Partial Differential Equations Example sheet 2

David Stuart
dmas2@cam.ac.uk

2 Background analysis

2.1 Fourier series

Consider the following spaces of 2π-periodic functions on the real line:

$$C^r_{per}([-\pi,\pi]) = \{u \in C^r(\mathbb{R}) : u(x+2\pi) = u(x)\}\,,$$

for $r \in [0,\infty]$. The case $r = 0$ is the continuous 2π-periodic functions, while the case $r = \infty$ is ths smooth 2π-periodic functions. For functions $u = u(x_1, \ldots x_n)$ we define the corresponding spaces $C^r_{per}([-\pi,\pi]^n)$ of C^r functions which are 2π-periodic in each coordinate. (All of these definitions generalize in obvious ways for classes of functions with periods other than 2π, e.g. $C^r_{per}(\prod_{j=1}^n [0, L_j])$ consists of C^r functions $u = u(x_1, \ldots x_n)$ which are L_j-periodic in x_j.)

Given a function $u \in C^\infty_{per}([-\pi,\pi])$ the Fourier coefficients are the sequence of numbers $\hat{u}_m = \hat{u}(m)$ given by

$$\hat{u}(m) = \hat{u}_m = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{-imx}u(x)\,dx\,, \qquad m \in \mathbb{Z}\,.$$

Integration by parts gives the formula $\widehat{\partial^\alpha u}(m) = (im)^\alpha \hat{u}(m)$ for positive integral α, which shows that the sequence of Fourier coefficients is a rapidly decreasing (bi-infinite) sequence: this means that $\hat{u} \in s(\mathbb{Z})$ where

$$s(\mathbb{Z}) = \{\hat{u} : \mathbb{Z} \to \mathbb{C} \text{ such that } |\hat{u}|_\alpha = \sup_{m\in\mathbb{Z}} |m^\alpha \hat{u}(m)| < \infty \;\forall \alpha \in \mathbb{Z}_+\}\,.$$

This in turn means that the series $\sum_{m\in\mathbb{Z}} \hat{u}(m)e^{imx}$ converges absolutely and uniformly to a smooth function. The central fact about Fourier series is that this series actually converges to u, so that each $u \in C^\infty_{per}([-\pi,\pi])$ can be represented as:

$$u(x) = \sum \hat{u}(m)e^{imx}\,, \quad \text{where} \quad \hat{u}(m) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{-imx}u(x)\,dx\,.$$

The whole development works for periodic functions $u = u(x_1, \ldots x_n)$ with the sequence space generalized to

$$s(\mathbb{Z}^n) = \{\hat{u} : \mathbb{Z}^n \to \mathbb{C} \text{ such that } |\hat{u}|_\alpha = \sup_{m\in\mathbb{Z}^n} |m^\alpha \hat{u}(m)| < \infty \;\forall \alpha \in \mathbb{Z}^n_+\}\,.$$

Here we use multi-index notation, in terms of which we have:

Theorem 2.1.1 *The mappping*

$$\begin{aligned} C^\infty_{per}([-\pi,\pi]^n) &\to s(\mathbb{Z}^n)\,, \\ u &\mapsto \hat{u} = \{\hat{u}(m)\}_{m\in\mathbb{Z}^n} \qquad \textit{where } \hat{u}(m) = \frac{1}{(2\pi)^n}\int_{[-\pi,\pi]^n} e^{-im\cdot x}u(x)\,dx \end{aligned}$$

is a linear bijection whose inverse is the map which takes $\hat{u}$ *to* $\sum_{m\in\mathbb{Z}^n}\hat{u}(m)e^{im\cdot x}$ *and the following hold:*

1. $u(x) = \sum_{m\in\mathbb{Z}^n}\hat{u}(m)e^{im\cdot x}$ *where* $\hat{u}(m) = \frac{1}{(2\pi)^n}\int_{[-\pi,\pi]^n} e^{-im\cdot x}u(x)\,dx$ *(Fourier inversion),*
2. $\widehat{\partial^\alpha u}(m) = (im)^\alpha\hat{u}(m)$ *for all* $m\in\mathbb{Z}^n, \alpha\in\mathbb{Z}^n_+$*,*
3. $\frac{1}{(2\pi)^n}\int_{[-\pi,\pi]^n}|u(x)|^2\,dx = \sum_{m\in\mathbb{Z}^n}|\hat{u}(m)|^2$ *(Parseval-Plancherel).*

2.2 Fourier transform

Define the Schwartz space of test functions:

$$\mathcal{S}(\mathbb{R}^n) = \{u\in C^\infty(\mathbb{R}^n) : |u|_{\alpha,\beta} = \sup_{x\in\mathbb{R}^n}|x^\alpha\partial^\beta u(x)| < \infty\,,\ \forall\alpha\in\mathbb{Z}^n_+, \beta\in\mathbb{Z}^n_+\,.\}$$

This is a convenient space on which to define the Fourier transform because of the fact that Fourier integrals interchange rapidity of decrease with smoothness, so the space of functions which are smooth and rapidly decreasing is invariant under Fourier transform:

Theorem 2.2.1 *The Fourier transform, i.e. the mapping*

$$\begin{aligned} \mathcal{F}:\mathcal{S}(\mathbb{R}^n) &\to \mathcal{S}(\mathbb{R}^n)\,, \\ u &\mapsto \hat{u} \qquad \textit{where } \hat{u}(\xi) = \mathcal{F}_{x\to\xi}(u(x)) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x}u(x)\,dx \end{aligned}$$

is a linear bijection whose inverse is the map $\mathcal{F}^{-1}$ *which takes* v *to the function* $\check{v} = \mathcal{F}^{-1}(v)$ *given by*

$$\check{v}(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{+i\xi\cdot x}v(\xi)\,d\xi\,,$$

and the following hold:

1. $u(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\hat{u}(\xi)e^{i\xi\cdot x}d\xi$ *where* $\hat{u}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x}u(x)\,dx$ *(Fourier inversion),*
2. $\widehat{\partial^\alpha u}(\xi) = \mathcal{F}_{x\to\xi}(\partial^\alpha u(x)) = (i\xi)^\alpha\hat{u}(\xi)$ *and* $(\partial^\alpha\hat{u})(\xi) = \mathcal{F}_{x\to\xi}((-ix)^\alpha u(x))$ *for all* $x,\xi\in\mathbb{R}^n, \alpha\in\mathbb{Z}^n_+$*,*
3. $\int_{\mathbb{R}^n}\hat{v}(\xi)u(\xi)\,d\xi = \int_{\mathbb{R}^n} v(x)\hat{u}(x)\,dx\,,$
4. $\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\overline{\hat{v}(\xi)}\hat{u}(\xi)\,d\xi = \int_{\mathbb{R}^n}\overline{v(x)}u(x)\,dx$ *(Parseval-Plancherel),*
5. $\widehat{u*v} = \hat{u}\hat{v}$ *where* $u*v = \int u(x-y)v(y)\,dy$ *(convolution).*

2.3 Banach spaces

A norm on a vector space X is a real function $x \mapsto \|x\|$ such that

1. $\|x\| \geq 0$ with equality iff $x = 0$,
2. $\|cx\| = |c|\|x\|$ for all $c \in \mathbb{C}$,
3. $\|x + y\| \leq \|x\| + \|y\|$.

(If the first condition is replaced by the weaker requirement $1'$ that $\|x\| \geq 0$ then the modified conditions $1', 2, 3$ define a semi-norm.) A normed vector space is a metric space with metric $d(x, y) = \|x - y\|$. Recall that a metric on X is a map $d : X \times X \to [0, \infty)$ such that

1. $d(x, y) \geq 0$ with equality iff $x = y$,
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all x, y, z in X.

(This definition does not require that X be a vector space.) The metric space (X, d) is complete if every Cauchy sequence has a limit point: to be precise if $\{x_j\}_{j=1}^{\infty}$ has the property that $\forall \epsilon > 0$ there exists $N(\epsilon) \in \mathbb{N}$ such that $j, k \geq N(\epsilon) \implies d(x_j, x_k) < \epsilon$ then there exists $x \in X$ such that $\lim_{j \to \infty} d(x_j, x) = 0$.

Definition 2.3.1 *A Banach space is a normed vector space which is complete (using the metric $d(x, y) = \|x - y\|$).*

Examples are

- $\mathbb{C}^n$ with the Euclidean norm $\|z\| = (\sum_j |z_j|^2)^{\frac{1}{2}}$.
- $C([a, b])$ with $\|f\| = \sup_{[a,b]} |f(x)|$ (uniform norm).
- Spaces of p-summable (bi-infinite) sequences $\{u_m = u(m)\}_{m \in \mathbb{Z}}$

$$l^p(\mathbb{Z}) = \{u : \mathbb{Z} \to \mathbb{C} \text{ such that } \|u\|_p = (\sum |u(m)|^p)^{\frac{1}{p}} < \infty\}$$

  and generalizations such as $l^p(\mathbb{Z}^n)$ and $l^p(\mathbb{N})$.
- For $s \in \mathbb{R}$ the spaces of bi-infinite sequences $\{u_m = u(m)\}_{m \in \mathbb{Z}}$

$$l^p_s(\mathbb{Z}) = \{u : \mathbb{Z} \to \mathbb{C} \text{ such that } \|u\|_{l^p_s} = (\sum |(1 + m^2)^{\frac{s}{2}} u(m)|^p)^{\frac{1}{p}} < \infty\}$$

  and generalizations such as $l^p_s(\mathbb{Z}^n)$ and $l^p_s(\mathbb{N})$.

- Spaces of measurable L^p functions for $1 \le p < \infty$

$$L^p(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{C} \text{ measurable with } \|u\|_p = (\int |u(x)\|^p \, dx)^{\frac{1}{p}} < \infty\}$$

and generalizations such as $L^p([-\pi,\pi]^n)$ and $L^p([0,\infty))$ etc. For $p = \infty$ the space $L^\infty(\mathbb{R}^n)$ consists of measurable functions which are bounded on the complement of a null set, and the least such bound is called the essential supremum and gives the norm $\|u\|_{L^\infty}$. In this example we identify functions which agree on the complement of a null set (almost everywhere). Read the appendix for an informal introduction to the Lebesgue spaces and a list of results from integration that we will use in this course[1].

Completeness is important because methods for proving that an equation has a solution typically produce a sequence of "approximate solutions", e.g. Picard iterates for the case of ode. If this sequence can be shown to be Cauchy in some norm then completeness ensures the existence of a limit point which is the putative solution, and in good situations can be proved to be a solution. The basic result is the *contraction mapping principle* which can be used to prove existence of solutions of equations of the form $Tx = x$, i.e. to find fixed points of mappings $T : X \to X$ defined on complete (non-empty) metric spaces:

Theorem 2.3.2 *Let (X,d) be a complete, non-empty, metric space and $T : X \to X$ a map such that*

$$d(Ty_1, Ty_2) \le kd(y_1, y_2)$$

with $k \in (0,1)$. Then T has a unique fixed point in X; in fact if $y_0 \in X$, then $T^m y_0$ converges to a fixed point as $m \to \infty$.

Proof We first prove uniqueness of any fixed point: notice that if $Ty_1 = y_1$ and $Ty_2 = y_2$ then the contraction property implies

$$d(y_1, y_2) = d(Ty_1, Ty_2) \le kd(y_1, y_2)$$

and therefore, since $0 < k < 1$, we have $d(y_1, y_2) = 0$ and so $y_1 = y_2$.

To prove existence, firstly, by the triangle inequality:

$$\begin{aligned} d(T^{n+r+1}y_0, T^n y_0) &\le \sum_{m=0}^{r} d(T^{n+m+1}y_0, T^{n+m}y_0) \\ &\le \sum_{m=0}^{r} k^m d(T^{n+1}y_0, T^n y_0) \le \sum_{m=0}^{r} k^{n+m} d(Ty_0, y_0)\,. \end{aligned}$$

But $0 < k < 1$ implies that $\sum_{m \ge 0} k^m = \frac{1}{1-k} < \infty$ and hence $T^m y_0$ forms a Cauchy sequence in X. So by completeness of X, $T^m y_0 \to y$ some y. But then $T^{m+1}y_0 \to Ty$, because T is continuous by the contraction property, and so $Ty = y$ and y is a fixed point. Thus T has a unique fixed point. □

Existence theorem for ordinary differential equations (ode). We now review the proof of the existence theorem for ode via the contraction mapping theorem in the Banach space of continuous functions with the uniform norm. We first note the following result:

[1] You will not be examined on any subtle points connected with the Lebesgue integral

Theorem 2.3.3 (Corollary to the contraction mapping principle) *Let (X, d) be a complete, non-empty, metric space and suppose $T : X \to X$ is such that T^n is a contraction for some $n \in \mathbb{N}$. Then T has a unique fixed point in X; in fact if $y_0 \in X$, then $T^m y_0$ converges to a fixed point as $m \to \infty$.*

Proof By Theorem 2.3.2, T^n has a unique fixed point, y. We also have that

$$T^n(Ty) = T^{n+1}y = T(T^n y) = Ty.$$

So Ty is also a fixed point of T^n and fixed points are unique so $Ty = y$. Also $T^{mn}y_0 \to y$ implies that $T^{mn+1}y_0 \to Ty = y$ and so on, until $T^{mn+(n-1)}y_0 \to y$ as ($m \to \infty$). All together this implies that $T^m y_0 \to y$. □

Theorem 2.3.4 (Existence theorem for ode) *Let $f(t, x)$ be a vector-valued continuous function defined on the region*

$$\{(t, x) : |t - t_0| \leq a, \|x - x_0\| \leq b\} \subset \mathbb{R} \times \mathbb{R}^n$$

which also satisfies a Lipschitz condition in x:

$$\|f(t, x_1) - f(t, x_2)\| \leq L\|x_1 - x_2\|.$$

Define $M = \sup |f(t, x)|$ and $h = \min(a, \frac{b}{M})$. Then the differential equation

$$\frac{dx}{dt} = f(t, x), \;\; x(t_0) = x_0 \tag{2.1}$$

has a unique solution for $|t - t_0| \leq h$.

Proof This is a consequence of the contraction mapping theorem applied in the complete metric space:

$$X = \{x \in C([t_0 - h, t_0 + h], \mathbb{R}^n) : \|x(t) - x_0\| \leq Mh \; \forall t \in [t_0 - h, t_0 + h]\},$$

endowed with the metric $d(x_1, x_2) = \sup\limits_{|t-t_0|\leq h} \|x_1(t) - x_2(t)\|$. (Recall that a limit in the uniform norm of continuous functions is itself continuous.)

Introduce the integral operator T by the formula

$$(Tx)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds. \tag{2.2}$$

The condition $Mh \leq b$ implies that $T : X \to X$. Notice that $x \in X$ solves (2.1) if and only if $Tx = x$, by the fundamental theorem of calculus. In particular, observe that $Tx = x$ implies that $x \in X$ is in fact continuously differentiable.

We now assert that, for $|t - t_0| \leq h$,

$$\|T^k x_1(t) - T^k x_2(t)\| \leq \frac{L^k}{k!}|t - t_0|^k d(x_1, x_2).$$

For $k=0$, this is obvious, and in general it follows by induction since

$$\begin{aligned}\|T^k x_1(t) - T^k x_2(t)\| &\le \int_{t_0}^t \|f(s, T^{k-1}x_1(s)) - f(s, T^{k-1}x_2(s))\| ds \\ &\le L\int_{t_0}^t \|T^{k-1}x_1(s) - T^{k-1}x_2(s)\| ds \\ &\le \frac{L^k}{(k-1)!}\int_{t_0}^t |s-t_0|^{k-1}\, ds\; d(x_1,x_2) \\ &\le \frac{L^k}{k!}|t-t_0|^k d(x_1,x_2).\end{aligned}$$

This implies that T^k is a contraction mapping for sufficiently large k, and so the result follows from Theorem 2.3.3. □

Theorem 2.3.5 (Gronwall Lemma) *Let* $u \in C([t_0,t_1])$ *satisfy*

$$u(t) \le A + B\int_{t_0}^t u(s)ds$$

for $t_0 \le t \le t_1$ *and some positive constants* A, B *. Then*

$$u(t) \le Ae^{B(t-t_0)}, \quad \textit{for } t_0 \le t \le t_1 .$$

Proof Define $F(t) = A + B\int_{t_0}^t u(s)ds$ - by the fundamental theorem of calculus this is a C^1 function which verifies $\dot{F} = Bu \le BF$. It follows that $e^{-Bt}F(t)$ is non-increasing, so that $e^{-Bt}F(t) \le e^{-Bt_0}F(t_0)$, and hence

$$u(t) \le F(t) \le F(t_0)e^{B(t-t_0)} = Ae^{B(t-t_0)}$$

for $t_0 \le t \le t_1$. □

Theorem 2.3.6 *In the situation of Theorem 2.3.4, let* $y(t), w(t)$ *be two solutions of* (2.1) *defined for* $t_0 \le t \le t_1 \le t_0 + a$ *and such that* $\|y(t) - x_0\| \le b$ *and* $\|w(t) - x_0\| \le b$ *on* $[t_0, t_1]$ *. Then*

$$\|y(t) - w(t)\| \le \|y(t_0) - w(t_0)\| e^{L(t-t_0)}$$

for $t_0 \le t \le t_1$ *.*

Proof $u(t) = y(t) - w(t)$ satisfies, by the Lipschitz property:

$$\begin{aligned}\|u(t)\| = \|y(t) - w(t)\| &= \|y(t_0) - w(t_0) + \int_{t_0}^t \big(f(s,y(s)) - f(s,w(s))\big) ds\ \| \\ &\le \|u(t_0)\| + L\int_{t_0}^t \|u(s)\| ds\,.\end{aligned}$$

The result is now an immediate consequence of Gronwall's inequality. □

Theorem 2.3.6 says that for ode defined by Lipschitz vector fields the solutions vary continuously with the initial data: this is the crucial stability property which is central to the notion of *well-posedness*. To be precise, we would say the initial value problem (2.1) is well-posed with Lipschitz continuous f is well-posed because for each initial value $x(0)$ in a neighbourhood of x_0 there exists a unique local solution which depends continuously on $x(0)$.

For pde the same definition is used, and is very important, however the issue is more subtle: this is because norms (or, possibly, some other topological notion which allows one to define continuity) are used in the definition of well-posed. If a pde can be solved for a solution u which is *uniquely* determined by some set of initial and/or boundary data $\{f_j\}$ then the problem is said to be well-posed in a norm $\|\cdot\|$ if in addition the solution changes a small amount in this norm as the data change. This would be satisfied if, for example, for any other solution v determined by data $\{g_j\}$ there holds the stability estimate:

$$\|u-v\| \le C(\sum_j \|f_j - g_j\|_j)\,, \qquad \text{for some } C > 0\,, \tag{2.3}$$

where $\|\cdot\|_j$ are some collection of norms which measure what kind of changes of data produce small changes of the solution. *Finding the appropriate norms such that* (2.3) *holds for a given problem is a crucial part of understanding the problem - they are generally not known in advance.* Once this is understood, it is helpful with development of numerical methods for solving problems on computers, and tells you in an experimental situation how accurately you need to measure the data to make a good prediction.

To fix ideas consider the problem of solving an evolution equation of the form

$$\partial_t u = P(\partial_x)u$$

where P is a constant coefficient polynomial; e.g. the case $P(\partial_x) = i\partial_x^2$ corresponds to the Schrödinger equation $\partial_t u = i\partial_x^2 u$. If we are solving this with periodic boundary conditions $u(x,t) = u(x+2\pi,t)$ and with given initial data $u(x,0) = u_0(x)$ for $u_0 \in C^\infty_{per}([-\pi,\pi])$. *Formally* the solution can be given as

$$u(x,t) = \sum_{m\in\mathbb{Z}} e^{tP(im)+imx}\hat{u}_0(m) \tag{2.4}$$

and if the initial data $u_o = \sum \hat{u}_0(m)e^{imx}$ is a finite sum of exponentials then (2.4) is easily seen to define a solution since it reduces also to a finite sum. In the general case it is necessary to investigate convergence of the sum so that it does define a solution, then to prove uniqueness of this solution, and finally to find norms for which (2.3) holds. For this final step the Parseval identity is often very helpful, and for the case of the Schrödinger equation the series (2.4) does indeed define a solution for smooth periodic data u_0, v_0 and

$$\max_{t\in\mathbb{R}} \int |u(x,t) - v(x,t)|^2\,dx \le \int |u_0(x) - v_0(x)|^2\,dx\,.$$

This inequality would be interpreted as saying that the Schrödinger equation is well posed in L^2 (globally in time since there is no restriction on t.)

In general an equation defines a well posed problem with respect to specific norms, which encode certain aspects of the behaviour of the solutions and have to be found as part of the investigation: *the property of being well posed depends on the norm.* This is related to the fact that norms on infinite dimensional vector spaces (like spaces of functions) can be inequivalent (i.e. can correspond to different notions of convergence), unlike in Euclidean space $\mathbb{R}^n$.

2.4 Hilbert spaces

A Hilbert space is a Banach space which is also an inner product space: the norm arises as $\|x\| = (x,x)^{\frac{1}{2}}$ where $(\cdot,\cdot) : X \times X \to \mathbb{C}$ satisfies:

1. $(x,x) \geq 0$ with equality iff $x = 0$,
2. $\overline{(x,y)} = (y,x)$,
3. $(ax+by,z) = \overline{a}\,(x,z) + \overline{b}\,(y,z)$ and $(x, ay+bz) = a\,(x,y) + b\,(x,z)$ for complex numbers a, b and vectors x, y, z.

(Funtions $X \times X \to \mathbb{C}$ like this which are linear in the second variable and anti-linear in the first are sometimes called sesqui-linear.) Crucial properties of the inner product in a Hilbert space are the Cauchy-Schwarz inequality $|\,(x,y)\,| \leq \|x\|\|y\|$ and the fact that the inner product can be recovered from the norm via

$$(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 - i\|x+iy\|^2 + i\|x-iy\|^2\right)\,, \qquad \text{(polarization)}\,.$$

Examples include $l^2(\mathbb{Z}^n)$ with inner product $\sum_m \overline{u(m)}v(m)$ and $L^2(\mathbb{R}^n)$ with inner product $(u,v) = \int \overline{u(x)}v(x)\,dx$. Another example is the Sobolev spaces: firstly in the periodic case

$$H^1(\mathbb{R}/2\pi\mathbb{Z}) = \{u \in L^2([-\pi,\pi]) : \|u\|_1^2 = \sum_{m\in\mathbb{Z}} (1+|m|^2)|\hat{u}(m)|^2 < \infty\}\,, \tag{2.5}$$

where $u = \sum \hat{u}(m)e^{imx}$ is the Fourier representation, and secondly

$$H^1(\mathbb{R}^n) = \{u \in L^2(\mathbb{R}^n) : \|u\|_1^2 = \int_{\mathbb{R}^n} (1+|\xi|^2)|\hat{u}(\xi)|^2\,d\xi < \infty\}\,, \tag{2.6}$$

where $\hat{u}$ is the Fourier transform.

The new structure in Hilbert (as compared to Banach) spaces is the notion of orthogonality coming from the the inner product. A set of vectors $\{e_n\}$ is called orthonormal if $(e_n, e_m) = \delta_{nm}$. We will consider only Hilbert spaces which have a countable orthonormal basis $\{e_n\}$ (*separable* Hilbert spaces). In such spaces it is possible to decompose arbitrary elements as $u = \sum u_n e_n$ where $u_n = (e_n, u)$. (The case of Fourier series with $e_m(x) = e^{imx}/\sqrt{2\pi}, m \in \mathbb{Z}$ is an example.) The Parseval identity in abstract form reads $\|u\|^2 = \sum |\,(e_n,u)\,|^2$ and:

Theorem 2.4.1 *Given an orthonormal set $\{e_n\}$ the following are equivalent:*

- $(e_n, u) = 0\ \forall n$ *implies* $u = 0$*, (completeness)*
- $\|u\|^2 = \sum |\,(e_n,u)\,|^2\ \forall u \in X$*, (Parseval),*
- $u = \sum (e_n,u)\,e_n\ \forall u \in X$ *(orthonormal basis).*

A closed subspace $X_1 \subset X$ of a Hilbert space is also a Hilbert space, and there is an orthogonal decomposition

$$X = X_1 \oplus X_1^{\perp}$$

where $X_1^{\perp} = \{y \in X : (x_1, y) = 0 \forall x_1 \in X_1\}$. This means that any $x \in X$ can be written uniquely as $x = x_1 + y$ with $x_1 \in X_1$ and $y \in X_1^{\perp}$, and there is a correponding projection $P_{X_1} x = x_1$.

Associated to a Hilbert space X is its dual space X' which is defined to be the space a bounded linear maps:

$$X' = \{L : X \to \mathbb{C}\,,\ \text{with } L \text{ linear and } \|L\| = \sup_{x \in X, \|x\|=1} |Lx| < \infty\}.$$

The definition of the norm on X' ensures that $|L(x)| \leq \|L\| \|x\|$.

Theorem 2.4.2 (Riesz representation) *Given a bounded linear map L on a Hilbert space X there exists a unique vector $y \in X$ such that $Lx = (y, x)$; also $\|L\| = \|y\|$. The correspondence between L and y gives an identification of the dual space X' with the original Hilbert space X.*

Proof $\operatorname{Ker} L$ is a closed subspace whose orthogonal complement $(\operatorname{Ker} L)^{\perp}$ is one-dimensional: to see this let l_1, l_2 lie in $(\operatorname{Ker} L)^{\perp}$ with $Ll_j = c_j$ for some non-zero numbers c_1, c_2. Then $L(c_1 l_2 - c_2 l_1) = 0$, so that $c_1 l_2 - c_2 l_1 \in (\operatorname{Ker} L) \cap (\operatorname{Ker} L)^{\perp}$ and hence l_1, l_2 are linearly dependent, establishing that $\dim(\operatorname{Ker} L)^{\perp} = 1$. Choose any non-zero vector $v \in (\operatorname{Ker} L)^{\perp}$ - it is unique up to multiplication by scalars. Let $x \in X$ be arbitrary, then

$$x - \frac{(v, x)}{\|v\|^2} v \in ((\operatorname{Ker} L)^{\perp})^{\perp} = \operatorname{Ker} L\,,$$

and hence $Lx = (y, x)$ where $y L(v) v / \|v\|^2$. □

A generalization of this (for non-symmetric situations) is:

Theorem 2.4.3 (Lax-Milgram lemma) *Given a bounded linear map $L : X \to \mathbb{R}$ on a Hilbert space X, and a bilinear map $B : X \times X \to \mathbb{R}$ which satisfies (for some positive numbers $\|B\|, \gamma$):*

- $|B(x, y)| \leq \|B\| \|x\| \|y\| \quad \forall x, y \in X$ *(continuity),*
- $B(x, x) \geq \gamma \|x\|^2 \quad \forall x \in X$ *(coercivity),*

there exists a unique vector $z \in X$ such that $Lx = B(z, x) \forall x \in X$.

Proof The uniqueness statement is a consequence of the coercivity assumption. To prove existence, first apply Riesz representation to the map $y \mapsto B(x, y)$ for fixed x, to deduce the existence of a vector $w_x \in X$ such that $B(x, y) = (w_x, y)$. Put $y = w_x$, then we deduce that $\|w_x\|^2 \leq \|B\| \|x\| \|w_x\|$, and hence $\|w_x\| \leq \|B\| \|x\|$. But also the assignment $x \mapsto w_x$ is linear on account of the bilinearity of B, and therefore we can write $w_x = Ax$ where A is a bounded linear map $X \to X$ with $\|A\| \leq \|B\|$, and

$$B(x, y) = (Ax, y)\,.$$

We now make three assertions about A. Firstly: by the coercivity assumption $(Ax, x) \geq \gamma \|x\|^2$, so that *$A$ is injective*. Secondly: *AX, the range of A, is a closed subspace*. Indeed if $y_n \to y$ is a convergent sequence with $y_n = Ax_n$, then coercivity implies $\gamma \|x_n - x_m\|^2 \leq (A(x_n - x_m), x_n - x_m) = (y_n - y_m, x_n - x_m)$, so that $\|x_n - x_m\| \leq \gamma^{-1} \|y_n - y_m\|$. It follows that $\{x_n\}$ is cauchy, and so there exists $x \in X$ such that $x_n \to x$, and by continuity $Ax = y$, so AX is indeed closed as claimed. Thirdly: $AX = X$ because otherwise there exists a non-zero vector u with $(y, u) = 0 \forall y \in AX$, i.e. $(Ax, u) = B(x, u) = 0$ for all x, which gives a contradiction to coercivity by choosing $x = u$. Thus A is a bounded linear bijection, with bounded inverse.

Finally, apply Riesz representation again to write $L(x) = (w, x)$ for some $w \in X$ (and all $x \in X$.) Then, since we just proved A is surjective, there exists $z \in X$ such that $Az = w$, and so

$$L(x) = (w, x) = (Az, x) = B(z, x), \qquad \forall x \in X$$

completing the proof. □

A bounded linear operator $B : X \to X$ means a linear map $X \to X$ with the property that there exists a number $\|B\| \geq 0$ such that $\|Bu\| \leq \|B\| \|u\| \, \forall u \in X$. As in Sturm-Liouville theory we say a bounded linear operator is diagonalizable if there is an orthonormal basis $\{e_n\}$ such that $Be_n = \lambda_n e_n$ for some collection of complex numbers λ_n which are the eigenvalues.

2.5 Distributions

Definition 2.5.1 *A periodic distribution $T \in C^\infty_{per}([-\pi, \pi]^n)'$ is a continuous linear map $T : C^\infty_{per}([-\pi, \pi]^n) \to \mathbb{C}$, where continuous means that if f_n and all its partial derivatives $\partial^\alpha f_n$ converge uniformly to f then $T(f_n) \to T(f)$. Here we call $C^\infty_{per}([-\pi, \pi]^n)$ the space of test functions.*

A tempered distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ is a continuous linear map $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$, where continuous means that if $\|f_n - f\|_{\alpha,\beta} \to 0$ for every Schwartz semi-norm then $T(f_n) \to T(f)$. Here we call $\mathcal{S}(\mathbb{R}^n)$ the space of test functions.

In both cases for $x_0 \in \mathbb{R}^n$ any fixed point (which may be taken to lie in $[-\pi, \pi]^n$ in the periodic case) the Dirac distribution defined by $\delta_{x_0}(f) = f(x_0)$ gives an example.

Remark 2.5.2 *The notion of convergence on $C^\infty_{per}([-\pi, \pi]^n$ and $\mathcal{S}(\mathbb{R}^n)$ used in this definition makes these spaces into topological spaces in which the convergence must be with respect to a countable family of semi-norms. These are examples of Frechet spaces, a class of topological vector spaces which generalize the notion of Banach space by using a countable family of semi-norms rather than a single norm to define a notion of convergence. Using this notion of convergence one can check that the Fourier transform $\mathcal{F}$ is continuous as is its inverse, and the Fourier inversion theorem can be summarized by the assertion that $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a linear homeomorphism with inverse $\mathcal{F}^{-1}$.*

Remark 2.5.3 *Notice that integrable functions define distributions in a natural way: in the simplest case if g is continuous 2π-periodic function then the formula $T_g(f) = \int_{[-\pi,\pi]} g(x)f(x)\,dx$ defines a periodic distribution and clearly the mapping $g \mapsto T_g$ is an injection of $C_{per}([-\pi, \pi])$ into $(C^\infty_{per}([-\pi, \pi]))'$. Similarly if g is absolutely integrable on $\mathbb{R}^n$ then the formula $T_g(f) = \int_{\mathbb{R}^n} g(x)f(x)\,dx$ defines a tempered distribution. The* *mapping $g \mapsto T_g$ is, properly interpreted, injective: if $g \in L^1(\mathbb{R}^n)$ then $T_g(f) = 0$ for all $f \in \mathcal{S}(\mathbb{R}^n)$ implies that $g = 0$ almost everywhere. On account of this remark distributions are often called "generalized functions". The Dirac example indicates that there are distrubutions which do not arise as T_g.*

Remark 2.5.4 *In these definitions distributions are elements of the dual space of a space of test functions with a specified notion of convergence (a topology). Another frequently used class of distributions is the dual space of $C_0^\infty(\mathbb{R}^n)$ the space of compactly supported smooth functions, topologized as follows: $f_n \to f$ in C_0^∞ if there is a fixed compact set K such that all f_n, f are supported in K and if all partial derivatives of $\partial^\alpha f_n$ converge (uniformly) to $\partial^\alpha f$. This class of distributions is more convenient for some purposes, but not for using the Fourier transform, for which purpose the tempered distributions are most convenient because of remark 2.5.2, which allows the fourier transform to be defined on tempered distributions "by duality" as we now discuss.*

Operations are defined on distributions by using duality to transfer them to the test functions, e.g.:

- Given $T \in \mathcal{S}'(\mathbb{R}^n)$ an arbitrary partial derivative $\partial^\alpha T$ is defined by $\partial^\alpha T(f) = (-1)^{|\alpha|} T(\partial^\alpha f)$.
- Given $T \in \mathcal{S}'(\mathbb{R}^n)$ its fourier transform $\hat{T}$ is defined by $\hat{T}(f) = T(\hat{f})$.
- Given $T \in \mathcal{S}'(\mathbb{R}^n)$ and $\chi \in \mathcal{S}(\mathbb{R}^n)$ the distribution χT is defined by $\chi T(f) = T(\chi f)$. This is also the definition if χ is a polynomial - it makes sense because a polynomial times a Schwartz function is again a Schwartz function.

It is useful to check, with reference to the fact in remark 2.5.3 that distributions are generalized functions, that all such defintions of operations on distributions are designed to extend the corresponding definitions on functions: e.g. for a Schwartz function g we have

$$\partial^\alpha T_g = T_{\partial^\alpha g}\,,$$

where on the left ∂^α means distributional derivative while on the right it is the usual derivative from calculus applied to the test function g. The same principle is behind the other definitions.

There are various alternate notations used for distributions:

$$T(f) \,=\, \langle T\,,\, f\rangle \,=\, \int T(x) f(x)\, dx$$

where in the right hand version it should be remembered that the expression is purely formal in general: the putative function $T(x)$ has not been defined, and the integral notation is not an integral - just shorthand for the duality pairing of the definition. It is nevertheless helpful to use it to remember some formulae: for example the formula for the distributional derivative takes the form

$$\partial^\alpha T(f) \,=\, \int \partial^\alpha T(x) f(x)\, dx \,=\, (-1)^{|\alpha|} \int T(x) \partial^\alpha f(x)\, dx = (-1)^{|\alpha|} T(\partial^\alpha f)\,,$$

which is "familiar" from integration by parts. The formula $\int \delta(x - x_0) f(x)\, dx = f(x_0)$ and related ones are to be understood as formal expressions for the proper definition of the delta distribution above.

2.6 Positive distributions and Measures

In this section[2] we restrict to 2π-periodic distributions on the real line for simplicity. The delta distribution δ_{x_0} has the property that if $f \geq 0$ then $\delta_{x_0}(f) \geq 0$; such distributions are called positive. Positive distributions have an important continuity property as a result: if T is any positive periodic distribution, then since

$$-\|f\|_{L^\infty} \leq f(x) \leq \|f\|_{L^\infty}, \qquad \|f\| = \sup |f(x)|$$

for each $f \in C^\infty_{per}([-\pi,\pi])$ it follows from positivity that $T(\|f\|_{L^\infty} \pm f) \geq 0$ and hence by linearity that

$$-c\|f\|_{L^\infty} \leq T(f) \leq c\|f\|_{L^\infty}$$

where $c = T(1)$ is a positive number. This inequality, applied with f replaced by $f - f_n$, means that if $f_n \to f$ uniformly then $T(f_n) \to T(f)$, i.e. *positive distributions* are automatically continuous with respect to *uniform convergence*, in strong contrast to the continuity property required in the original definition. In fact this new continuity property ensures that a positive distribution can be extended uniquely as a map

$$L : C_{per}([-\pi,\pi]) \to \mathbb{R}$$

i.e. as a continuous linear functional on the space of continuous functions. This extension is an immediate consquence of the density of smooth functions in the continuous functions in the uniform norm (which can be deduced from the Weierstrass approximation theorem). A much more lengthy argument allows such a functional to be extended as an integral $L(f) = \int f\, d\mu$ which is defined for a class of measurable functions f which contains and is bigger than the class of continuous functions. To conclude: positive distributions automatically extend to define continuous linear functional on the space of continuous functions, and hence can be identified with a class of *measures* (Radon measures) which can be used to integrate much larger classes of functions (extending further the domain of the original distribution).

2.7 Sobolev spaces

We define the Sobolev spaces for $s = 0, 1, 2, \ldots$ on various domains:
On $\mathbb{R}^n$: we have the following equivalent definitions:

$$\begin{aligned} H^s(\mathbb{R}^n) &= \{u \in L^2(\mathbb{R}^n) : \|u\|^2_{H^s} = \sum_{\alpha:|\alpha|\leq s} \|\partial^\alpha u\|^2_{L^2} < \infty\} \\ &= \{u \in L^2(\mathbb{R}^n) : \int_{\mathbb{R}^n} (1 + \|\xi\|^2)^s |\hat{u}(\xi)|^2\, d\xi < \infty\} \\ &= \overline{C_0^\infty(\mathbb{R}^n)}^{\|\cdot\|_{H^s}} . \end{aligned}$$

In the first line the partial derivatives are taken in the distributional sense: the precise meaning is that all *distributional (=weak)* partial derivatives up to order s of the distribution T_u determined by u are distributions which are determined by square integrable funtions which are designated $\partial^\alpha u$ (i.e. $\partial^\alpha T_u = T_{\partial^\alpha u}$ with $\partial^\alpha u \in L^2$ in the notation introduced previously). The final line means that H^s is the closure of the space of

[2] This is an optional section, for background only

smooth compactly supported functions $C_0^\infty(\mathbb{R}^n)$ in the Sobolev norm $\|\cdot\|_{H^s}$. The quantity $\tilde{\|}u\tilde{\|}^2_{H^s} = \int_{\mathbb{R}^n}(1+\|\xi\|^2)^s|\hat{u}(\xi)|^2\,d\xi$ appearing in the middle definition defines a norm which is equivalent to the norm $\|u\|_{H^s}$ appearing in the first definition. (Recall that $\|\cdot\|$ and $\tilde{\|}\cdot\tilde{\|}$ are equivalent if there exist positive numbers C_1, C_2 such that $\|u\| \le C_1\tilde{\|}u\tilde{\|}$ and $\tilde{\|}u\tilde{\|} \le C_2\|u\|$ for all vectors u; equivalent norms give rise to identical notions of convergence (i.e. they define the same topologies).

Theorem 2.7.1 *For $s = 0, 1, 2\ldots$ the Sobolev space $H^s(\mathbb{R}^n)$ is a Hilbert space, and so complete in either of the norms*

$$\|u\|^2_{H^s} = \sum_{\alpha:|\alpha|\le s}\|\partial^\alpha u\|^2_{L^2} \quad \textit{or} \quad \tilde{\|}u\tilde{\|}^2_{H^s} = \int_{\mathbb{R}^n}(1+\|\xi\|^2)^s|\hat{u}(\xi)|^2\,d\xi$$

which are equivalent. Given any $u \in H^s(\mathbb{R}^n)$ there exists a sequence u_ν of $C_0^\infty(\mathbb{R}^n)$ functions such that $\|u-u_\nu\|_{H^s} \to 0$ as $\nu \to +\infty$. If $u \in H^s(\mathbb{R}^n)$ for $s > \frac{n}{2}+k$ with $k \in \mathbb{Z}_+$ then $u \in C^k(\mathbb{R}^n)$ and there exists $C > 0$ such that

$$\|u\|_{C^k} = \sum_{|\alpha|\le k}\sup_{x\in\mathbb{R}^n}|\partial^\alpha u(x)| \le C\|u\|_{H^s}\,. \tag{2.7}$$

- The fact that H^s functions can be approximated by $C_0^\infty(\mathbb{R}^n)$ functions means that for many purposes calculations can be done with $C_0^\infty(\mathbb{R}^n)$, or $\mathcal{S}(\mathbb{R}^n)$, functions and in the end the result extended to H^s. See the second worked problem to see how this goes to prove (2.7) with $k = 0$.
- For $s = 0$ we have $H^0 = L^2$ and the norm $\|\cdot\|_{H^0}$ is exactly the L^2 norm, while $\tilde{\|}\cdot\tilde{\|}_{H^0}$ is proportional (and hence equivalent) to the L^2 norm by the Parseval-Plancherel theorem.
- Strictly speaking the assertion $u \in C^k(\mathbb{R}^n)$ in the last sentence of the theorem only holds after possibly redefining u on a set of zero measure. This subtle point, which will generally be ignored in the following, arises because u is really only a distribution which can be represented by an L^2 function, and as such is only defined up to sets of zero measure.

On $(\mathbb{R}/(2\pi\mathbb{Z}))^n$: In the 2π-periodic case the following definitions are equivalent:

$$\begin{aligned} H^s_{per}([-\pi,\pi]^n) &= \{u \in L^2([-\pi,\pi]^n) : \|u\|^2_{H^s} = \sum_{\alpha:|\alpha|\le s}\|\partial^\alpha u\|^2_{L^2} < \infty\} \\ &= \{\sum_{m\in\mathbb{Z}^n}\hat{u}(m)e^{im\cdot x} : \sum_{m\in\mathbb{Z}^n}(1+\|m\|^2)^s|\hat{u}(m)|^2 < \infty\} \\ &= \overline{C^\infty_{per}([-\pi,\pi]^n)}^{\|\cdot\|_{H^s}}\,. \end{aligned}$$

Again the quantity apearing in the middle line defines an equivalent norm which can be used when it is more convenient. Since we are considering only the case $s = 0, 1, 2, \ldots$ the Fourier series $\sum_{m\in\mathbb{Z}^n}\hat{u}(m)e^{im\cdot x}$ always defines a square integrable function, and as s increases the function so defined is more and more regular (exercise), and as above we have:

Theorem 2.7.2 *For $s = 0, 1, 2\dots$ the periodic Sobolev space $H^s_{per}([-\pi,\pi]^n)$ is a Hilbert space, and so complete in either of the norms*

$$\|u\|^2_{H^s} = \sum_{\alpha:|\alpha|\le s} \|\partial^\alpha u\|^2_{L^2} \quad \textit{or} \quad \tilde{\|}u\tilde{\|}^2_{H^s} = \sum_{m\in\mathbb{Z}^n} (1+\|m\|^2)^s |\hat{u}(m)|^2$$

which are equivalent. Given any $u \in H^s_{per}([-\pi,\pi]^n)$ there exists a sequence u_ν of $C^\infty_{per}([-\pi,\pi]^n)$ functions such that $\|u-u_\nu\|_{H^s} \to 0$ as $\nu \to +\infty$. If $u \in H^s([-\pi,\pi]^n)$ for $s > \frac{n}{2}+k$ with $k \in \mathbb{Z}_+$ then $u \in C^k(\mathbb{R}^n)$ and there exists $C > 0$ such that

$$\|u\|_{C^k} = \sum_{|\alpha|\le k} \sup_{x\in[-\pi,\pi]^n} |\partial^\alpha u(x)| \le C\|u\|_{H^s}\,. \tag{2.8}$$

Similar comments to those made after Theorem 2.7.1 apply of course. To keep the notation clean we do not indicate "periodic" in the notation for *norm* on a space of periodic functions, only the space - it should be clear from the context.

With the concept of tempered distribution understood, it is possible to extend the definition of the Sobolev spaces $H^s(\mathbb{R}^n)$ for all $s \in \mathbb{R}$: since the Fourier transform $\hat{u}$ of any tempered distribution is well-defined, we just say

$$H^s(\mathbb{R}^n) \;=\; \{u \in \mathcal{S}'(\mathbb{R}^n) : \int_{\mathbb{R}^n} (1+\|\xi\|^2)^s |\hat{u}(\xi)|^2\, d\xi < \infty\}\,. \tag{2.9}$$

(Notice, it is implicit in this definition that the tempered distribution $\hat{u}$ is in fact represented by a measurable function.)

These various definitions of Sobolev spaces require some modifications for the case of general domains Ω, starting with the notion of the weak partial derivative (since we did not define distributions in Ω).

Definition 2.7.3 *A locally integrable function u defined on an open set Ω admits a weak partial derivative corresponding to the multi-index α if there exists a locally integrable function, designated $\partial^\alpha u$, with the property that*

$$\int_\Omega u\,\partial^\alpha\chi\,dx = (-1)^{|\alpha|}\int_\Omega \partial^\alpha u\,\chi\,dx\,,$$

for every $\chi \in C^\infty_0(\Omega)$.

A useful fact is that in this situation:

$$\|\partial^\alpha u\|_{L^2} \;=\; \sup\{\int_\Omega u\,\partial^\alpha\chi\,dx \;:\; \chi\in C^\infty_0(\Omega) \text{ and } \|\chi\|_{L^2} = 1\,.\} \tag{2.10}$$

Then employing this notion of partial derivative we define (for $s = 0, 1, 2, \dots$):

$$H^s(\Omega) = \{u \in L^2(\Omega) : \|u\|^2_{H^s} = \sum_{\alpha:|\alpha|\le s} \|\partial^\alpha u\|^2_{L^2} < \infty\}$$

(with all L^2 norms being defined by integration over Ω). This space is to be distinguished from the corresponding closure of the space of smooth functions supported in a compact subset of Ω:

$$H^s_0(\Omega) = \overline{C^\infty_0(\Omega)}^{\|\cdot\|_{H^s}}\,.$$

Since these functions are limits of functions which vanish in a neighbourhood of Ω they are to be thought of as vanishing in some generalized sense on $\partial\Omega$ (at least in the case $s = 1, 2, \ldots$ and if Ω has a smooth boundary $\partial\Omega$.) The case $s = 1$ gives the space $H_0^1(\Omega)$ which is the natural Hilbert space to use in order to give a weak formulation of the Dirichlet problem for the elliptic equation $Pu = f$ on Ω.

In one dimension, with $\Omega = (a, b) \subset \mathbb{R}$ the relation between $H^1((a, b))$ and $H_0^1((a, b))$ can be stated simply because all $f \in H^1((a, b))$ are uniformly continuous (similar to the argument in the last part of the 2nd worked problem) and

$$H_0^1((a, b)) = \{f \in H^1((a, b)) : f(a) = f(b) = 0\} .$$

More details and proofs can be found in the relevant chapter of the book of Brezis. In this course we need to be able to use Sobolev spaces to study pde, and we will see that the H^s spaces are easy to work with for various reasons:

1. The "energy" methods often give rise to information about $\int u^2 dx$ or $\int \|\nabla u\|^2 dx$ where u is a solution of a pde, and this translates into information about the solution in Sobolev norms. As a specific example: conservation of energy

$$\frac{1}{2} \int (u_t^2 + \|\nabla u\|^2)\, dx \ = E \ = constant\, ,$$

when u is a solution of the wave equation .

2. The Parseval-Plancherel theorem means that information on Sobolev norms is often easily obtainable when the solution is written down using Fourier methods.

3. The Sobolev spaces H^s are Hilbert spaces (complete) whose elements can be approximated by smooth functions: in practice, this means one has the dual advantages of smoothness of the functions and completeness of the space of functions.

Thus typically we will do some computations for smooth solutions of pde which give information about their Sobolev norms, and then using density we will extend the information to more general (weak) solutions lying in the Sobolev spaces themselves. The use of the full Sobolev space is crucial in any argument relying on completeness, typically in proving existence of a solution e.g. by variational methods or by the Lax-Milgram lemma.

2.8 Appendix: integration

The aim of this appendix[3] is to give a brief review of facts from integration needed - completeness of the L^p spaces, dominated convergence and other basic theorems. We first consider the case of functions on the unit interval $[0, 1]$. A main achievement of the Lebesgue integral is to construct *complete* vector spaces of functions where the completeness is with respect to a norm defined by an integral such as the L^2 norm $\|\cdot\|_{L^2}$ defined by

$$\|f\|_{L^2}^2 = \int_0^1 |f(x)|^2\, dx\, .$$

[3]This section gives a brief introduction to the results on Lebesgue integral which we make use of. You should be able to use the results listed here but will not be examined on the proofs or on any subtleties connected with the results.

This is a perfectly good norm on the space of continuous functions $C([0,1])$, but the resulting normed vector space is not complete (and so not a Banach space) and is not so useful as a setting for analysis. The Lebesgue framework provides a larger class of functions which can be potentially integrated - the *measurable functions*. The complete Lebesgue space L^2 which this construction leads to then consists of (equivalence classes of) measurable functions f with $\|f\|_{L^2}^2 < \infty$; here it is necessary to consider equivalence classes of functions because functions which are non-zero only on sets which are very small (in a certain precise sense) are invisible to the integral, and so have to be factored out of the discussion. The "very small" sets in question are called null sets and are now defined.

2.8.1 Null sets and measurable functions on $[0,1]$

An interval in $[0,1]$ is a subset of the form (a,b) or $[a,b]$ or $(a,b]$ or $[a,b)$ (respectively open,closed, half open). In all cases the length of the interval is $|I| = b-a$. A collection of intervals $\{I_\alpha\}$ *covers* a subset A if $A \subset \cup_\alpha I_\alpha$.

Definition 2.8.1 (Null sets) *For a set $A \subset [0,1]$ we define the* outer measure *to be*

$$|A|_* = \inf_{\{I_n\}_{n=1}^\infty \in \mathcal{C}} \Big\{ \sum_n |I_n| \ :\ A \subset \cup I_n \Big\} ,$$

where $\mathcal{C}$ consists of countable families of intervals in $[0,1]$. A set $N \subset [0,1]$ is null if $|I|_* = 0$, i.e. if for all $\epsilon > 0$ there exists $\{I_n\}_{n=1}^\infty \in \mathcal{C}$ which covers A with $\sum |I_n| < \epsilon$.

Definition 2.8.2 *We say $f = g$ almost everywhere (a.e.) if $f(x) = g(x)$ for all $x \notin N$ for some null set N. We say a sequence of functions f_n converges to f a.e. if $f_n(x) \to f(x)$ for all $x \notin N$ for some null set N.*

Equality a.e. defines an equivalence relation, and two equivalent functions f, g are said to be Lebesgue or measure theoretically equivalent. One way to think about measurable functions is provided by the Lusin theorem, which says a measurable function is one which is "almost continuous" in the sense that it agrees with a continuous function on the complement of a set of arbitrarily small outer measure:

Definition 2.8.3 (Measurable functions) *A function $f : [0,1] \to \mathbb{R}$ is measurable if for every $\epsilon > 0$ there exists a continuous function $f^\epsilon : [0,1] \to \mathbb{R}$ and a set F^ϵ such that $|F^\epsilon|_* < \epsilon$ and $f(x) = f^\epsilon(x)$ for all $x \notin F^\epsilon$. We write $L([0,1])$ for the space of all measurable functions so defined.*

Theorem 2.8.4 *$L([0,1])$ is a linear space closed under almost everywhere convergence: given a sequence $f_n \in L([0,1])$ of measurable functions which converges to a function f a.e. it follows that $f \in L([0,1])$.*

Definition 2.8.3 is not the usual definition of measurability - which involves the notion of a distinguished collection of sets, the σ-algebra of measurable sets - but is equivalent to it by what is called the *Lusin theorem* (see for example §2.4 and §7.2 in the book *Real Analysis* by Folland). The Lusin theorem gives a helpful way of thinking about measurability (the Littlewood 3 principles - see §3.3 in the book *Real Analysis* by Royden and Fitzpatrick). A companion to the Lusin theorem is the *Egoroff theorem* which states that given a sequence $f_n \in L([0,1])$ of measurable functions which converges to a function f a.e. then for every $\epsilon > 0$ it is possible to find a set $E \subset [0,1]$ with $|E|_* < \epsilon$ such that $f_n \to f$ uniformly on $E^c = [0,1] - E$. Thus two of Littlewood's principles say that " a measurable function is one which agrees with a continuous function except on a set which may be taken to have arbitrarily small size" and "a sequence of measurable functions which converges almost everywhere converges uniformly on the complement of a set which may be assumed to be arbitrarily small".

2.8.2 Definition of $L^p([0,1])$

An integral $\int_0^1 f(x)dx$ can be defined of any non-negative measurable function, although the value can be $+\infty$. When the function is continuous, or indeed Riemann integrable, this integral agrees with the Riemann integral, and it has the following properties (for arbitrary non-negative measurable functions f, g):

1. $\int_0^1 cf(x)dx = c\int_0^1 f(x)dx$ if $c > 0$,
2. $\int_0^1 (f(x) + g(x))\,dx = \int_0^1 (f(x) + g(x))\,dx$,
3. $\int_0^1 f(x)\,dx \le \int_0^1 g(x)\,dx$ if $f \le g$ a.e.

Exercise For non-negative measurable functions f, g, show that if $f = g$ a.e. then $\int_0^1 f(x)dx = \int_0^1 g(x)dx$.

Accepting that such a definition exists, we can now define the $L^p([0,1])$ spaces, which are Banach spaces of functions on which there exists a well-defined notion of the integral (called the Lebesgue integral).

Definition 2.8.5 *For $1 \le p < \infty$ define $L^p([0,1])$ to be the linear space of measurable functions on $[0,1]$ with the property that*

$$\|f\|_{L^p}^p = \int_0^1 |f(x)|^p\,dx < \infty\,.$$

For the case $p = \infty$: firstly, say that f is essentially bounded above with upper (essential) bound M if $f(x) \le M$ for $x \notin N$ for some null set N. Then let ess $\sup f$ be the infimum of all upper essential bounds. Then:

Definition 2.8.6 $L^\infty([0,1])$ *is the linear space of measurable functions on $[0,1]$ with the property that*

$$\|f\|_{L^\infty} = \text{ess}\,\sup|f| < \infty\,.$$

A crucial fact for us is that considering the spaces of equivalence classes of functions which agree almost everywhere we obtain *Banach* spaces, also written $L^p([0,1])$: these "Lebesgue spaces" are vector spaces of (equivalence classes of) functions which are *complete* with respect to the norm $\|\cdot\|_{L^p}$. (The fact that, strictly speaking, the elements of these spaces are equivalence classes of functions which agree almost everywhere, is often taken as understood and not repeatedly mentioned each time the spaces are made use of.)

The spaces $L^p([0,1])$ contain the continuous functions, and the Lebesgue integral, which is defined on the whole of these spaces, is equal to the Riemann integral when restricted to Riemann integrable functions. These $L^p([0,1])$ spaces are special cases of $L^p(\mathcal{M})$ spaces which arise from abstract measure spaces $\mathcal{M}$ on which a measure μ (and a σ-algebra of measurable sets) is given; μ measures the "size" of elements of this collection of measurable sets. In the general setting the integral of a function is often defined in terms of the measure of sets on which the function takes given values: for example, one development of the integral takes as starting point the following definition for the integral of a non-negative measurable function:

$$\int f\, d\mu = \int_0^\infty \mu(\{f > \lambda\})\, d\lambda\,. \tag{2.11}$$

The point here is that as λ increases, the sets $\{f > \lambda\}$ decrease and their measure $\mu(\{f > \lambda\})$ decreases also, so that (2.11) is well-defined as the Riemann integral of a monotone function. See the book *Analysis* by Lieb and Loss for a development along these lines.

Other examples of measure spaces used in this course are

- $L^p([a,b])$ with norm $(\int_a^b |f(x)|^p\, dx)^{\frac{1}{p}}$,
- $L^p([-\pi,\pi]^n)$ with norm $(\int_{[-\pi,\pi]^n} |u(x)|^p\, dx)^{\frac{1}{p}}$, and
- $L^p(\Omega)$ with norm $(\int_\Omega |f(x)|^p\, dx)^{\frac{1}{p}}$, where $\Omega \subset \mathbb{R}^n$ is open; the case $\Omega = \mathbb{R}^n$ will occur most often.

2.8.3 Assorted theorems on integration

Theorem 2.8.7 (Hölder inequality) $\int fg dx \le \|f\|_{L^p}\|g\|_{L^q}$ *for any pair of functions* $f \in L^p, g \in L^q$ *(on any measure space) with* $p^{-1} + q^{-1} = 1$ *and* $p, q \in [1,\infty]$.

Corollary 2.8.8 (Young inequality) *If* $f \in L^p(\mathbb{R}^n)$ *and* $g \in L^1(\mathbb{R}^n)$ *then* $f*g \in L^p(\mathbb{R}^n)$ *and* $\|f*g\|_{L^p} \le \|f\|_{L^p}\|g\|_{L^1}$ *for* $1 \le p \le \infty$.

Theorem 2.8.9 (Dominated convergence theorem) *Let the sequence* $f_n \in L^1$ *converge to* f *almost everywhere (on any measure space) and assume that there exists a nonnegative measurable function* $\Phi \ge 0$ *such that* $|f_n(x)| \le \Phi(x)$ *almost everywhere and* $\int \Phi < \infty$*. Then* $\lim_{n\to\infty}\int f_n = \int f$ *and* $\lim_{n\to\infty}\|f_n - f\|_{L^1} = 0\,.$

Corollary 2.8.10 (Differentiation through the integral) *Let* $g \in C^1(\mathbb{R}^n \times \Omega)$ *where* $\Omega \subset \mathbb{R}^m$ *is open, and consider* $F(\lambda) = \int_{\mathbb{R}^n} g(x,\lambda)dx$*. Assume there exists a measurable function* $\Phi(x) \ge 0$ *such that*

- $\int_{\mathbb{R}^n} \Phi(x)\, dx < \infty\,,$
- $\sup_\lambda(|g(x,\lambda)| + |\partial_\lambda g(x,\lambda)|) \le \Phi(x)\,.$

Then $F \in C^1(\Omega)$ *and* $\partial_\lambda F = \int_{\mathbb{R}^n} \partial_\lambda g(x,\lambda)\, dx$.

Corollary 2.8.11 *If f is a $C^k(\mathbb{R}^n)$ function with all partial derivatives $\partial^\alpha f$ of order $|\alpha| \le k$ bounded, and $g \in L^1(\mathbb{R}^n)$ then $f * g \in C^k(\mathbb{R}^n)$ and $\partial^\alpha(f * g) = (\partial^\alpha f) * g$ for $|\alpha| \le k$.*

Theorem 2.8.12 (Tonelli) *If $f \ge 0$ is a nonnegative measurable function $f : \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}$ then*

$$\iint\limits_{\mathbb{R}^l \times \mathbb{R}^m} f(x,y)\,dxdy = \int\limits_{\mathbb{R}^l} \left(\int\limits_{\mathbb{R}^m} f(x,y)\,dy \right) dx = \int\limits_{\mathbb{R}^m} \left(\int\limits_{\mathbb{R}^l} f(x,y)\,dx \right) dy\,.$$

Theorem 2.8.13 (Fubini) *If f is a measurable function $f : \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}$ such that*

$$\iint\limits_{\mathbb{R}^l \times \mathbb{R}^m} |f(x,y)|\,dxdy < \infty$$

then

$$\iint\limits_{\mathbb{R}^l \times \mathbb{R}^m} f(x,y)\,dxdy = \int\limits_{\mathbb{R}^l} \left(\int\limits_{\mathbb{R}^m} f(x,y)\,dy \right) dx = \int\limits_{\mathbb{R}^m} \left(\int\limits_{\mathbb{R}^l} f(x,y)\,dx \right) dy\,.$$

Remark 2.8.14 *In these two results it is to be understood that when we write down repeated integrals that an implicit assertion is that the functions $y \mapsto \int f(x,y)dx$ and $x \mapsto \int f(x,y)dy$ are measurable and integrable.*

Theorem 2.8.15 (Minkowski inequality) *If f is a measurable function $f : \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}$ is measurable, then*

$$\| \int\limits_{\mathbb{R}^m} f(x,y)g(y)\,dy \|_{L^p(dx)} \le \int\limits_{\mathbb{R}^m} \|f(x,y)\|_{L^p(dx)} |g(y)|\,dy\,. \tag{2.12}$$

where

$$\|f(x,y)\|^p_{L^p(dx)} = \int\limits_{\mathbb{R}^l} |f(x,y)|^p\,dx\,,$$

with the understanding as above that this means that if the right hand side of (2.12) *is finite then the function $f(x,y)g(y)$ is integrable in y for almost every x and the resulting function $x \mapsto \int f(x,y)g(y)\,dy$ is measurable and* (2.12) *holds.*

2.9 Worked problems

1. Prove that if a continuous 2π-periodic function $f \in C_{per}([-\pi,\pi])$ satisfies

   $$\hat{f}(m) = (2\pi)^{-1} \int_{-\pi}^{+\pi} e^{-imx} f(x)dx = 0$$

   for all $m \in \mathbb{Z}$, then f is identically zero. Deduce that if $f \in C^\infty_{per}([-\pi,\pi])$ then $f = \sum_{m\in\mathbb{Z}} \hat{f}(m)e^{imx}$.

Answer Assume for the sake of contradiction that there exists x_0 with $f(x_0) \neq 0$. Replacing $f(\cdot)$ by $\pm f(\cdot - x_0)$ we may assume w.l.o.g. that $f(0) > 0$. Now:

$$\int_{-\pi}^{+\pi} f(x)(\epsilon + \cos x)^k \, dx = 0$$

for all $k \in \mathbb{Z}_+$ and $\epsilon \in \mathbb{R}$, by the assumption that all Fourier coefficients vanish. By continuity of f there exists $\delta \in (0, \pi/2]$ such that $f(x) > f(0)/2 > 0$ for $|x| < \delta$. Now $\max_{\delta \leq |x| \leq \pi} \cos x < 1$ and $\cos 0 = 1$, so there exists

- $\epsilon > 0$ such that $|\epsilon + \cos x| < 1 - \epsilon/2$ for $\delta \leq |x| \leq \pi$;
- $\eta \in (0, \delta)$ such that $|\epsilon + \cos x| > 1 + \epsilon/2$ for $|x| < \eta$.

Now

$$\begin{aligned} \int_{-\pi}^{+\pi} f(x)(\epsilon + \cos x)^k \, dx \;&=\; \int_{|x| \leq \eta} + \int_{|\eta| < |x| < \delta} + \int_{\delta \leq |x| \leq \pi} f(x)(\epsilon + \cos x)^k \, dx \\ &\geq\; (1 + \frac{\epsilon}{2})^k \frac{f(0)}{2} - 2\pi \sup |f| (1 - \frac{\epsilon}{2})^k \,, \end{aligned}$$

since the middle integral is ≥ 0 because $f > 0$ on $(-\delta, +\delta)$, and also since $\delta \leq \pi/2$ and $\cos x \geq 0$ on $[0, \pi/2]$. Now let $k \to +\infty$: the final term has limit zero, while the first term has limit $+\infty$ providing a contradiction.

For the last part observe that for $f \in C_{per}([-\pi, \pi])$ the Fourier coefficients satisfy

$$\sup_m m^N |\hat{f}(m)| < \infty$$

for all N (rapidly decreasing) and therefore the series $\sum_{m \in \mathbb{Z}} \hat{f}(m) e^{imx}$ converges absolutely to define a continuous periodic function whose Fourier coefficients are $\hat{f}(m)$. (The latter assertion follows from the fact that the sum and integral can be interchanged when integrating an absolutely and uniformly convergent power series.) Therefore $f(x) - \sum_{m \in \mathbb{Z}} \hat{f}(m) e^{imx}$ is a continuous 2π-periodic function whose Fourier coefficents all vanish. It therefore vanishes itself by the previous part, completing the proof.

2. For positive s the Sobolev space is defined as

$$H^s(\mathbb{R}^n) \;=\; \{ f \in L^2(\mathbb{R}^n) : \|f\|^2_{H^s} = \int_{\mathbb{R}^n} (1 + \|\xi\|^2)^s |\hat{f}(\xi)|^2 \, d\xi < \infty \} \,.$$

Show that if $s > n/2$ then $H^s(\mathbb{R}^n) \subset C(\mathbb{R}^n)$ and there exists a positive number C such that

$$\sup_{x \in \mathbb{R}^n} |f(x)| \leq C \|f\|_{H^s} \,. \tag{2.13}$$

In the case $n = 1$ prove, using calculus only, the inequality

$$\sup_{x \in \mathbb{R}} |f(x)| \leq C' \left(\int (f^2 + (\partial_x f)^2) \, dx \right)^{\frac{1}{2}} . \tag{2.14}$$

for all $f \in \mathcal{S}(\mathbb{R})$ and for some positive C'. Comment on the relation with the first part of the question. Prove that all functions $f \in H^1(\mathbb{R})$ are uniformly continuous on $\mathbb{R}$.

Answer We will first establish the inequality for $f \in \mathcal{S}(\mathbb{R}^n)$. By the Hölder inequality:

$$(2\pi)^n \big|\mathcal{F}^{-1}(\hat{f})\big| = \left|\int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi\cdot x}\, dx\right| = \left|\int_{\mathbb{R}^n} \frac{(1+\|\xi\|^2)^{\frac{s}{2}}\hat{f}(\xi)e^{i\xi\cdot x}}{(1+\|\xi\|^2)^{\frac{s}{2}}}\, dx\right|$$
$$\leq \left(\int_{\mathbb{R}^n} \frac{1}{(1+\|\xi\|^2)^s}\, dx\right)^{\frac{1}{2}} \|f\|_{H^s}\,.$$

The integral on the second line is finite, since adopting polar coordinates (r, Ω) it is just

$$\int_{S^{n-1}} d\Omega \int_0^\infty \frac{r^{n-1}}{(1+r^2)^s} dr$$

which is finite for $2s - n + 1 > 1$, i.e. for $s > n/2$. This establishes the stated inequality (2.13). Exactly the same calculation shows that $\|\hat{f}\|_{L^1} \leq C\|f\|_{H^s}$. Now to complete the proof, just approximate $f \in H^s(\mathbb{R}^n)$ by a sequence f_ν of Schwartz functions as in Theorem 2.7.1: since the constant in (2.13) is independent of ν we can take the limit $\nu \to \infty$. Then since $\|f - f_\nu\|_{H^s} \to 0$ the sequence f_ν is Cauchy in the norm H^s, and hence also Cauchy in the uniform C^0 norm by (2.13). This implies that the limit $f \in C^0(\mathbb{R}^n)$ and obeys (2.13).

For the one dimensional inequality calculate, by the Hölder inequality that

$$|f(x) - f(y)| = |\int_y^x f'(z)dz| \leq |x-y|^{\frac{1}{2}} \left|\int (f')^2 dz\right|^{\frac{1}{2}}$$

and

$$|f^2(x) - f^2(y)| = |2\int_y^x f(z)f'(z)dz| \leq \int_{\mathbb{R}} f^2 + (f')^2\, dz\,.$$

For $f \in \mathcal{S}(\mathbb{R})$ let $y \to +\infty$ in the second inequality and the result follows for such f. It holds for general $f \in H^1(\mathbb{R})$ by density (strictly speaking up to sets of measure zero). The first inequality implies uniform continuity.

Relation with the first part of the question: the Parseval-Plancherel theorem implies that:

$$\int_{\mathbb{R}} f^2 + (f')^2\, dz = \frac{1}{2\pi}\int_{\mathbb{R}} (|\hat{f}(\xi)|^2 + |\xi|^2|\hat{f}(\xi)|^2)\, d\xi\,.$$

So that the result of the second part is really a special case of the first part, but the proof is different.

2.10 Example sheet 2

1. Obtain and solve the ODE satisfied by characteristic curves $y = y(x)$ for the equation $(x^2 + 2)^2 u_{xx} - (x^2+1)^2 u_{yy} = 0$. Show that there are two families of such curves which can be written in the form $y - x + 2^{-\frac{1}{2}} \arctan 2^{-\frac{1}{2}} x = \xi$ and $y + x - 2^{-\frac{1}{2}} \arctan 2^{-\frac{1}{2}} x = \eta$, for arbitrary real numbers ξ, η. Now considering the change of coordinates $(x, y) \to (\xi, \eta)$ so determined find the form of the equation in the coordinate system (ξ, η).

2. Which of the following functions of x lie in Schwartz space $\mathcal{S}(\mathbb{R})$: (a) $(1 + x^2)^{-1}$,(b) e^{-x}, (c)$e^{-x^4}/(1 + x^2)$? Show that if $f \in \mathcal{S}(\mathbb{R})$ then so is $f(x)/P(x)$ where P is any strictly positive polynomial (i.e. $P(x) \geq \theta > 0$ for some real θ.

3. Show that any $\hat{u} = \{\hat{u}_m\}_{m\in\mathbb{Z}}$ in $l^p_s(\mathbb{Z})$ defines (for any $p \geq 1, s \in \mathbb{R}$) a periodic distribution via the formula

$$F_{\hat{u}}(\varphi) \;=\; 2\pi \sum_{m\in\mathbb{Z}} \hat{u}_{-m}\hat{\varphi}(m) \tag{2.15}$$

for each test function $\varphi = \sum \hat{\varphi}(m)e^{imx} \in C^\infty_{per}([-\pi,\pi])$. Show that if $\hat{u} \in s(\mathbb{Z})$, then this distribution agrees with the periodic distribution T_u determined in the usual way by the the smooth function $u = \sum \hat{u}_m e^{imx}$.

4. Use Fourier series to solve the following initial value problem

$$\partial_t u = \partial_x^3 u \qquad u(0,x) = f(x)$$

for $x \in [-\pi,\pi]$ with periodic boundary conditions $u(t,-\pi) = u(t,\pi)$ and f smooth and 2π-periodic. Discuss well-posedness properties of your solutions for the L^2 norm, i.e. $\|u(t)\|_{L^2} = (\int_{-\pi}^{+\pi} |u(t,x)|^2 dx)^{\frac{1}{2}}$, using the Parseval-Plancherel theorem. Extend your result to the H^s norm.

*Show that if $f \in L^2([-\pi,\pi])$ your Fourier series formula gives a distributional solution of the equation, in a precise sense which you should define.

5. Show that the heat equation $\partial_t u = \partial_x^2 u$, with 2π-periodic boundary conditions in x, is well-posed forwards in time in L^2 norm, but not backwards in time (even locally). (Hint compute the L^2 norm of solutions u_n for negative t corresponding to initial values $u_n(0,x) = n^{-1}e^{inx}$.)

6. (i) Use Fourier series to solve the Schrödinger equation

$$\partial_t u = i\partial_x^2 u \qquad u(0,x) = f(x)$$

for initial value f smooth and periodic. Prove in two different ways that there is only one smooth periodic solution.

(ii) Use the fourier transform to solve the Schrodinger equation for $x \in \mathbb{R}$ and initial value $f \in \mathcal{S}(\mathbb{R})$. Find the solution for the case $f = e^{-x^2}$.

7. (i) Verify that the tempered distribution u on the real line defined by the function $K_m(x) = (2m)^{-1}e^{-m|x|}$, (for positive m), solves

$$\left(\frac{-d^2}{dx^2} + m^2\right)u = \delta_0$$

in $\mathcal{S}'(\mathbb{R})$. For which $s \in \mathbb{R}$ does the function K_m lie in $H^s(\mathbb{R})$? For which $s \in \mathbb{R}$ does the δ-function lie in $H^s(\mathbb{R})$?[Use the definition (2.9).]

(ii) Verify that the function on the real line $g(x) = 1$ for $x \leq 0$ and $g(x) = e^{-x}$ for $x > 0$ defines a tempered distribution T_g which solves in $\mathcal{S}'(\mathbb{R})$

$$T'' + T' = -\delta_0.$$

8. (a) Write down the precise distributional meaning of the equation

$$-\Delta(|x|^{-1}) = 4\pi\delta_0 \qquad \text{in } \mathcal{S}'(\mathbb{R}^3)$$

in terms of test functions, and then use the divergence theorem to verify that it holds. (Hint: apply the divergence theorem on the region $\{0 < |x| < R\} - \{0 < |x| < \epsilon\}$ for R sufficiently large and take the limit $\epsilon \to 0$ carefully).

(b) Find the fundamental solution $G_m \in \mathcal{S}'(\mathbb{R}^3)$ of the operator $(-\Delta + m^2)$ with $m > 0$ and in the case of domain $\mathbb{R}^3$. Indicate the modifications of (a) required to prove this.

9. For each of the following equations, find the most general tempered distribution T which satisfies it.

$$xT = 0, \quad xdT/dx = 0, \quad x^2T = \delta_0, \quad xdT/dx = \delta_0$$

$$dT/dx = \delta_0, \quad dT/dx + T = \delta_0 \quad T - (d/dx)^2 T = \delta_0.$$

(Hint: see Friedlander §2.7).

10. Solve the equation $x^m T = 0$ in $\mathcal{S}'(\mathbb{R})$.