j x}$$

By separation of variables (see (41)) we get

$$u(x,t) = \sum_{n\in\mathbb{Z}} a_n e^{-4\pi^2 n^2 t} e^{2\pi i n x} \tag{104}$$

[6]Note that the point values of the function at the discontinuity are irrelevant, as they wash out as a result of the integration involved in calculating Fourier coeffcients.

In discrete Fourier space, $\hat{u}$, the solution is given by the product of the Fourier coefficients of $u(x,0)$ and the Fourier coefficients of the **heat kernel for the circle**,

$$H_t(x) = \sum_{n\in\mathbb{Z}} e^{-4\pi n^2 t} e^{2\pi i n x} \tag{105}$$

and therefore, with convolution on the circle given by $(f * g)(y) = \int_0^1 f(x)g(y-x)dx$, we get (check!)

$$u = f * H_t \tag{106}$$

In terms of the Jacobi theta function,

$$\vartheta(z;\tau) = \sum_{n=-\infty}^{\infty} \exp\left(\pi i n^2 \tau + 2\pi i n z\right) = 1 + 2\sum_{n=1}^{\infty}\left(e^{\pi i \tau}\right)^{n^2}\cos(2\pi n z) = \sum_{n=-\infty}^{\infty} q^{n^2}\eta^n,\ \Re\tau > 0 \tag{107}$$

we have $H_t(x,t) = \vartheta(x;4\pi i t)$.

17.2 The heat equation on the line; smoothening by convolution

This is the same as (103), except with $x \in \mathbb{R}$.

Theorem 125 (Uniqueness)**.** *Assume u is continuous on $\overline{\mathbb{R}^+ \times \mathbb{R}}$, $u(t,\cdot) \in \mathcal{S}$ with bounded seminorms uniformly in $0 < t < T$, $u(x,0) = 0$, and that u satisfies* (103)*. Then $u = 0$.*

Proof. The proof is very similar to that in §9. The decay and smoothness of $\mathcal{S}$ are obviously too strong. □

Let's now build the solutiuon. The Fourier transform of (103) on the line reads

$$\frac{\partial \hat{u}}{\partial t} = -4\pi^2\xi^2\hat{u} \tag{108}$$

which is now an ODE, with ξ as a parameter. This gives,

$$\hat{u}(t,\xi) = \hat{f}(\xi)e^{-4\pi^2\xi^2 t} \tag{109}$$

The convolution theorem implies

$$u = f * \mathcal{H}_t \tag{110}$$

where

$$\mathcal{H}_t(x) = \mathcal{F}^{-1}\left(e^{-4\pi^2\xi^2 t}\right) = (4\pi t)^{-1/2} e^{-x^2/4t} \tag{111}$$

Theorem 126. *If $f \in \mathcal{S}$, then the solution of* (103) *on $\mathbb{R}$ is* (110)*, and $u(t,\cdot) \in \mathcal{S}$; $\|u(t,\cdot) - f\|_\infty \to 0$ and $\|u(t,\cdot) - f\|_2 \to 0$ as $t \to 0$*

Proof. First, note that $\hat{u}(t,\cdot) \in \mathcal{S}$ uniformly in t, implying that $u(t,\cdot) \in \mathcal{S}$ uniformly in t. Next,

$$|u(x,t) - f(x)| = \left|\int_{\mathbb{R}} \hat{f}(\xi)\left(e^{-4\pi^2\xi^2 t} - 1\right)e^{2\pi i \xi x} d\xi\right| \leqslant \int_{\mathbb{R}} |\hat{f}(\xi)| \left|e^{-4\pi^2\xi^2 t} - 1\right| d\xi \to 0 \tag{112}$$

by dominated convergence. For the L^2 norm, we use Plancherel:

$$\|u(t,\cdot) - f\|_2^2 = \|\hat{u}(t,\cdot) - \hat{f}\|^2 = \int_{\mathbb{R}} |\hat{f}(\xi)|^2 \left| e^{-4\pi^2\xi^2 t} - 1 \right|^2 d\xi \to 0 \tag{113}$$

again by dominated convergence. □

Corollary 127 (Smoothening by convolution). *Let $f \in C_c(\mathbb{R})$. Then $g_t = f * H_t \in \mathcal{S}$ (in fact, g_t is* entire*) and $g_t \to f$ uniformly as $t \to 0$.*

Proof. Indeed, if $f \in \mathbb{C}(\mathbb{R})$, then $\hat{f} \in C^\infty(\mathbb{R}) \cap C_0(\mathbb{R})$, hence $\hat{f}(\xi)e^{-4\pi^2\xi^2 t} \in \mathcal{S}$ (g_t is in fact entire because $\hat{f}(\xi)e^{-4\pi^2\xi^2 t}$ decays superexponentially). The rest follows from Theorem 126. □

Theorem 128. *The heat kernel on the circle is the periodization of the heat kernel on the line:*

$$H_t(x) = \sum_{n\in\mathbb{Z}} \mathcal{H}_t(x+n) \tag{114}$$

Proof. This follows immediately from (111), (105) and the general form of Poisson's summation formula. □

Corollary 129. *The heat kernel is positive, and the family $\{H_t\}_{t\geqslant 0}$ is an approximation to the identity.*

Proof. Positivity follows from (114). It is clear from (105) that $\int_{-1/2}^{1/2} H_t(x)dx = 1$. We have to show that the integral of H_t over an interval not containing 0, say $J = (\alpha, \beta)$ where $0 < \alpha < \beta < 1/2$ vanishes in the limit $t \to 0$. Note that for $x \in J$ and $0 \neq n \in \mathbb{Z}$ we have $|1 + x/n| \geqslant |1-\beta| := \epsilon$, implying $|x+n| \geqslant |n|\epsilon$ and thus

$$\sum_{|n|\geqslant 1} \mathcal{H}_t(x+n) \leqslant \sum_{|n|\geqslant 1} (4\pi t)^{-1/2} e^{-\epsilon^2 n^2/4t} \to 0 \text{ as } t \to 0 \tag{115}$$

by monotone convergence, which implies, by dominated convergence,

$$\int_J H_t(x)dx \leqslant \frac{e^{-\alpha^2/4t}}{(4\pi t)^{1/2}}(\beta - \alpha) + \int_J \sum_{|n|\geqslant 1} \mathcal{H}_t(x+n)dx \to 0 \text{ as } t \to 0 \tag{116}$$

□

Corollary 130. *For any* **continuous** *initial condition f, the heat equation on the circle has a unique smooth solution, $u(x,t) = (H_t * f)(x)$.*

Proof. Indeed, $H_t * f$ is smooth and solves the heat equation for any $t > 0$ and, by Corollary 129, $\lim_{t\to 0} H_t * f = f$. □

17.3 Linear PDEs

A differential operator is an operator L of degree m has the general form

$$L = \sum_{|\alpha|\leqslant m} a_\alpha(x) D^\alpha; \quad Lf = \sum_{|\alpha|\leqslant m} (2\pi i)^{-|\alpha|} a_\alpha(x) \partial^\alpha f$$

and it is with constant coefficients if $a_\alpha(x) = b_\alpha$ are independent of x. Let $f \in \mathcal{S}$. Then,

$$\mathcal{F}(Lf)(\xi) = \sum_{|\alpha| \leqslant m} b_\alpha \xi^\alpha \hat{f}(\xi) = P(\xi)\hat{f}(\xi) \tag{117}$$

The polynomial $P(\xi)$ is called **total symbol** of L, or simply **symbol**. The part of the polynomial containing the terms of highest degree only,

$$\sum_{|\alpha|=m} b_\alpha \xi^\alpha$$

is called **principal symbol**. For a second order partial differential operator L with principal symbol

$$\sum_{i+j=2} b_{ij} \xi_1^i \xi_2^j$$

the operator is called **elliptic** if the matrix $B = \{b_{ij}\}_{i,j}$ is positive or negative definite, **hyperbolic** if B is not definite but $\det(B) \neq 0$ and parabolic if exactly one eigenvalue of B is zero. Thus, Δ is elliptic, $\partial_x^2 - \partial_y^2$ is hyperbolic, and $\partial_t - \partial_x^2$ is hyperbolic. The names derive from the form of the symbol: for the Laplacian, the symbol is $-\xi_1^2 - \xi_2^2$ whose level lines are ellipses; the level lines are hyperbolas for $\xi_1^2 - \xi_2^2$; the heat equation has total symbol $\xi_1 + \xi_2^2$ whose level lines are parabolas; whether the parabola is concave or convex is also important. Let's examine these four types of equations on the circle, with conditions (initial, boundary, etc) in $\mathcal{S}$ performing (discrete) Fourier transform in one variable only.

For the wave equation, we get

$$[\hat{u}_{tt}]_j = -4\pi^2 j^2 \hat{u}_j$$

with solutions $u_j = a_j e^{-2\pi ijt} + b_j e^{2\pi ijt}$, meaning

$$u(x,t) = \sum_{j\in\mathbb{N}} a_j e^{-2\pi ij(x+t)} + \sum_{j\in\mathbb{N}} b_j e^{2\pi ij(x-t)}$$

and the solution is completely determined if we provide $u(x,0), u_t(x,0)$ (we note that $u(x,t) = f(x+t) + g(x-t)$, more about this in a moment). For the Laplacian, $u_{xx} + u_{yy} = 0$, once more taking the Fourier transform in y only, we get $\hat{u}_{xx} = \xi^2 \hat{u}$, meaning that, formally,

$$u(x,t) = \sum_{j\in\mathbb{N}} a_j e^{-2\pi j(ix+y)} + \sum_{j\in\mathbb{N}} b_j e^{2\pi ij(ix-y)}$$

Initial conditions in $\mathcal{S}$ mean that the coefficients a_k, b_k decrease faster than any power of ξ, but not necessarily exponentially fast. This means that the terms with $jy < 0$ in the first sum and $-jy > 0$ in the second must vanish, and we get, if the problem is formulated in the half plane $x \in \mathbb{R}, y \in \mathbb{R}^+$,

$$u(x,y) = \sum_{j\in\mathbb{N}} a_j e^{-2\pi(|j|y+ijx)}$$

which is completely determined by just one initial condition, $u(x,0) = f$. We also note that for $y > 0$, the solution is analytic in x, with some finite (in general) radius of analyticity. We saw that the heat equation solution is also determined by one initial condition, and the solution becomes entire at any $t > 0$. Finally, the backward heat equation would have the formal solution

$$\sum_{j\in\mathbb{Z}} a_k e^{4\pi^2 j^2 t + ijx}$$

and, for generic initial conditions in $\mathcal{S}$, this is nonsense for any $t > 0$ (the solution, assumed C^2 in x, would have a convergent Fourier series if it existed at all).

The principal symbol of a second order hyperbolic operator $\xi_1^2 - \xi_2^2$ factorizes over the reals. In the physical domain we have the factorization $\partial_t^2 - \partial_x^2 = (\partial_t - \partial_x)(\partial_t + \partial_x) = \partial_{\tau\eta}$ where $\tau = t - x, \eta = t + x$ which effectively reduces the PDE to ODEs:

$$\partial_{\tau\eta} u = 0 \Rightarrow \partial_\tau u = f(\tau) \Rightarrow u = F(\tau) + G(\eta) = F(t-x) + G(t+x)$$

The Poisson equation $\Delta u = 0$ is elliptic, and in a given domain it needs **one** boundary condition: either $u_{\partial\Omega} = f$ or the normal derivative $\frac{\partial u}{\partial n}|_{\partial\Omega} = g$.

Note the important role of the *principal symbol*: it's nature dictates the growth of the Fourier coefficients, which control the existence and smoothness of solutions.

17.4 Operators and symmetries

If $\mathcal{G}$ is a group of transformations on a space of functions, then L is invariant under $\mathcal{G}$ if $L(\gamma f) = \gamma L f$ for all $\gamma \in \mathcal{G}$. For instance if $\mathcal{G} = \mathcal{T}$ is the group of translations, $f(x) \mapsto f(x+\gamma), \gamma \in \mathbb{R}$, commutation means that $L(f(\cdot + a)) = (Lf)(\cdot + a)$.

Another way to write this is to note that $f \mapsto \gamma f$ is a linear operator; call it Γ. Then the operators L and Γ commute, $L\Gamma - \Gamma L =: [L, \Gamma] = 0$. Symmetries often place such restrictions on L that the operator is virtually determined by them. In physics, this is an important way to determine the fundamental laws of various theories.

Let's look at the question of which second order operators commute with the isometries of $\mathbb{R}^n$, the group generated by $\mathcal{T}$ and $O(n)$. Recalling our more general analysis of isometries of Hilbert spaces, all elements of $O(n)$ must be (real-valued) unitary transformations, $R \in O(n) \Rightarrow RR^* = I = RR^t$. In particular, $|\det R| = 1$.

Lemma 131. *The Fourier transform commutes with $O(n)$: $R \in O(n) \Rightarrow R(\mathcal{F}f(\xi)) = (\mathcal{F}f)(R\xi) = (\mathcal{F}f(R\cdot)(\xi)$.*

Proof. Changing variable $Rx = y$,

$$\widehat{f(Rx)} = \int_{\mathbb{R}^n} e^{-2\pi i\langle\xi,x\rangle} f(Rx)dx = \int_{\mathbb{R}^n} e^{-2\pi i\langle\xi,R^t y\rangle} f(y)dy = \int_{\mathbb{R}^n} e^{-2\pi i\langle R\xi,y\rangle} f(y)dy = \hat{f}(R\xi)$$

$\square$

Theorem 132. *A differential operator L commutes with the isometries of $\mathbb{R}^n$ iff it is a polynomial in Δ, $L = Q(\Delta)$.*

Proof. It is easy to see, as in the beginning of the paragraph, that L must have constant coefficients. In Fourier space it is a polynomial in $P(\xi)$ which, by Lemma 131, commutes with $O(n)$. We decompose the polynomial by homogeneous components,

$$P(\xi) = \sum_{m=0}^{M} \sum_{|\alpha|=n} a_\alpha \xi^\alpha = \sum_{m=0}^{M} P_m(\xi)$$

Next, we note that

$$0 = \lambda^{-M}[P(R\lambda\xi) - P(\lambda\xi)] \Rightarrow \lim_{\lambda\to\infty} \lambda^{-M}[P(\lambda R\xi) - P(\lambda\xi)] = P_M(R\xi) - P_M(\xi) = 0$$

This means that the highest order homogeneous polynomial is itself $O(n)$-invariant. Subtracting P_M from P and repeating the argument implies that P_{M-1} commutes with $O(n)$ and inductively, all homogeneous components $P_j(\xi)$ do. Take the unit sphere, $S = \{\xi : |\xi| = 1\}$ and note that $O(n)$ acts transitively on S. This follows from the exercise below. Thus $P_j(R\xi) = P_j(\xi)$ on S implies $P_j = a_j = const$ on S, entailing $P_j(\xi) = a_j|\xi|^j$ which is only possible if j is even, and thus $a_{2k+1} = 0$ and $P_{2k}(\xi) = a_{2j}(\xi_1^2 + ... + \xi_n^2)^j$. □

Exercise 23. *Show that $SL(n)$ acts transitively on $\mathbb{R}^n \setminus \{0\}$ and (thus) $O(n)$ acts transitively on the unit sphere in $\mathbb{R}^n$.*

17.5 Fourier transform of functions analytic in the lower half plane

Let $f \in L^1 \cap C_0(\mathbb{R})$ be s.t. $\hat{f} \in L^1$. Recall that this implies that $\mathcal{F}^{-1}\hat{f} = f$.

Proposition 133. *(i) Assume that $f \in L^1 \cap C_0(\mathbb{R})$ is s.t. $\hat{f} \in L^1$, and that f is analytic in the upper half plane $\mathcal{H}$, and that $|f(z)| \to 0$ as $|z| \to \infty$ in $\mathcal{H}$. Then $\hat{f}(\xi) = 0$ if $\xi < 0$.*

(ii) Assume $f \in L^1 \cap C_0(\mathbb{R})$ and $f(\xi) = 0$ for $\xi < 0$. Then $\check{f}$ is analytic in the upper half plane and $|f(z)| \to 0$ as $|z| \to \infty$ in $\mathcal{H}$.

Proof. (i) Let $\xi > 0$. Take C_r to be the three upper sides of a box in $\mathbb{C}$: the segment from r to $r - ir$, followed by the segment from $r - ir$ to $-r - ir$ and finally from $-r - ir$ to $-r$. Check that $\int_{C_r} e^{-i\xi x} f(x)dx \to 0$ as $r \to \infty$. Fix an ϵ and choose r large enough so that $|\int_{|x|>r} e^{-i\xi x} f(x)dx| + |\int_{C_r} e^{-i\xi x} f(x)dx| < \epsilon$. We then have

$$\left| \int_{\mathbb{R}} e^{-i\xi x} f(x)dx - \int_{[-r,r]\cup C_r} e^{-i\xi x} f(x)dx \right| < \epsilon$$

where $\int_{[-r,r]\cup C_r}$ means the integral over $[-r, r]$ followed by the integral on C_r discussed above. On the other hand, since f is analytic, Cauchy's theorem implies that $\int_{[-r,r]\cup C_r} e^{-i\xi x} f(x)dx = 0$, and since ϵ is arbitrary, the result follows.

(ii) Simply use dominated convergence and the Riemann-Lebesgue lemma.

□

17.6 The Laplace transform

Definition 134. *Let $F \in L^1(\mathbb{R}^+)$. The Laplace transform of $\mathcal{L}$ is defined as*

$$(\mathcal{L}F)(x) = \int_0^\infty e^{-px} F(p)dp, \quad \Re x > 0$$

More generally, if $e^{-ax}F \in L^1$ for some $a > 0$, then $\mathcal{L}F$ is defined by the same formula, for $\Re x > a$.

Theorem 135. *If $F \in L^1(\mathbb{R}^+)$, then $f(x) = \mathcal{L}F$ is analytic in the right half plane $\mathbb{H}$ and continuous in $\overline{\mathbb{H}}$. If $f(ix) \in L^1(\mathbb{R})$, then, for $p > 0$, F is given by the inverse Laplace transform,*

$$F(p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ipx} f(ix)dx = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} f(z)e^{pz}dz$$

If $\sup |z^a f(z)| < \infty$ *for some $a > 1$, then we equivalently have*

$$F(p) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f(z)e^{pz}dz$$

for any $c \geqslant 0$.

Proof. Analyticity in $\mathbb{H}$ follows from the fact that $F \in L^1$: dominated convergence allows then for differentiation inside the integral. Continuity in $\overline{\mathbb{H}}$ also follows from dominated convergence. In the limit $x \to 2\pi i t$, we get

$$(\mathcal{L}F)(2\pi i t) = \int_0^{\infty} e^{-2\pi i t p} F(p)dp = \hat{F}(t)$$

The rest follows from the Fourier inversion theorem. □

17.7 The adjoint operator

Recall that for a bounded operator B in a Hilbert space $\mathcal{H}$, we can define the adjoint B^* by $\langle Bx, y\rangle = \langle x, B^*y\rangle$, where uniqueness is immediate and existence is guaranteed by the Riesz representation theorem. An operator L which is not necessarily bounded is defined on some domain $\mathrm{dom}(L) = \Omega$ (we may assume that Ω is dense in $\mathcal{H}$, otherwise the natural Hilbert space to work in would be $\mathcal{H}_1 = \overline{\Omega}$). Naturally, the adjoint of L would be an operator L^*, defined on some domain Ω^* with the property

$$\forall (x, y) \in \Omega \times \Omega^*,\ \langle Lx, y\rangle = \langle x, L^*y\rangle$$

Obvious questions are of course existence of such an L^*, and uniqueness. Uniqueness is easy: if we have two operators L_1^* and L_2^* with the property above, then for any y such that $L_{1,2}^*$ are both defined, we have, for any x in the dense set $\mathrm{dom}(L)$,

$$\langle x, (L_1^* - L_2^*)y\rangle = 0 \Rightarrow (L_1^* - L_2^*)y = 0$$

For existence, define

$$\mathrm{dom}(L^*) = \{y \in \mathcal{H} : \exists z \in \mathcal{H}, \langle Lx, y\rangle = \langle x, z\rangle\} \tag{118}$$

and define L^* on $\mathrm{dom}(L^*)$ by

$$L^*y := z \tag{119}$$

Definition 136. *An operator A on a dense domain $\Omega \subset \mathcal{H}$ is self-adjoint if $A^* = A$. Note that this means that* $\mathrm{dom}(A^*)$ *is no more, and no less than* $\mathrm{dom}(A)$*.*

Proposition 137. *Let U be unitary from $\mathcal{H}$ to $\mathcal{H}'$ and $A : \Omega \to \mathcal{H}$ a linear operator with dense domain. Then $U\Omega$ is dense in $\mathcal{H}'$, UA is well defined on $U\Omega$ and its adjoint in UA^*.*

Proof. Since U is an isomorphism, this is a straightforward verification. □

Example 138. Consider the operator $D = i\frac{d}{dx}$ on S^1. First, we see that for smooth functions, say in $C^\infty(S^1)$, $\langle Df, g\rangle = \langle f, Dg\rangle$ and D^* exists at least on $C^\infty(S^1)$, and on it $D^* = D$. From the definition of the adjoint, it is clear that the domain of D^* gets larger if the domain of D shrinks. Suppose we want to determine first the "maximal" set of functions in $L^\infty \subset L^2$ on which we can define differentiation. We keep then $C^\infty(S^1)$ as a domain for D (or choose an even smoother space if it helps), and determine the corresponding domain of D^*.

Let $U = \mathcal{F}$, the discrete Fourier transform, a unitary map between $L^2(S^1)$ and $\ell^2(\mathbb{Z})$. Then UD is the operator of multiplication by $-k$ and, to understand what the adjoint of D is, it is enough to determine the adjoint of $-k$. We have

$$\langle Df, g\rangle = \sum_{k\in\mathbb{Z}}(-2\pi k f_k)\overline{g_k} = \sum_{k\in\mathbb{Z}} f_k\overline{(-2\pi k g_k)} =: \sum_{k\in\mathbb{Z}} f_k z_k$$

which implies $z_k = (-kg_k), k \in \mathbb{Z}$. Thus, $\mathrm{dom}(D^*) = \Omega^* = \{g \in \mathcal{H} : (kg_k)_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})\}$. Let $h_k = -2\pi i k g_k, k \in \mathbb{Z}$; for $k \neq 0$ we have $g_k = i/(2\pi k)h_k$ which means that $g = \int_0^x h + \Lambda(f)$ where $\Lambda(g)$ is an additive constant, which is a **bounded linear functional** on L^2 (why?). Now, $h \in L^2$ implies $g \in AC(S^1)$ (with derivative in L^2). This is the largest domain of D, with range in L^2. In this simple example, if we extend $\mathrm{dom}(D)$ to Ω^*, the same argument shows that this extended D is self-adjoint.

Example 138 indicates that if we want to extend D even further, then the extended domain, or range, or both cannot consist of usual functions, even allowing for the generalizations used in the L^p spaces.

Let us first relax the restriction on the range. The dual of $C_c([-a, a])$ is the space of Radon measures on $[-a, a]$. The Heaviside function $\Theta(x)$ is not in AC (it's not even continuous, of course). As an element of the dual of C_c it acts as $\langle\Theta, \phi\rangle = \int_0^a \phi(x)dx$. Proceeding as in the previous example, taking ϕ in the dense set $C^1([-a, a])$, we would define $\Lambda = \frac{d}{dx}\Theta$, as an element of C_c^* by

$$\Lambda\phi = -\langle\Theta, \frac{d}{dx}\phi\rangle = -\int_0^a \phi'(s)ds = \phi(0) \Rightarrow \Lambda = \delta(x)$$

where $\delta(x)$ is the Dirac mass measure at zero. Thus $\Theta'(x) = \delta(x)$ exists, as a measure, $\delta(x)$. In the same manner, we would get

$$\Theta''(x) = (\phi \mapsto \phi'(0))$$

This is obviously not defined as a bounded functional on $C_c([-a, a])$, but it is in $(C^1([-a, a]))^*$. This logic prompts us to consider the baseline space of **test functions** $\mathcal{D} = C_c^\infty(\mathbb{R}^n)$.

18 Distribution theory

"Il y a plus de 50 ans que l'ingénieur Heaviside introduisit ses régles de calcul symbolique, dans un mémoire audacieux où des calculs mathématiques fort peu justifiés sont utilisés pour la solution de problèmes de physique. Ce calcul symbolique, où opérationnel, n'a cessé de se développer depuis, et sert de base aux études théoriques des électriciens. Les ingénieurs l'utilisent systématiquement, chacun avec sa conception personnelle, avec la conscience plus ou moins tranquille ; c'est devenu une technique «qui n'est pas rigoureuse mais qui réussit bien». Depuis l'introduction par Dirac de la fameuse fonction $\delta(x)$, qui serait nulle partout sauf pour $x = 0$, de telle sorte que $\int_{-\infty}^{\infty} \delta(x)dx = +1$, les formules du calcul symbolique sont devenues encore plus inacceptables pour la rigueur des mathématiciens. Écrire que la fonction d'Heaviside $Y(x)$ égale 0 pour $x < 0$ et a 1 pour $x \geqslant 0$ a pour dérivée la fonction de Dirac $\delta(x)$ dont la définition même est contradictoire, et parler des dérivées $\delta'(x), \delta''(x)$,... de cette fonction denude d'existence réelle, c'est dépasser les limites qui nous est permises. Comment expliquer le succès de ces méthodes? Quand une telle situation contradictoire se présente, il est bien rare qu'il n'en résulte pas une théorie mathématique nouvelle qui justifie, sous une forme modifiée, le langage des physiciens ; il y a même là une source importante de progrès des mathématiques et de la physique."

"More than 50 years ago the engineer Heaviside introduced his symbolic calculus rules, in an audacious memoir in which mathematical calculations with scant justification were used to solve physical problems. This symbolic calculus, or operational calculus, has not ceased to be developed since, and serves as a foundation for the theoretical studies of electricians. The engineers use it systematically, everyone using his own conception, with a more or less peaceful conscience; it has become a technique "which is not rigorous, but is successful". Ever since Dirac's introduction of the famous function $\delta(x)$, which would be zero everywhere except at $x = 0$, in such a way that $\int_{-\infty}^{\infty} \delta(x)dx = +1$, the formulas of symbolic calculus have become even more unacceptable for the rigor of mathematicians. To write that the Heaviside function $Y(x)$ which equals 0 fo $x < 0$ and 1 for $x \geqslant 1$ has as a derivative the Dirac function $\delta(x)$, whose very definition is contradictory, and then talk about the derivatives $\delta'(x), \delta''(x)$,... of this function devoid of real existence, is to exceed the limits that are permitted to us. How can one explain the success of these methods? When such a contradictory situation presents itself, it is rarely not the case that a new mathematical theory emerges, which justifies, in a modified form, the language of of physicists; there is even, in this, an important source of progress of mathematics and physics."

Laurent Schwartz, Théorie des Distributions

18.1 The space of test functions $\mathcal{D} = C_c^\infty(\mathbb{R}^n)$

In the following we will say that a collection $\mathcal{F}$ of functions f are compactly supported in the compact set K if there is an open set $\mathcal{O}$ with $\overline{O} = K$ such that $f \prec \mathcal{O}$ for all functions in the family.

18.2 The topology on $\mathcal{D}$

The topology on $\mathcal{D}$ is that of an inductive limit of Fréchet space (called "an LF space". It has the following properties:

(i) A sequence $\{f_n\}_{n\in\mathbb{N}}$ converges to $f \in \mathcal{D}$ iff there is an n_0 and a compact K such that **all** $f_n, n \geqslant n_0$ are **supported in** K, and

$$\forall \alpha \in (\mathbb{N} \cup \{0\})^n, \lim_{n\to\infty} \|\partial^\alpha (f_n - f)\|_\infty = 0$$

(ii) A set $S \subset \mathcal{D}$ is bounded iff there is a compact K s.t. S is a bounded subset of $C_c^\infty(K)$.

(iii) A sequence is Cauchy if there is a K s.t. all functions are supported in K and the sequence is Cauchy in $C_c^\infty(K)$.

(iv) Let Y be a locally convex topological space. A mapping $A : \mathcal{D} \to Y$ is continuous if it is continuous on every $C_c^\infty(K)$.

(v) A linear functional $\Lambda : \mathcal{D} \to \mathbb{C}$ is continuous iff there is an N and a K s.t.

$$|\Lambda\phi| \leqslant c_K \sup\{|\partial^\alpha \phi| : x \in K, |\alpha| \leqslant N\}$$

Note 139. *The topology on $\mathcal{D}$ is not metrizable. Indeed, take a sequence of compact sets s.t. $K_j \uparrow \mathbb{R}^n$ as $j \to \infty$. Clearly, $\cup_j C_c^\infty(K_j) = \mathcal{D}$, but for every j, the interior of $C_c^\infty(K_j)$ is empty (why?).*

More precisely, $\mathcal{D}$ is an inductive limit of Fréchet spaces (see Appendix A) as follows. Let $K_1 \subset K_2 \subset \ldots \subset K_i \subset \ldots \subset \mathbb{R}^n$ be compact sets s.t. for all i, K_i is contained in the interior of K_{i+1} (e.g., the balls of radius i centered at the origin). Then $\mathcal{D}$ is the inductive limit of the sequence of Fréchet spaces $C_c^\infty(K_i)$.

Definition 140. *$\mathcal{D}'$, the dual of $\mathcal{D}$, is* **the space of distributions**. *If $F \in \mathcal{D}'$ its value on the function $\phi \in \mathcal{D}$ is denoted by by $\langle F, \phi\rangle$, or even, by abuse of notation, $\int F(x)\phi(x)dx$.*

Definition 141. *It is often useful to restrict test functions to smaller sets: If $\mathcal{O}$ is open (K is compact), $\mathcal{D}(\mathcal{O})$ ($\mathcal{D}(K)$, resp.) denote the compactly supported infinitely differentiable functions whose support is contained in $\mathcal{O}$) (K resp.).*

18.3 Examples of distributions

Check that the following are examples of elements of $\mathcal{D}'$:

1. (Distributions generalize functions.) Any $f \in L^1(\mathbb{R})$ is a distribution, if interpreted as the functional $\langle f, \phi\rangle = \int_{\mathbb{R}^n} f(x)\phi(x)dx$.

2. More generally, Radon measures are distributions acting by $\langle \mu, \phi\rangle = \int_{\mathbb{R}^n} \phi d\mu$.

3. The Dirac mass at zero, at times called "delta function", is the functional $\delta(x)$ defined by $\langle \delta, \phi\rangle = \phi(0)$. More generally, the Dirac mass at x_0, $\delta_{x_0}(x)$ is the distribution $\langle \delta_{x_0}, \phi\rangle = \phi(x_0)$.

4. Derivatives of the Dirac mass at a point: $\langle \partial^\alpha \delta_{x_0}, \phi\rangle = (-1)^{|\alpha|}\partial^\alpha \phi|_{x_0}$.

5. Let $F_N(x, y) = \sum_{k=-N}^N e^{-2\pi i k(x-y)}$. Then, with $\phi \in C_c^\infty([-1/2, 1/2])$, and ϕ_k the Fourier coefficients of ϕ, we have

$$\langle F_N(\cdot, y), \phi\rangle = \sum_{k=-N}^{N} \phi_k e^{iky} \to \phi(y) \text{ as } N \to \infty$$

   and thus $\sum_{k=-N}^N e^{-2\pi i k(x-y)} \to \delta_y(x)$ as $N \to \infty$, in $\mathcal{D}'(-1/2, 1/2)$.

Proposition 142 (Fundamental seqences)**.** *Assume $f \in L^1(\mathbb{R}^n)$ and $\|f\|_1 = 1$; for $t > 0$ define $f_t(x) = t^{-n} f(x/t)$. Then $f_t \to \delta$ as $t \to 0$.*

Proof. Let $\phi \in \mathcal{D}$. Then,

$$\langle f_t, \phi\rangle = t^{-n}\int_{\mathbb{R}^n} f(x/t)\phi(x)dx = \int_{\mathbb{R}^n} f(x)\phi(xt)dx \to \phi(0)$$

by dominated convergence. □

18.4 Support of a distribution

General distributions are not functions, as we have seen. In particular, if $F, G \in \mathcal{D}'$ we cannot meaningfully speak in general of $F(x_0)$, or say that F and G agree at some point. But agreement on an open set is a meaningful notion:

Definition 143. *$F \in \mathcal{D}'$ is zero on the open set $\mathcal{O}$ if F restricted to $\mathcal{D}(\mathcal{O})$ is zero. Similarly, if $F, G \in \mathcal{D}'$ we say that F and G agree on $\mathcal{O}$ if $F - G = 0$ on $\mathcal{O}$.*

Note that this notion coincides with usual equality of functions (a.e.) if F and G are functions.

Proposition 144. *Let $\mathcal{O}_\alpha$ be open sets with $\cup_\alpha \mathcal{O}_\alpha = \mathcal{O}$. If $F \in \mathcal{D}'(\mathcal{O})$ and $F = 0$ on each $\mathcal{O}_\alpha$, then $F = 0$ on $\mathcal{O}$.*

Proof. Let $\phi \in \mathcal{D}(\mathcal{O})$. Since $\operatorname{supp}(\phi)$ is a compact set, $\phi \prec \mathcal{O}_1 \cup \mathcal{O}_2 \cup ... \cup O_m$ for some m (where we re-indexed the sets). Let $\psi_j, j = 1, ..., m$ be a smooth partition of unity on $\operatorname{supp}(\phi)$ with $\psi_j \prec \mathcal{O}_j$. Then $\langle F, \phi\rangle = \sum_j \langle F, \psi_j\phi\rangle = 0$, by assumption. □

Definition 145. *For $F \in \mathcal{D}'$, there is a maximal open set $\mathcal{O}$ in $\mathbb{R}^n$ on which F is zero (which could be empty, of course). Then, the* **support** *of F is $\mathbb{R}^n \setminus \mathcal{O}$.*

Example 146. *The Dirac mass at x_0 has $\{x_0\}$ as a support.*

Note that $\mathcal{D}$ is embedded in $\mathcal{D}'$ by

$$\langle \psi, \phi\rangle := \int_{\mathbb{R}^n} \psi(s)\phi(s)ds \tag{120}$$

Definition 147. *Let T be a linear continuous operator on $\mathcal{D}$. T has a transpose if there is a linear continuous operator $T^\times$ on $\mathcal{D}$ s.t.*

$$\langle T^\times \psi, \phi\rangle =: \langle \psi, T\phi\rangle \tag{121}$$

As an example, the transpose of ∂^α is $(-1)^{|\alpha|}\partial^\alpha$. Note that the transpose is uniquely defined by $(T^\times)^\times = T$ and (121). Check that the transposes below exist and satisfy the rules in 1. and 2.

1. $(aT + bS)^\times = aT^\times + bS^\times$.
2. $(TS)^\times = S^\times T^\times$.

18.5 Extension of operators from functions to distributions

Proposition 148. *Assume T is linear and continuous from $\mathcal{D}$ to $\mathcal{D}$. Define $T^\times$ by*

$$\langle T^\times F, \phi\rangle = \langle F, T\phi\rangle$$

Then $T^\times$ is linear and continuous on $\mathcal{D}'$.

Proof. For linearity:

$$\langle T^{\times}(aF_1 + bF_2), \phi\rangle = \langle aF_1 + bF_2, T\phi\rangle = a\langle F_1, T\phi\rangle + b\langle F_2, T\phi\rangle = a < T^{\times}F_1, \phi\rangle + b\langle T^{\times}F_2, \phi\rangle$$

Continuity: By the definition of the topology on $\mathcal{D}'$, and since $T\phi \in \mathcal{D}$, if a net $\{F_\alpha\}_{\alpha\in A}$ is s.t. $F_\alpha \to F$, then $\langle F_\alpha, T\phi\rangle \to \langle F, T\phi\rangle$ [7]. (Where do we use the continuity of T?) □

Examples. 1. (Differentiation) Let F be any L^1_{loc} function. Then, F has derivatives of all orders in the sense of distributions, since

$$\langle \partial^\alpha F, \phi\rangle = (-1)^{|\alpha|}\int_{\mathbb{R}^n} F(x)\partial^\alpha\phi(x)$$

is a continuous functional on $\mathcal{D}$. Note that if $\phi \in \mathcal{D}$, then $\partial^\alpha\phi \in \mathcal{D}$ and $\phi \mapsto \partial^\alpha\phi$ is in $\mathcal{D}'$ (why?). Thus, if F is a distribution, then $\partial^\alpha F$ is a distribution, defined by $(-1)^{|\alpha|}\langle F, \partial^\alpha\phi\rangle$.

2. (Multiplication with smooth functions) If $F \in \mathcal{D}'$ and $\psi \in \mathcal{D}$, then $F\psi \in \mathcal{D}'$, since $T^{\times} := \phi \mapsto \psi\phi$ satisfies the hypotheses of the Proposition above (it acts continuously on $\mathcal{D}$), and $F\psi$ is then the distribution $\langle F\psi, \phi\rangle := \langle F, \psi\phi\rangle$. Note that smoothness is needed in this definition; if $F = \delta$ and ψ is a function in the sense of L^∞ only, $(\psi\phi)(0)$ is undefined, in general. Likewise, if $\psi \in C^n$only were in $\mathcal{D}'$, then, since the derivative of any order of a distribution is a distribution, it would follow that $\psi^{(n+k)}(0)$ are all definied, which is clearly false in general. **The product of two distributions is not defined, in general.**

3. (Translation) Since $\int_{\mathbb{R}^n} f(x+a)g(x)dx = \int_{\mathbb{R}^n} f(x)g(x-a)dx$ if $f, g \in \mathcal{D}$, the extension to $\mathcal{D}'$ of translation is $\langle \tau_a F, \phi\rangle = \langle F, \tau_{-a}\phi\rangle$, and the proposition applies since $T = \tau_{-a}$ is continuous.

4. (Composition with linear transformations of $\mathbb{R}^n$.) Let M be linear and invertible on $\mathbb{R}^n$. Then, $(T^{\times}\phi)(x) = \phi(M^{-1}x)$ is continuous, and the natural definition of $F \circ M$ (check by taking $F = \psi \in \mathcal{D}$)
$$\langle F \circ M, \phi\rangle = |\det M|^{-1}\langle F, \phi \circ M^{-1}\rangle$$
In particular, if R is the reflection $\mathsf{R}\phi(x) = \phi(-x)$, we have
$$\langle \mathsf{R}F, \phi\rangle = \langle F, \mathsf{R}\phi\rangle$$

Theorem 149 (Regularity). *For any distribution F and compact K, there is a positive integer $N(K)$ and a positive constant $c(K)$ s.t. for all $\phi \in \mathcal{D}(K)$,*

$$|\langle F, \phi\rangle| \leqslant c(K)|\phi|_N, \text{ where } |\phi|_N = \max_{|\alpha|\leqslant N}\|\partial^\alpha\phi\|_\infty \tag{122}$$

In other words, $F \in [C^N(K)]'$.

Proof. By contradiction: assume the inequality is false for all N. Then, for any N there is a $\phi_N \in \mathcal{D}(K)$ s.t. $|\langle F, \phi_N\rangle| = 1$ and $|\phi_N|_N \leqslant 1/N$. However, the sequence $\{\phi_N\}_{N\in\mathbb{N}}$ converges to zero in $\mathcal{D}(K)$ which contradicts $|\langle F, \phi_N\rangle| = 1$ for all N.

□

[7] The spaces $\mathcal{D}$ and $\mathcal{D}'$ are not sequential: there exist sequentially open sets which are not open. The class of all sequentially open sets is not compatible with the vector space structure on $\mathcal{D}'$, [2].

We can generalize the results in Exercise 7 as follows.

Proposition 150. *Consider the space of functions which are in $C^m(\mathbb{T}^n)$ for all $m \leqslant K-1$, with absolutely continuous derivatives of order $K-1$, and derivatives of order K in L^2. Define the norms*

$$\|g\|_{2,K}^2 = \sum_{|\alpha|\leqslant K} \|\partial^\alpha g\|_2^2 < \infty \tag{123}$$

(i) The space $\mathcal{H}_K$ of such functions is a Hilbert space.

(ii) $\mathcal{H}_K$ is equivalently characterized by the Fourier coefficient norm

$$\|g\|_{2,K}^2 = \sum_{k\in\mathbb{Z}^n} (1+|k|^2)^K |\hat{g}_k|^2 < \infty \tag{124}$$

(iii) Smooth functions are dense in $\mathcal{H}_K$ in the norm above.

(iv) If $u \in \mathcal{H}_K$ and $K > n/2$, then u is continuous and $\|u\|_\infty < const.\|u\|_K$ where the constant does not depend on u. More generally, if $K > n/2 + M$, then $u \in C^M(\mathbb{T}^n)$ and $|u|_M \leqslant const\|u\|_K$, where the constant does not depend on u. Consequently, iu $K > n/2 + M$, then $\mathcal{H}_K$ is continuously embedded in $C^M(\mathbb{T}^n)$, and is a dense subset of $C^M(\mathbb{T}^n)$.

Proof. (i) Straightforward: this norm comes from an inner product.

(ii) Parseval.

(iii) Smooth functions are those for which all norms above indexed by $K \in \mathbb{N}$ are finite. Density is obvious, as if we simply truncate the series in (124) at $k = k_N$, then the function corresponding to it is smooth for any k_N and in the limit $k_N \to \infty$ we recover the infinite sum.

(iv) We have, by Cauchy-Schwarz

$$(\sup_{\mathbb{T}^n} |u|)^2 \leqslant \left(\sum_{k\in\mathbb{Z}^n} |u_k|\right)^2 \leqslant \left(\sum_{k\in\mathbb{Z}^n} (1+|k|^2)^K |\hat{u}_k|^2\right) \sum_{k\in\mathbb{Z}^n} \frac{1}{(|k|^2+1)^K} \leqslant const.\|u\|_K \tag{125}$$

The case $K > n/2 + M$ is similar and left as an exercise. □

Theorem 151. *Any distribution with compact support can be written in the form*

$$F = \sum_{|\alpha|\leqslant L} \partial^\alpha g_\alpha \tag{126}$$

where g_α are continuous functions and L is some nonnegative integer.

Proof. Let $\psi \in \mathcal{D}$ be supported in an open cube containing the support of F and which is one on $\operatorname{supp}(F)$. Without loss of generality, we may assume that this is a cube side 1 centered at the origin. By the definition of the support, we see that $F\psi = F$, and since $\psi\phi$ is zero outside C, by Theorem 149, there exist $c = c(C)$ and $M = M(C)$ s.t for all $\phi \in \mathcal{D}$ we have

$$|\langle F, \phi\rangle| = |\langle F, \psi\phi\rangle| \leqslant const.|\psi\phi|_M \leqslant const.|\phi|_M$$

Let $N > M + n/2$. If $u \in \mathcal{H}_N$ then $u \in C^M(\mathbb{T}^n)$ and F is a linear functional on $\mathcal{H}_N$. F is also continuous in $\|\cdot\|_N$ since $|F(u)| \leqslant const.|u|_M \leqslant const\|u\|_N$. Therefore, F is the inner product with an element $g \in \mathcal{H}_N$, and, if $\phi \in \mathcal{D} \subset \mathcal{H}_N$, we have

$$F(\phi) = \sum_{|\alpha|\leqslant N} \int_{\mathbb{T}^n} (\partial^\alpha \phi)(\partial^\alpha \overline{g}) = (-1)^{|\alpha|} \sum_{|\alpha|\leqslant N} \int_{\mathbb{T}^n} \overline{g}(\partial^{2\alpha}\overline{\phi}) = \sum_{|\alpha|\leqslant N} \langle (-1)^{|\alpha|}\partial^{2\alpha} g, \phi \rangle$$

□

Note 152. *The functions g_α can be chosen to be compactly supported. Indeed let $\chi \in \mathcal{D}$ be s.t $\chi = 1$ on the support of F. Then $F = F\chi$ and for any $\phi \in \mathcal{D}$ we have*

$$\langle F, \phi \rangle = \langle F, \chi\phi \rangle = \langle \partial^\alpha g, \chi\phi \rangle = \langle g, \partial^\alpha(\chi, \phi) \rangle = \sum_{\beta_1+\beta_2=\alpha} \langle g, c_{\beta_1\beta_2} \partial^{\beta_1}\chi \partial^{\beta_2}\phi \rangle = \sum_{\beta_1+\beta_2=\alpha} \langle c_{\beta_1\beta_2} g \partial^{\beta_1}\chi, \partial^{\beta_2}\phi \rangle$$

and thus

$$F = \sum_{\beta_1+\beta_2=\alpha} c_{\beta_1\beta_2} \partial^{\beta_2}(g\partial^{\beta_1}\chi) = \sum_{|\gamma|\leqslant|\alpha|} \partial^\gamma g_\gamma$$

where $\operatorname{supp} g_\gamma \subset \operatorname{supp}\chi$.

Corollary 153. *For any K, $\mathcal{D}$ is embedded densely in $\mathcal{D}'(K)$.*

Proof. If $F \in \mathcal{D}'(K)$, then $F = \partial^\beta g$ for some continuous g, by the previous theorem. Let $\{\psi_n\}_{n\in\mathbb{N}}$ be a set of functions in $\mathcal{D}$ converging to g. In the topology of $\mathcal{D}'$, $\partial^\beta\psi_n \to \partial^\beta g$. □

Convolution with elements of $\mathcal{D}$. This is defined, generalizing the convolution within $\mathcal{D}$ by (check that it is a generalization!)

$$\langle (F * \psi), \phi \rangle = \langle F, \phi * \mathsf{R}\psi \rangle$$

For example,

$$\langle (\delta * \psi), \phi \rangle = \langle \delta, \phi * \mathsf{R}\psi \rangle = \int_{\mathbb{R}^n} \phi(s)\psi(-(0-s))ds = \langle \psi, \phi \rangle \Rightarrow \delta * \psi = \psi$$

An alternative formula is obtained as follows. If ϕ and F are in $\mathcal{D}$ then

$$(F * \phi)(x) = \int_{\mathbb{R}^n} F(s)\phi(x-s)ds = \langle F(s), \tau_x \mathsf{R}\phi \rangle \tag{127}$$

and simple estimates show that the operation is continuous in $\mathcal{D}$; thus the extension of convolution to $\mathcal{D}'$ is given by (127). The two definitions coincide, by continuity and density of $\mathcal{D}$ in $\mathcal{D}'$.

Proposition 154 (Smoothing of distributions by convolution)**.** *$F * \phi := \langle F, \tau_x \mathsf{R}\phi \rangle$ is C^∞ and $\partial^\alpha F * \phi = \langle F, \partial^\alpha \tau_x \mathsf{R}\phi \rangle = \langle F, \tau_x \partial^\alpha \mathsf{R}\phi \rangle = (\partial^\alpha F) * \phi = F * \partial^\alpha\phi$.*

Proof. Note first that the continuity of ϕ implies that $\lim_{\epsilon\to 0}(\tau_\epsilon \mathsf{R}\phi - \mathsf{R}\phi) = 0$ in the topology of $\mathcal{D}$. Thus

$$\langle F, \tau_{x+\epsilon}\mathsf{R}\phi \rangle \to \langle F, \mathsf{R}\tau_x\phi \rangle \text{ as } \epsilon \to 0$$

and thus the (usual) function $g(x) = \langle F, \tau_x\phi \rangle$ is continuous. Next (take first $n = 1$), we see that $\epsilon^{-1}(\tau_\epsilon \mathsf{R}\phi - \mathsf{R}\phi) \to \mathsf{R}\phi'$ in the topology of $\mathcal{D}$, and thus

$$\epsilon^{-1}(\langle F, \tau_{x+\epsilon}\mathsf{R}\phi \rangle - \langle F, \tau_x \mathsf{R}\phi \rangle) \to \langle F, \tau_x \mathsf{R}\phi' \rangle = \langle F', \tau_x \mathsf{R}\phi \rangle > \text{ as } \epsilon \to 0$$

and g defined above is differentiable. Inductively, it is infinitely differentiable. Since proving differentiability involves on variable at a time, the result follows. $\square$

18.6 The Hadamard finite part

Distributions can be used to regularize certain divergent integrals, as first anticipated by Hadamard in the theory of hyperbolic PDEs. I adapt this example from [4]. The integral we want to regularize is

$$\int_0^\infty \phi(x)x^{-3/2}dx$$

Let $f(x) = 0$ for $x < 0$ and $f(x) = -2x^{-1/2}$ for $x > 0$. Then

$$\langle f', \phi\rangle = -\langle f, \phi'\rangle = -\int_0^\infty \frac{\phi'(x)}{x^{1/2}} = -\lim_{\epsilon\to 0}\int_\epsilon^\infty \frac{\phi'(x)}{x^{1/2}} = \lim_{\epsilon\to 0}\left(\int_\epsilon^\infty \frac{\phi(x)}{x^{3/2}} - 2\phi(0)\epsilon^{-1/2}\right)$$

which, at least when $\phi \in C^1$, is a finite number.

18.7 Green's function

This is a very important method to solve inhomogenneous PDEs (or ODEs). Let L_x be a differential or partial differential operator in some domain with specified boundary conditions. Suppose we solve the non-homogeneous problem $L_x G(x, y) = \delta(x - y)$ (here, we take some licence in the notation, and we agree that $x \in \mathbb{R}^n$ is the variable of the equation $Lg = f$ and $y \in \mathbb{R}^n$ is a parameter). Then,

$$L_x\langle G(x, y), f(y)\rangle = \langle L_x G(x, y), f(y)\rangle = \langle \delta(x - y), f(y)\rangle = f(x)$$

and thus, the solution of the non-homogeneous equation is obtained from a universal kernel for the given equation, the Green function $G(x, y)$ by

$$h(x) = \int G(x, y)f(y)dy \Rightarrow L_x h = f$$

19 The dual of $C^\infty(\mathcal{O})$

The topology on $C^\infty(\mathcal{O})$ is that of uniform convergence on compact sets. Take an increasing sequence of precompact open subsets of $\mathcal{O}$, $\{\mathcal{O}_j\}_j$ with closures $\{K_j\}_j, K_{j+1} \supset \mathcal{O}_j$ whose union is $\mathcal{O}$ and introduce the seminorms

$$\|f\|_{[j,\alpha]} = \sup_{x\in K_j} |\partial^\alpha f(x)|$$

These seminorms define a Fréchet space structure on $C^\infty(\mathcal{O})$.

Proposition 155. *$\mathcal{D}(\mathcal{O})$ is dense in $C^\infty(\mathcal{O})$.*

Proof. Indeed, we can take a sequence of smooth functions ψ_j s.t. $K_j \prec \psi_j \prec \mathcal{O}_{j+1}$ and it is clear that $\lim_j \psi_j = 1$ in the seminorms above, hence, for any $\phi \in C^\infty(\mathcal{O})$, $\lim_j \psi_j\phi = \phi$ in these same seminorms. $\square$

Let $\mathcal{E}'(\mathcal{O})$ be the set of distributions compactly supported in $\mathcal{O}$.

Proposition 156. *The dual of $C^\infty(\mathcal{O})$ is $\mathcal{E}'(\mathcal{O})$. (*Since the elements of $\mathcal{E}'(\mathcal{O})$ are not defined on $C^\infty(\mathcal{O})$, we need to be more precise: any $F \in \mathcal{E}'(\mathcal{O})$ extends uniquely to a linear continuous functional on $C^\infty(\mathcal{O})$, and conversely, the restriction of a linear continuous functional on $C^\infty(\mathcal{O})$ to $\mathcal{D}(\mathcal{O})$ is in $\mathcal{D}'(\mathcal{O})$.*)*

Proof. Let $F \in \mathcal{E}'(\mathcal{O})$. Since $\operatorname{supp} F \subset \cup_j \mathcal{O}_j$, there is an m s.t. $\operatorname{supp} F \subset \mathcal{O}_{m-1}$. If $K_{m-1} \prec \psi \prec \mathcal{O}_m$ then $F\psi = F$, and, by the regularity theorem, for any $\phi \in \mathcal{D}$ and some N we have

$$|\langle F, \phi\rangle| = |\langle F\psi, \phi\rangle| = |\langle F, \psi\phi\rangle| \leqslant C_2 \sum_{|\alpha|\leqslant N} \|\phi\|_{[m,\alpha]}$$

By continuity, F extends uniquely to $\phi \in C^\infty(\mathcal{O})$ by $\langle F, g\rangle = \langle F, \psi g\rangle$ with ψ as above.

Conversely, by the same argument as in the regularity theorem, for any continuous functional G on $C^\infty(\mathcal{O})$ there must be an N and m such that, for any $\phi \in C^\infty(\mathcal{O})$,

$$|\langle G, \phi\rangle| \leqslant const \sum_{|\alpha|\leqslant K} \|\phi_{[m,\alpha]}\| \leqslant const \sum_{|\alpha|\leqslant K} \|\partial^\alpha \phi\|$$

In particular, G is compactly supported in $\mathcal{O}_m$. Thus G is a continuous linear functional when restricted to $\mathcal{D}(\mathcal{O})$; in other words, this restriction is in $\mathcal{D}'(\mathcal{O})$. □

19.0.1 Convolution of distributions

Let $F \in \mathcal{D}'$ and $G \in \mathcal{E}'$. Then, the natural definition of convolution is by dualization,

$$\langle F * G, \phi\rangle = \langle F, \mathsf{R}G * \psi\rangle$$

It can be shown that $F * G = G * F$ in a number of ways, e.g. Exercises 20,21 in Folland, or by density!

20 The Fourier transform

The Fourier transform of a compactly supported function (say in $\mathbb{R}$),

$$\hat{\phi}(\xi) = \int_{\mathbb{R}} e^{-2\pi i \xi x} \phi(x) dx$$

is never compactly supported (unless it is zero). Indeed $\hat{\psi}$ is entire, and analytic functions that vanish on a set with an accumulation point, then are identically zero. We need to **enlarge** the space of test functions by allowing for slower decay at infinity, thus preventing the analyticity of the Fourier transform. An enlarged space which is invariant under $\mathcal{F}$ is $\mathcal{S}$.

Recall the topolgy of $\mathcal{S}$, and that $C_c^\infty(\mathbb{R}^n) = \mathcal{D}$ is dense in $\mathcal{S}$ in the topology of $\mathcal{S}$.

Definition 157. *$\mathcal{S}'$, the dual of $\mathcal{S}$, is the space of* **tempered distributions**.

While functions in $\mathcal{D}$ are zero outside a compact set, functions in $\mathcal{S}$, while still going to zero fast at infinity, they are fast-decaying only when compared to polynomials. This induces by duality a constraint on the growth of the elements of the dual.

Examples 158. (a). Let f be in $L^1_{loc}(\mathbb{R}^n)$ and assume that, for some N, $(1+|x|)^{-N}|f(x)|$ is bounded in $\mathbb{R}^n$. Then $\int_{-\infty}^{\infty} f(x)\phi(x)dx$ is a continuous functional on $\mathcal{S}$ (check!).

(b). $e^{ax} \in \mathcal{S}(\mathbb{R})'$ **iff** $\Re a = 0$. Indeed, if $\Re a = 0$, this follows from the previous example. Otherwise, by symmetry, we can reduce to the case $\Re a > 0$; let ϕ be compactly supported with $\int \phi = 1$ and let $\psi_j = \phi(x-j)e^{-ax}$. We see that $\psi_j \to 0$ as $j \to \infty$ in $\mathcal{S}$, while $\int e^{-ax}\psi_j = 1$ for all j.

Definition 159. *A C^∞ function ψ is* **slowly increasing** *if, together with its derivatives it does not grow faster than polynomially. More precisely, for any α $\exists N(\alpha) \in \mathbb{N}$ s.t.*

$$\left\| \frac{\partial^\alpha \psi}{(1+|x|)^{N(\alpha)}} \right\|_\infty < \infty \tag{128}$$

Proposition 160. *(i) If $F \in \mathcal{S}'$ and ψ is slowly increasing, then $F\psi \in \mathcal{S}'$.*
(ii) If $F \in \mathcal{S}'$ and $\psi \in \mathcal{S}$, then $F * \psi$ is slowly increasing, and, for $\phi \in \mathcal{S}$, we have

$$\langle F, \phi * \mathsf{R}\psi \rangle = \int_{\mathbb{R}^n} \phi(x)(F*\psi)(x)dx \tag{129}$$

Proof. (i) $\langle F\psi, \phi\rangle := \langle F, \psi\phi\rangle$ is an element of $\mathcal{S}$ since $\psi\phi$ is in $\mathcal{S}$, as it is easy to check.

(ii) First, we have checked that $F*\phi$ is smooth if ϕ is. As in the proof of the regularity theorem, for a given F to be in $\mathcal{S}'$, F must be bounded with respect to a finite number of seminorms that define the Fréchet space $\mathcal{S}$, that is, $\exists m, N, C$ s.t. for all $\phi \in \mathcal{S}$

$$|\langle F, \phi\rangle| \leqslant C \max\{\|\phi\|_{m,\alpha} : |\alpha| \leqslant N\}$$

Note also that for any $x, y \in \mathbb{R}^n$, $1+|x| \leqslant 1+|x-y|+|y| \leqslant (1+|x-y|)(1+|y|)$. Since $\partial^\beta F * \phi = F * \partial^\beta \phi = \langle F, \tau_x \mathsf{R}\partial^\beta\phi\rangle$ we have

$$\begin{aligned}|\partial^\beta F * \phi|(x) \leqslant \max_{|\alpha|\leqslant N}\sup_s (1+|s|)^m|\partial^{\alpha+\beta}\phi(x-s)| &\leqslant (1+|x|)^m \max_{|\alpha|\leqslant N}\sup_s (1+|x-s|)^m|\partial^{\alpha+\beta}\phi(x-s)| \\ &\leqslant C(1+|x-s|)^m \max_{|\alpha|\leqslant N+|\beta|} \|\phi\|_{m,\alpha}\end{aligned} \tag{130}$$

Since $\mathcal{D}$ is dense in $\mathcal{S}$, its embedding in $\mathcal{D}'$ is dense in $\mathcal{D}' \supset \mathcal{S}'$ we can check that $\mathcal{D}$ is dense in $\mathcal{S}'$. (129) is abvious if $F \in \mathcal{D}$, and the rest follows by continuity and dominated convergence. $\square$

We note that for $f, g \in \mathcal{S}$ we have

$$\langle \hat{f}, g\rangle = \iint f(x)g(y)e^{-2\pi ixy}dxdy = \langle f, \hat{g}\rangle \tag{131}$$

the definition of the Fourier transform of a distribution should be: for $F \in \mathcal{S}'$ and $g \in \mathcal{S}$,

$$\langle \hat{F}, g\rangle := \langle F, \hat{g}\rangle \tag{132}$$

It follows by duality that the basic properties of the Fourier transform that we established for functions in $\mathcal{S}$ hold for functions in $\mathcal{S}'$. Check also the following Fourier transforms:

$$\mathcal{F}\delta(x-x_0) = e^{-2\pi i x_0 k} \Rightarrow \mathcal{F}e^{2\pi i k_0 x} = \delta(k-k_0) \tag{133}$$

Note that the last equality can be interpreted as a generalized orthonormality relation of $e^{2\pi i k_1 x}$ and $e^{2\pi i k_2 x}$!

Exercise 24. *Show that:*

(i)

$$\mathcal{F}(PV\frac{1}{x}) = -\pi i \operatorname{sgn}(k)$$

(One way is to regularize the PV distribution, and show first that $\mathcal{F}\left(\operatorname{sgn}(k)e^{-\epsilon|k|}\right) = \dfrac{4i\pi x}{4\pi^2 x^2 + \epsilon^2}$*.), and also*

(ii)

$$\mathcal{F}\chi_{[0,\infty)} = \frac{1}{2}\delta(k) + \frac{1}{2\pi i}\mathrm{PV}\left(\frac{1}{k}\right)$$

$$\mathcal{F}\left(\delta(x)^{(n)}\right) = (-2\pi i)^n k^n$$

(iii) In particular check that the Fourier transform of linear combinations of the delta function and its derivatives are precisely the polynomials (and vice-versa).

(iv) Let $\{a_n\}_{n\in\mathbb{Z}}$ *be an* ℓ^1 *sequence. Then*

$$\mathcal{F}\left(\sum_{n\in\mathbb{Z}} a_n e^{2\pi i n x}\right) = \sum_{n\in\mathbb{Z}} a_n \delta(k-n)$$

which is the "Fourier spectrum " of a periodic functions.

Proposition 161. *Let* $F \in \mathcal{E}'$*. Then* $\hat{F} = \langle F, e^{-2\pi i x\cdot\xi}\rangle$*.* $\hat{F}$ *is an entire function of slow growth.*

Proof. We use the decomposition in Note 152. Clearly, it is enough to prove the result for one distribution of the form $\partial^\gamma g$ where g is continuous and compactly supported. Let $\{\phi_j\}_{j\in\mathbb{N}} \subset \mathcal{D}$ be a sequence compactly supported in some K converging to g in $\mathcal{D}'$. We then have

$$\widehat{\partial^\gamma \phi_j} = \int_K \partial^\gamma \phi_j(x) e^{-2\pi i \xi\cdot x} dx = \langle \partial^\gamma \phi_j, e^{-2\pi i\xi\cdot x}\rangle = \langle \phi_j, \partial^\gamma e^{-2\pi i \xi\cdot x}\rangle$$
$$\underset{j\to\infty}{\to} \langle g, \partial^\gamma e^{-2\pi i\xi\cdot x}\rangle = (-2\pi i\xi)^\gamma \int_K g(x))e^{-2\pi i \xi\cdot x} dx \tag{134}$$

and the rest is straightforward. □

21 Sobolev Spaces

The Fourier transform has the important feature of transforming smoothness properties into decay ones (and vice-versa). Furthermore, the Fourier transform is a unitary operator between L^2 spaces. In many applications (PDEs notably) it is convenient to bring together these features: we can introduce L^2 spaces whose norms enforce a given degree of smoothness. We have already noted that the norm

$$\|f\| = \sum_{|\alpha|\leqslant k} \|\partial^\alpha f\|_2^2 \tag{135}$$

comes from an inner product $(\cdot,\cdot)$, and then the space of functions

$$\{f : \|f\| < \infty\} \tag{136}$$

is a Hilbert space, the **Sobolev space** H_k. Taking the Fourier transform of H_k, we obtain the following dual (Fourier) norm

$$\|\hat{f}\|^2 = \sum_{|\alpha|\leqslant k} \langle \xi^\alpha \hat{f}, \xi^\alpha \hat{f}\rangle_2 = \sum_{|\alpha|\leqslant k} \langle (|\xi|^2)^\alpha \hat{f}, \hat{f}\rangle_2 \leqslant const\|(1+|\xi|^2)^{k/2}\hat{f}\|^2$$

where const only depends on k. Noting that $(1+|z|)^m \leqslant const.(1+|z|^m)$, we see that the norm above is *equivalent* to

$$\|\hat{f}\| = \|(1+|\xi|^2)^{k/2}\hat{f}\|_2$$

In Fourier space we can immediately generalize the norms from $k \in \mathbb{N}$ to any $s \in \mathbb{R}$, which can be interpreted as a norm weighted by $(1+\Delta)^{s/2}$. In fact, we have the following map:

$$\Lambda_s f = \mathcal{F}^{-1}((1+|\xi|^2)^{s/2}\mathcal{F}f) \tag{137}$$

We are now in position to define the Sobolev space $H_s = W^{s,2}$ by

$$H_s = \{F \in \mathcal{S}' : \Lambda_s f \in L^2\} \tag{138}$$

The spaces $W^{s,p}$ generalize H_s by using L^p norms,

$$W^{k,p}(\mathbb{R}^n) := \{f : \Lambda_s f \in L^p(\mathbb{R}^n)\}$$

Note that the elements of Sobolev spaces are distributions. Nonetheless, we have the following:

Proposition 162. *If $f \in H_s$, then $\hat{f}$ and $\check{f}$ are tempered functions.*

Proof. Since $\check{f} = \mathsf{R}\hat{f}$, we only check the statements about $\hat{f}$. The fact that $\Lambda_s f$ is a function (an element of L^2, more precisely), means $(1+|\xi|^2)^{s/2}\hat{f}$, and therefore $\hat{f}$, are functions.

Now, treating f as an element of $\mathcal{S}'$, and using the fact that $\hat{f}$ is a function, we have

$$\langle f, \phi\rangle = \langle \check{f}, \hat{\phi}\rangle = \int \hat{\phi}\mathsf{R}\hat{f}$$

which means that $\hat{f}$ is a tempered distribution and thus a tempered function. □

The inner product that we get by polarization is clearly

$$\langle f, g\rangle_{(s)} = \int_{\mathbb{R}^n} \hat{f}(\xi)(1+|\xi|)^s\overline{\hat{g}(\xi)}d^n\xi$$

The following properties follow easily from the definition

Proposition 163. 1. *$H_0 = L^2$ with $\|\cdot\|_{(0)} = \|\cdot\|_2$.*

2. *The Fourier transform is an isomorphism between H_s and $L^2(\mathbb{R}^n, \mu)$ where $d\mu = (1+|\xi|^2)^s d\xi$.*

3. *$\mathcal{S}$ is dense in H_s for all s (this is most easily seen based on 1. above).*

4. *If $s > t$, then $\|\cdot\|_{(t)} \leqslant \|\cdot\|_{(s)}$ and H_s is dense in H_t in $\|\cdot\|_{(t)}$.*

5. *Λ_t is a unitary isomorphism between H_s and H_{s-t} for all $s, t \in \mathbb{R}$.*

6. *Since $|\xi^\alpha| \leqslant (1+|\xi|^2)^{|\alpha|/2}$, ∂^α is a bounded map between H_s and $H_{s-|\alpha|}$.*

In one dimension $\delta(x)$ is in H_s if $s < -\frac{1}{2}$, and in n dimensions if $s < -\frac{n}{2}$. We see that regularity is measured more finely in this way.

Which Sobolev spaces consist of functions? The following theorem answers this important question.

Theorem 164 (The Sobolev embedding theorem)**.** *If $s > k + n/2$, then*
(i) $H_s \subset C_0^k$.
(ii) $f \in H_s$ implies $\mathcal{F}(\partial^\alpha f) \in L^1$ and $\|\mathcal{F}(\partial^\alpha f)\|_1 \leqslant C\|f\|_{(s)}$ where C only depends on $k - s$.

Proof. We prove (ii) first as (i) follows from it and the Riemann-Lebesgue lemma. We apply Cauchy-Schwarz:

$$\frac{1}{(2\pi)^{|\alpha|}}\int |\mathcal{F}\partial^\alpha f| = \int |\xi^\alpha f| \leqslant \int (1+|\xi|^2)^{\frac{k}{2}}\hat{f} \leqslant \|(1+|\xi|^2)^s \hat{f}\|_2 \; \|\frac{1}{(1+|\xi|^2)^{s-k}}\|_2$$

□

Theorem 165. *If $f \in H_{-s}$, then the functional $\phi \mapsto \langle f, \phi\rangle$ extends continuously to a functional on H_s with norm $\|f\|_{(-s)}$, and any element in the dual of H_s is of this form.*

(Does this mean that the Hilbert space H_{-s} "is the dual of" H_s?)

Proof. By Proposition 162 $\check{f}$ is a tempered function. Cauchy-Schwarz implies

$$|\langle f, \phi\rangle = \int \hat{f}\hat{\phi} \leqslant \|(1+|\xi|^2)^{-s/2}\hat{f}\| \;\; \|(1+|\xi|^2)^{s/2}\hat{\phi}\| = \|f\|_{(-s)}\|\phi\|_{(s)} \tag{139}$$

Conversely, we can start in Fourier space with $\hat{f}$, an element of $\mathcal{F}H_{-s}$ and let it act on an element of H_s by

$$\langle f, \phi\rangle = \int_{\mathbb{R}^n} \hat{f}(\xi)\hat{\phi}(\xi)d\xi \left('' = \int_{\mathbb{R}^n} f(x)\phi(x)dx''\right) = \int_{\mathbb{R}^n} \hat{f}(\xi)(1+|\xi|^2)^{-\frac{s}{2}}\hat{\phi}(\xi)(1+|\xi|^2)^{\frac{s}{2}}d\xi \tag{140}$$

which, again by Cauchy-Schwarz shows that f is an element of $\mathcal{S}'$ which is also in H_s.

□

22 Appendix A: inductive limits of Fréchet spaces

Let V be a topological vector space over $\mathbb{R}$ or $\mathbb{C}$.

Definition 166. *The set $A \subset V$ is*

(a). Convex if $a_1, a_2 \in A$ implies $ta_1 + (1-t)a_2 \in A$ for any $t \in [0, 1]$.

(b). Balanced if $a \in A$ implies $\lambda a \in A$ if $|\lambda| \leqslant 1$;

(c). *Bounded if for any neighborhood $\mathcal{V}$ of 0 there is a $\gamma > 0$ s.t. $\gamma\mathcal{V} \supset A$.*

(d). *Absorbent or absorbing if $\{\cup_{t>0} tA\} = \mathcal{V}$. (The set A can be scaled out to absorb every point in the space.)*

Definition 167. (a). *A family of seminorms on a vector space $\mathcal{V}$ is called separating if for any $0 \neq v \in \mathcal{V}$ there is a seminorm $\|\cdot\|_\alpha$ s.t. $\|v\|_\alpha > 0$.*

(b). *$\mathcal{V}$ is called locally convex if the origin has a local base of absolutely convex absorbent sets.*

Proposition 168. *The topological vector space $\mathcal{V}$ is a locally convex space* **iff** *the topology is given by a family of seminorms.*

Proof. For the "if" part, the proof is immediate; the converse requires Minkovky's functionals and the Hahn-Banach separation theorem, see [3]. □

Theorem 169. *Let $\mathcal{V}$ be a topological vector space whose topology is given by a family of seminorms. Then $\mathcal{V}$ is metrizable, and a translation-invariant metric is determined by*

$$\rho(x,0) = \sum_{k=1}^{\infty} 2^{-k} \frac{\|x\|_k}{1+\|x\|_k} \tag{141}$$

The balls $B(0,r) := \{x : \rho(x,0) < r\}$ are balanced. If $\mathcal{V}$ is complete with respect to ρ, then it is a Fréchet space.

Proof. Largely a straightforward verification, see [3] , p. 437 and on. □

Definition 170. *Let*

$$\mathcal{V}_1 \subset \mathcal{V}_2 \subset \cdots \subset \mathcal{V}_j \cdots$$

be a sequence of Fréchet spaces. Thw inductive limit topology on $\mathcal{V}$ is the strongest locally convex topology, in which the injections $\mathcal{V}_i \to \mathcal{V}$ are continuous.

Theorem 171. *Let $\mathcal{V}$ be an inductive limit of Fréchet spaces as in Definition 170.*

(i). *The open, convex, balanced neighborhoods of zero are the sets W s.t. $W \cap \mathcal{V}_j$ are open, convex, balanced neighborhoods of zero for all j, and these sets uniquely determine the topology of $\mathcal{V}$.*

(ii). *$A \subset \mathcal{V}$ is bounded* **iff** *A is a bounded subset of some fixed $\mathcal{V}_{n_0}$.*

(iii). *A sequence is Cauchy in $\mathcal{V}$* **iff** *there is some n_0 a.t. the sequence is contained $\mathcal{V}_{n_0}$ and is Cauchy there.*

(iv). *Let $\mathcal{X}$ be a locally convex topological vector space. The linear map $T : \mathcal{X} \to \mathcal{V}$ is continuous* **iff** *the restriction of T to $\mathcal{V}_j$ is continuous for every j.*

References

[1] Reed, M. and Simon, B., Methods of Modern Mathematical Physics, vol II, Academic Press (1980).

[2] , Convergence of sequences of distributions, Proceedings Of The American Mathematical Society, v. 27, no 3, (1971).

[3] Gerd Grubb, Distributions and Operators, Springer-Verlag New York (2009).

[4] , L. Schwartz, Théorie des Distributions, Hermann, Paris, 1966.