Notes on Functional Analysis

Alexander Terenin

Table of Contents

Chapter 1

Elements of Metric Topology

1.1. Definitions of Metric Topology

Metric space

Definition 1. *Let X be a set. A metric M on X is a function $d : X \times X \to \mathbb{R}$ such that*

1. $d(x, y) \geq 0$ *for all* $x, y \in X$.
2. $d(x, y) = 0 \iff x = y$
3. $d(x, y) = d(y, x)$
4. *Triangle Inequality:* $d(x, z) \leq d(x, y) + d(y, z)$.

Call X a METRIC SPACE.

Open ball

Definition 2. *Let (X, d) be a metric space. For all $x \in X, r > 0$, define*

$$B(x, r) = \{y \in X : d(x, y) < r\}. \tag{1.1}$$

Call B BALL *entered at x of radius r.*

Open set

Definition 3. *Let $V \subset X$. We say that V is open if and only if for all x, there exists an $r > 0$ such that $B(x, r) \subset V$.*

The above means that a set is *open* if, for every point, there exists at least one ball fully contained inside that set.

Metric topology

DEFINITION 4. *Consider the set*

$$O = \{V \in P(X) : V \text{ open}\} \subset P(X) \tag{1.2}$$

where $P(X)$ is the power set of X. Call O the TOPOLOGY *induced by M.*

Topology

DEFINITION 5. *For a given set X, a topology is a set O satisfying the following properties.*

1. *$\varnothing \in O$. ($\iff$ $\varnothing$ is open.)*
2. *$X \in O$. ($\iff$ X is open.)*
3. *$(V_i)_{i\in I}$, $V_i \in O$, for all $i \in I$, implies $\bigcup_{i\in I} V_i \in O$ (arbitrary unions are open)*
4. *$(V_i)_{i\in I}$, $V_i \in O$, for all $i \in I$ for I finite, implies $\bigcap_{i\in I} \in O$ (finite intersections are open)*

Call (X, O) a topological space.

EXAMPLE 6. *Take R, $d(x, y) = |y - x|$. Every interval (a, b) is open. On the other hand, take*

$$\bigcap_{n\in\mathbb{N}^+} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\} \tag{1.3}$$

which is not open. However, take

$$\bigcup_{n\in\mathbb{N}^+} \left(0, 1 - \frac{1}{n}\right) = (0, 1) \tag{1.4}$$

which is open.

As shown before, every metric space induces a topology. However, not every topology induces a metric.

Definition 7. *If a topological space X induces a metric, we say that X is metrizable.* Metrizable topology

The distance induced by a metrizable topology is not necessarily unique.

Example 8. *Take $\mathbb{R}^d$. Define*

$$d_1(x,y) = \sum_{i=1}^{d} |x_i - y_i| \quad d_2(x,y) = \sqrt{\sum_{i=1}^{d} (x_i - y_i)^2} \quad d_\infty(x,y) = \max_i |x_i - y_i|. \tag{1.5}$$

These define different balls – on R^2 d_1 looks like a circle, d_2 looks like a diamond, d_∞ looks like a square. Nonetheless, all three induce the same topology. We can draw a circle both outside and inside a given diamond – this property is enough to show they induce the same topology.

Definition 9. *Let $F \subset X$ where X is a topological space. We say F is closed if $F^C = X \setminus F$ is open.* Closed set

Some properties of closed sets.

1. $\varnothing$ is closed. (It is also open.)
2. X is closed. (It is also open.)
3. Spaces where X and $\varnothing$ are the only sets that are both closed and open are called connected.
4. $(F_i)_{i \in I}$, F_i closed for all $i \in I$, implies $\cap_{i \in I} V_i \in O$ (arbitrary intersections of closed sets are closed).
5. $(F_i)_{i \in I}$, F_i open for all $i \in I$ for I finite, implies $\bigcup_{i \in I} \in O$ (finite unions of closed sets are closed).

Definition 10. *For $x \in X$, $W \subset X$, W is a neighborhood of x if and only if there exists an open set V such that $x \in V \subset W$.* Neighborhood

Example 11. *The previous distances extend to the complex numbers – d_1, d_2, d_∞ are well-defined on $\mathbb{C}$ if absolute value is replaced by the complex modulus, and all induce the same topology.*

Interior

Definition 12. *Let $A \subset X$. Define*

$$A^\circ = \{x \in A : A \text{ is a neighborhood of } X\} = \{x \in A : \text{ there exists} r > 0, B(x,r) \subseteq A\} \tag{1.6}$$

Call A° the interior of A. A° is open.

Closure

Definition 13. *Let $A \subset X$. Define*

$$\overline{A} = \{x \in A : \text{ for all } r > 0, B(x,r) \cap A \neq \varnothing\}. \tag{1.7}$$

Call $\overline{A}$ the closure of A. $\overline{A}$ is closed. We have $A^\circ \subset A \subset \overline{A}$.

Example 14. *Consider the set*

$$A = \{(x,y) : x^2 + y^2 < 1\}. \tag{1.8}$$

The interior is the ball without the circle on the boundary. The closure is the ball with the circle on the boundary.

Dense set

Definition 15. *A set $A \subseteq X$ is dense if and only if $\overline{A} = X$, i.e. if and only if*

$$\text{for all } x \in X, r > 0, \text{ we have } B(x,r) \cap A \neq \varnothing. \tag{1.9}$$

Example 16. *The set $\mathbb{Q}$ is dense in $\mathbb{R}$, because for any set $(x-r, x+r)$ there exists a set $y \in Q$ such that $y \in (x-r, x+r)$.*

Bounded set

Definition 17. *A set A is bounded if and only if there exists $x \in X$, $r > 0$ such that $A \subset B(x,r)$.*

1.2. Convergence

Sequence

Definition 18. *A sequence $(x_n)_{n \geq 1}$ in X is a map $\mathbb{N} \to X$.*

Definition 19. *A sequence $(x_n)_{n\geq 1}$ is convergent if and only if there exists $x \in X$ such that $d(x_n, x) \to 0$. The latter means that for every $\varepsilon > 0$ there exists an integer $n_0 \geq 1$ such that for all $n \geq n_0$ we have $d(x_n, x) \leq \varepsilon$.* Convergence

Proposition 20. *For a metrizable topology, if $(x_n)_{n\geq 1}$ is convergent, then its limit x is unique. Call*

$$x = \lim_{n\to\infty} x_n \tag{1.10}$$

the limit of the sequence.

Remark 21. *The above is not true in general – in a general topological space, we at minimum need a Hausdorff property to guarantee uniqueness of the limit.*

Definition 22. *Let $A \subset X$. Define*

$$d_A(x, y) = d(x, y) \text{ for } (x, y) \in A \times A. \tag{1.11}$$

Call d_A the restriction of M to $A \times A$, i.e. the induced distance on A by M.

Proposition 23. *Let $U \subset A$, and let d_A be the restriction of M to $A \times A$. U is an open set of (A, d_A) if and only if there exists an open set V of (X, d) such that $U = V \cap A$.*

Example 24. *Let $A \in [0, 1]$ and let $d(x, y) = |x - y|$. The set*

$$\left(\frac{1}{x}, 1\right] \tag{1.12}$$

is an open set of $[0, 1]$, but not an open set of $\mathbb{R}$.

Proposition 25. *G is closed for (A, d_A) if and only if there exists F closed in (X, d) and $G = F \cap A$*

1.3. Continuity

Continuous at a point

Definition 26. *Define a function* $(X, d) \xrightarrow{f} (Y, D^*)$. *We say that* f *is continuous at* x *if for all neighborhoods* W *of* $f(x)$ *in* Y, *there exist a neighborhood* U *of* $x \in X$ *such that* $f(U) \subset W$ *for*

$$f(U) = \{y \in Y : \text{ there exists } x \in U \text{ s.t. } f(x) = y\}. \tag{1.13}$$

This means if we choose a neighborhood W near a point $f(x)$, we need to be able to find a neighborhood of x that is mapped to inside W. In a metric space, we can take neighborhoods to be balls, which provides good intuition.

Example 27. *Translating the above with* ε-δ, *we can write for all* $\varepsilon > 0$ *(with* $W = B(f(x), \varepsilon)$*) there exists a* $\delta > 0$ *(*$U = B(x, \delta)$*) such that for all* $y \in B(x, \delta)$ *we have* $f(x) \in B(f(x), \varepsilon)$ *(*$f(B(x, \delta)) \subset B(f(x), \varepsilon)$*). Translating again using distances, this means for all* $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that for all* y *with* $d(x, y) < \delta$ *we have* $d(f(x), f(y)) < \varepsilon$.

Proposition 28. *The composition of continuous functions at a given point are continuous at that point.*

Continuous function

Definition 29. *A function* f *is* continuous *if and only if* f *is continuous at* x *for all* $x \in X$.

Proposition 30. *A function* f *is continuous if and only if for any open set* V *of* Y, *the preimage*

$$f^{-1}(V) = \{x \in X : f(x) \in V\} \tag{1.14}$$

is open in X.

If $f : X \xrightarrow{f} Y$ is continuous, and U is an open set of X, then $f(U)$ may not be an open set of Y. This means that for a continuous function the image of an open set need not be open – whereas the preimage must be.

Proposition 31. *A function f is continuous if and only if for all G closed in Y, $f^{-1}(G)$ is closed in X.*

Definition 32. *A function $f : (X, d) \to (Y, D^*)$ is a homeomorphism if and only if f is bijective, f is continuous, and f^{-1} is continuous.* Homeomorphism

The above means that if we know the open sets of one space, we immediately know the open sets of the other, which enables us to determine whether two topologies are, in an appropriate sense, identical.

Definition 33. *A function $f : (X, d) \to (Y, D^*)$ is an isometry if and only if f is bijective, and $D^*(f(x), f(y)) = d(x, y)$ for all $x, y \in X$.* Isometry

Remark 34. *An isometry is a much stronger than a homeomorphism – the identity mapping from $(\mathbb{R}^d, d_1)$ to $(\mathbb{R}^d, d_2)$ is a homeomorphism but not an isometry.*

Definition 35. *A function f between two metric spaces (X, d) and (Y, D^*) is uniformly continuous if and only if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in X$ and all $y \in X$, $d(x, y) < \delta)$ implies $d(f(x), f(y)) < \varepsilon$.* Uniform continuity

Recall that f is continuous if f is continuous at every point, i.e. for all $x \in X$ and all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $y \in X$, $d(x, y) < \delta$ implies $d(f(x), f(y)) < \varepsilon$. Uniform continuity means we must use the same δ for all x, rather than a different δ for every x. This is clearly a stronger property – f uniformly continuous implies f continuous, but the converse is false.

Definition 36. *For $k > 0$, a function f between two metric spaces (X, d) and (Y, D^*) is k-Lipschitz if and only if* Lipschitz continuity

$$D^*(f(x), f(y)) \leq kd(x, y) \tag{1.15}$$

for all $x, y \in X$.

Proposition 37. *If f is k-Lipschitz, then f is uniformly continuous.*

Completeness

Definition 38. *Let (X, d) be a metric space. Define a sequence $(x_n)_{n\geq 1}$. x_n is Cauchy if and only if for every $\varepsilon > 0$, there exist an $n_0 \in \mathbb{N}$ such that for all $n, p > n_0$, we have $d(x_n, x_p) < \varepsilon$.*

Proposition 39. *The following hold.*

1. *$(x_n)_{n=1}^{\infty}$ is Cauchy $\implies$ $(x_n)_{n=1}^{\infty}$ is bounded.*
2. *$(x_n)_{n=1}^{\infty}$ is convergent $\implies$ $(x_n)_{n=1}^{\infty}$ is Cauchy (the converse is not true in general).*
3. *Let $(x_n)_{n=1}^{\infty}$ is Cauchy and let $(x_{n_k})_{k=1}^{\infty}$ be a convergent subsequence ($x_{n_k} \to x$ as $k \to \infty$). Then $(x_n)_{n\geq 1}$ is Cauchy is Cauchy*

Complete metric space

Definition 40. *(X, d) is complete if and only if every Cauchy sequence is convergent.*

Complete subspace

Definition 41. *Let (X, d) be a metric space and $A \subset X$. A is complete if the metric space (A, d_A) is complete where $d_A : A \times A \to \mathbb{R}$ where $d_A(x, y) = d(x, y)$. {missed}*

Proposition 42. *Let $A \subset (X, d)$. Then the following hold.*

1. *A complete implies A closed in X.*
2. *If X is complete then A is complete if and only if A is closed in X.*

Example 43. *$\mathbb{R}, \mathbb{C}, \mathbb{R}^d, \mathbb{C}^d$ are complete. $\mathbb{Q}$, defined as ratios of integers, is a field, but is not complete – there is no root in $\mathbb{Q}$ to the equation $x^2 = 2$. $\mathbb{R}$ is constructed as the completion of $\mathbb{Q}$.*

Fixed point theorem

Theorem 44. *Let (X, d) be a complete metric space. Let $f : X \to X$ be a strict contraction, in the sense that there exists a k, $0 < k < 1$, such that*

$d(f(x), f(y)) < kd(x, y)$ for all x, y. Then there exists a unique fixed point $x^* \in X$, i.e. $f(x^*) = x$.

Proof. Sketch. Let $x_0 \in X$. Construct $(x_n)_{n\geq 1}$ be recursion with $x_{n+1} = f(x_n)$. Then $(x_n)_{n\geq 1}$ is Cauchy. Use that $\sum_{n=1}^{\infty} k^n < \infty$ because $k < 1$. ■

Example 45. *There are sequences which are non-strictly-contractive and have no fixed point.*

Example 46. *Let $u'(t) = f(u(t), t)$ with $u \in \mathbb{R}, f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, and $u(0) = u_0$ given. Suppose we have $|f(u, t) - f(v, t)| < L|u - v|$. Write $u(t) = u_0 + \int_0^t f(u(s), s)\,\mathrm{d}s$, we have a map $F : u \to v$ for $u \in C^0([0, T])$, $v \in C^0([0, T])$, defined by $v(t) = u_0 + \int_0^t f(u(s), s)\,\mathrm{d}s$. Letting $u_1, u_2 \in C([0, T])$ and $v_1 = F(u_1), v_2 = F(u_2)$, we get* Cauchy-Lipschitz Theorem

$$
\begin{aligned}
|(v_2 - v_1)(t)| &= \left| \int_0^t (f(u_2(s), s) - f(u_1(s), s))\,\mathrm{d}s \right| && (1.16)\\
&\leq \int_0^t |f(u_2(s), s) - f(u_1(s), s)|\,\mathrm{d}s && (1.17)\\
&\leq L \int_0^t |u_2(s) - u_1(s)|\,\mathrm{d}s && (1.18)\\
&\leq L \int_0^T |u_2(s) - u_1(s)|\,\mathrm{d}s && (1.19)\\
&\leq LT \max_{s\in[0,T]} |u_2(s) - u_1(s)|. && (1.20)
\end{aligned}
$$

Letting $d_\infty(u_1, u_2) = \max_{s\in[0,T]} |u_2(s) - u_1(s)|$, we see that

$$d_\infty(v_1, v_2) = \max_{t\in[0,T]} |v_2(t) - v_1(t)| \leq LT \max_{t\in[0,T]} |u_2(t) - u_1(t)| = d_\infty(u_1, u_2) \tag{1.21}$$

and if $LT < 1$, given $(C^0([0, T]), d_\infty)$ is complete, we may apply the fixed point theorem and see that there exists a unique solution to the given differential equation on $[0, T]$.

1.4. Compactness

Definition 47. *Let (X, d) be a metric space. $K \subset X$ is said to be* Compact set *compact if and only if it satisfies the Borel-Lebesgue Property.*

Borel-Lebesuge Property

Definition 48. *For all coverings of K by open sets $\{O_i\}_{i\in I}$, i.e. for all $i \in I$, where O_i is open in X, with $K \subset \bigcup_{i\in I} O_i$, there exists $J \subset I$ such that J is finite and $\{O_j\}_{j\in J}$ is a covering of K, in the sense that $K \subset \bigcup_{j\in J} O_j$.*

Part of the motivation for the above property is that if we have an infinite sequence over a finite set, the infinite sequence must repeat at least one element an infinite number of times, so it must contain an infinite constant subsequence.

Finite covering by balls of arbitrary radius

Definition 49. *Suppose that $K \subset X$ has the property of finite covering by balls of arbitrary radius, in the sense that for every $\varepsilon > 0$ there exists N_ε and there exists $x_1, .., x_{N_\varepsilon} \in K$ such that $K \subset \bigcup_{k=1}^{N_k} B(x_k, \varepsilon)$.*

As before, since N_k is finite, we expect

Proposition 50. *The Borel-Lebesgue property implies finite covering by balls of arbitrary radius.*

Proof. Fix ε. For all $x \in K$, consider $B(x, \varepsilon)$ which are open sets. Clearly, $K \subset \bigcup_{x\in K} B(x, \varepsilon)$, so K is a covering of X. By the Borel-Lebesgue property, there exists N_ε and a finite collection $x_1, .., x_{N_\varepsilon}$ such that $K \subset \bigcup_{k=1}^{N_\varepsilon} B(x_k, \varepsilon)$. ■

Proposition 51. *Let (X, d) be a metric space. If $K \subset X$, K compact implies K closed and bounded.*

Remark 52. *The converse is not true in general – however, it does hold for a finite-dimensional normed vector space.*

Proposition 53. *Let (X, d) be compact. If $F \in X$ is closed, then F is compact.*

Proposition 54. *Suppose* $f : (X, d) \to (Y, D^*)$ *is continuous. Let* $K \subset X$ *be compact. Then* $f(K)$ *is compact in* Y*.*

Proposition 55. *Suppose* (X, d) *is compact. Then if* $f : (X, d) \to (Y, D^*)$ *is continuous,* f *is uniformly continuous.*

Proposition 56. *Suppose* X *is compact. Let* f *be a bijection between* X *and* Y*. Then* f *is a homeomorphism – i.e.* f^{-1} *is continuous, and we can conclude that* Y *is compact.*

Theorem 57. *Let* (X, d) *be a metric space and* $K \subset X$*. The following are equivalent.* Bolzano-Weierstrass Theorem

1. K *is compact in* X*.*
2. *Given any sequence* $(x_n)_{n\geq 1}$ *such that* $x_n \in K$ *for all* $n \geq 1$*, there exists a subsequence* $(x_{n_k})_{k\geq 1}$ *and* $x \in K$ *such that* $x_{n_k} \to x$ *as* $k \to \infty$*.*

Example 58. $[a, b] \in \mathbb{R}$ *is compact. This is proven by using the reverse implication of the Bolazno-Weierstrass Theorem, by constructing a convergent subsequence through considering a sequence of nested intervals cut in half and constructing the sequence using a diagonal argument.*

Proposition 59. *Consider* $f : X \to \mathbb{R}$ *continuous for* (X, d) *compact. Then* f *is bounded, i.e. there exists* $c > 0$ *such that* $|f(x)| \leq c$ *for all* $x \in X$*. Furthermore, there exists two points* $x_0, x_1 \in X$ *such that* $f(x_0) = \min_{x\in X} f(x)$ *and* $f(x_1) = \max_{x\in X} f(x)$*.*

Example 60. *Take* (E, x) *compact. Prove that for all open coverings* $\{O_i\}_{i\in I}$ *of* E *there exists* $\delta > 0$ *such that for all* $x \in E$*, there exists* $i \in I$ *such that the ball* $B(x, \delta) \subset O_i$*.* Exercise

1. *Suppose the above is false. Show that there exists an open covering* $(O_i)_{i\in I}$ *of* E *and a sequence* $(x_n)_{n\geq 1}$ *such that* $x_n \in E$ *for all* $n \geq 1$ *and* $B(x_n, 1/n) \not\subset O_i$ *for all* $i \in I$*.*

- *The negation is as follows: there exists an open covering* $\{O_i\}_{i\in I}$ *of* E *such that for all* $\delta > 0$ *there exists* $x \in E$ *such that for all* $i \in I$, *the ball* $B(x,\delta) \not\subset O_i$. *Take* $\delta = 1/n$, *for all* $n \in \mathbb{N}$ *defines* x_n *for all* $n \geq 1$.

2. *Show that there exists* $(x_{n_k})_{k\geq 1}$ *and* $a \in E$ *such that* $x_{n_k} \to a$ *as* $k \to \infty$.

   - *We prove this by noting* E *is compact, and applying the Bolzano-Weierstrass Theorem.*

3. *Show that there exists* $i_0 \in I$ *and* $r > 0$ *such that* $B(a,r) \in O_{i_0}$.

   - *If* $a \in E = \bigcup_{i\in I} O_i$, *hence there exists* $i_0 \in I$ *such that* $a \in O_i$. *The set* O_{i_0} *is open, so by the definition of an open set, there exists an* r *such that* $B(a,r) \in O_{i_0}$.

4. *Using the previous part, show that there exists* $k_0 \in \mathbb{N}$ *such that for all* $k \geq k_0$, $d(a, x_{n_k}) < r/2$.

   - *We know that* $x_{n_k} \to a$ *as* $k \to \infty$, *so for every* $\varepsilon > 0$ *there exists a* $K > 0$ *such that for all* $k \geq K$ *we have* $d(a, x_{n_k}) < \varepsilon$. *Take the corresponding* K *for* k_0.

5. *Take* k_1 *such that* $1/n_{k_1} < r/2$. *Remark that for all* $k > k_1$, *we have* $1/n_k < r/2$. *Take* $k > \max\{k_0, k_1\}$. *Using previous part, show that* $B(x_{n_k}, 1/n_k) \subset B(a,r)$.

   - *Take a point* $x \in B(x_{n_k}, 1/n_k)$. *We want to show that* $x \in B(a,r)$, *so* $d(a,x) < r$. *We know* $d(x, x_{n_k}) < 1/n_k$, *so by the triangle inequality,* $d(a,x) \leq d(x, x_{n_k}) + d(x_{n_k}, a)$. *We know* $k > \max\{k_0, k_1\} \geq k_0$, *so* $d(x_{n_k}, a) < r/2$. *On the other hand,* $d(x, x_{n_k}) < 1/n_k < r/2$, *and adding the two inequalities completes the proof.*

Hence, there exists an $(x_n)_{n\geq 1}$ *such that* $B(x_n, 1/n) \not\subset O_i$ *for all* $i \in I$. *The conclusion of the last point contradicts the second one.*

Chapter 2

Lebesgue Integration Theory

If we use the Riemann integral to define a normed vector space with the Riemann-L^1 norm, this space is not complete and therefore not a Banach space. This was one of the motivations behind defining the Lebesgue integral in the early 20th century.

Definition 61. *Take a set X, and let $\mathcal{A} \subset \mathcal{P}(X)$. $\mathcal{A}$ is called a σ-algebra if and only if the following properties hold.* σ-algebra

1. $X \in \mathcal{A}$.
2. $A \in \mathcal{A} \implies A^C = X \setminus A \in \mathcal{A}$.
3. *If $(A_i)_{i=1}^{\infty}$ is a countable family in $\mathcal{A}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.*

Definition 62. *Let $\mu : \mathcal{A} \to [0, \infty]$ be a measure if and only if it satisfies the following.* Measure

1. $\mu(\varnothing) = 0$.
2. *If $(A_i)_{i=1}^{\infty}$ is a countable family in $\mathcal{A}$ such that, if $i \neq j$ then $A_i \cap A_j = \varnothing$, then $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.*

Definition 63. *The triple $(X, \mathcal{A}, \mu)$ is called a measure space.* Measure space

Lebesgue measure

Theorem 64. *Let $X = \mathbb{R}^d$ or $X = \mathbb{C}^d$ endowed with the standard topology. There exists a unique σ-algebra, called the σ-algebra of Lebesgue-measurable sets, and a unique measure μ on it, such that the following hold.*

1. *Every open set is measurable.*
2. $P = \bigtimes_{j=1}^{d}(a_j, b_j) \implies \mu(P) = \prod_{j=1}^{d}(b_j - a_j)$.
3. *If A is Lebesgue-measurable and $\mu(A) = 0$, then for all $B \subset A$, B is Lebesgue-measurable and $\mu(B) = 0$.*
4. *For all Lebesgue-measurable A, the set $A+x$, defined as the translation of A by $x \in \mathbb{R}^d$, is Lebesgue-measurable, and $\mu(A+x) = \mu(A)$.*

It follows immediately that every closed set is also measurable, as is every singleton, and every countable set. Let $\{x\}$ be a singleton. Then $\mu(\{x\}) = \mu(\bigtimes_{i=1}^{d}[x_i, x_i]) = \prod_{i=1}^{d}(x_i - x_i) = 0$. But then by countable additivity, every finite set also has measure zero, and every countable set has measure zero, so $\mu(\mathbb{Q}) = 0$.

Negligible set

Definition 65. *Let A be a Lebesuge-measurable set with $\mu(A) = 0$. Then A is called* NEGLIGIBLE.

Almost everywhere

Definition 66. *Let A be a Lebesgue-measurable set. A property $(P(x))$ is said to hold on A almost everywhere if and only if there exists a set B such that $(P(x))$ is true for $x \in B \subset A$ such that $\mu(A \setminus B) = 0$.*

{missed}

Proposition 67. *Let $A \subset \mathbb{R}^d$ be a measurable set. We have the following.*

1. *f is Lebesgue-measurable $\implies$ $|f|$ Lebesgue-measurable.*
2. *$(f_n)_{n\geq 1}$ Lebesgue-measurable $\implies$ $\sup_{n\geq 1} f_n$, $\inf_{n\geq 1} f_n$, $\limsup_{n\to\infty} f_n$, $\liminf_{\to\infty} f_n$ Lebesgue-measurable.*

3. *$f : A \to \mathbb{R}, g : A \to \mathbb{R}$ (not $\mathbb{R} \cup \{\infty\}$) Lebesgue-measurable $\implies$ $f + g, fg$ Lebesgue-measurable.*

Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. Then the following hold.

1. $\liminf_{n\to\infty} \sup_{n\geq 1}(\inf_{p\geq n} x_p) \in [-\infty, \infty]$.
2. $\limsup_{n\to\infty} \inf_{n\geq 1}(\sup_{p\geq n} x_p) \in [-\infty, \infty]$.

Note that $x_n = \inf_{p\geq n} x_p$ is an increasing sequence and $x_n = \sup_{p\geq n} x_p$ is a decreasing sequence. Therefore, either $\liminf_{n\to\infty} x_n = \pm\infty$ or $\liminf_{n\to\infty} x_n = \lim_{n\to\infty} x_n$. Similarly either $\limsup_{n\to\infty} x_n = \pm\infty$ or $\limsup_{n\to\infty} x_n = \lim_{n\to\infty} x_n$.

Definition 68. *Let (E, d) be a metric space and $(x_n)_{n\geq 1}$, $x_n \in E$. An element $a \in E$ is called a limit point of $(x_n)_{n\geq 1}$ iff there exists a subsequence $(x_{n_k})_{k\geq 1}$ such that $x_{n_k} \to a$ as $k \to \infty$.* Limit point

Proposition 69. $\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$

Proposition 70. *$(x_n)_{n\geq 1}$ converges $\iff$ $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$*

Proposition 71. *$(x_n)_{n\geq 1}$ converges $\iff$ $(x_n)_{n\geq 1}$ has a unique limit point.*

Proposition 72. *Let $(x_n)_{n\geq 1}$ be a sequence. Define*

$$L = \{l \in [-\infty, \infty] : l \text{ is a limit point of } (x_n)_{n\geq 1}\}. \tag{2.1}$$

Then $\liminf_{n\to\infty} x_n = \min L$ and $\limsup_{n\to\infty} x_n = \max L$.

Proposition 73. *The following hold.*

1. *$f : A \to \mathbb{R}$ continuous $\implies$ f Lebesgue-measurable.*
2. *$A \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}$, f Lebesgue-measurable, g continuous $\implies$ $f(g(\cdot))$ Lebesgue-measurable.*

Simple function

Definition 74. *Let A be Lebesgue-measurable set of $\mathbb{R}^d$. A simple function $s : A \to \mathbb{R}$ is a function such that there exists a finite collection of Lebesgue-measurable sets $(A_i)i = 1^n$ which are mutually disjoint, i.e. $A_i \cap A_j = \varnothing$, such that for all $i, j, i \neq j$, and a family of real numbers $(\alpha_i)_{i=1}^n$ such that*

$$s = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i} \qquad \mathbb{1}_A = \begin{cases} 1 & x \in A_i \\ 0 & x \notin A_i \end{cases}. \tag{2.2}$$

The function s is a Lebesgue-measurable function.

Lebesgue integral

Definition 75. *Let $s = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ with $\alpha_i \geq 0$ for all $i = 1, .., n$. Then we define*

$$\int_A s(x)\,\mathrm{d}x = \sum_{i=1}^{n} \alpha_i \mu(A_i) \in [0, \infty]. \tag{2.3}$$

Next, let $f : A \to [0, \infty]$ be a non-negative Lebesgue-measurable function. Then

$$\int_A f(x)\,\mathrm{d}x = \sup\Big\{ \int_A s(x)\,\mathrm{d}x : s \textit{ simple functions such that } 0 \leq s(x) \leq f(x) \text{ a.e. } x \in \tag{2.4}$$

For f non-negative Lebesgue-measurable on A, $\int_A s(x)\,\mathrm{d}x$ always exists as a element A of $[0, \infty]$.

Proposition 76. *For f non-negative Lebesgue-measurable on A, we have*

$$\int_A f(x)\,\mathrm{d}x < \infty \implies f(x) < \infty \text{ a.e. } x \in A \tag{2.5}$$

Lebesgue-integrable function

Definition 77. *Let $f : A \to [-\infty, \infty]$. We say f is Lebesgue-integrable iff*

$$\int_A |f(x)|\,\mathrm{d}x < \infty. \tag{2.6}$$

Then we define for $a \in \mathbb{R}$

$$a_+ = \max\{a, 0\} = \begin{cases} a & a \geq 0 \\ 0 & a < 0 \end{cases} \quad a_- = \max\{-a, 0\} = \begin{cases} 0 & a \geq 0 \\ |a| & a < 0 \end{cases} \tag{2.7}$$

so that we have for all $a \in \mathbb{R}$ that $a = a_+ - a_-$ and $|a| = a_+ + a_-$. Then, define

$$\int_A f(x)\,\mathrm{d}x = \int_A f_+(x)\,\mathrm{d}x - \int_A f_-(x)\,\mathrm{d}x. \tag{2.8}$$

Note for $a \in \mathbb{R}$ that $0 \le a_+ \le |a|$ and $0 \le a_- \le |a|$. Hence for any function f, $0 \le f_+ \le |f|$ and $0 \le f_- \le |f|$. For any f, g non-negative Lebesgue-measurable such that $0 \le f(x) \le g(x)$ a.e. in A, we have $0 \le \int_A f(x)\,\mathrm{d}x \le \int_A g(x)\,\mathrm{d}x$. Hence, we have $0 \le \int_A f_+(x)\,\mathrm{d}x \le \int_A |f(x)|\,\mathrm{d}x < \infty$ and $0 \le \int_A f_-(x)\,\mathrm{d}x \le \int_A |f(x)|\,\mathrm{d}x < \infty$ and we conclude $\int_A f(x)\,\mathrm{d}x \in \mathbb{R}$ – i.e. that the Lebesgue-integral of a Lebesgue-integrable function is finite.

We have

$$\int_0^1 \mathbb{1}_{\mathbb{Q}}(x)\,\mathrm{d}x = \mu(\mathbb{Q} \cap (0,1)) \tag{2.9}$$

by definition of simple functions. Since there exists a bijection $\mathbb{N} \to \mathbb{Q} \cap (0,1)$, there exists a sequence $(x_n)_{n=1}^{\infty}$ such that $\mathbb{Q} \cap (0,1) = \{x_n : x \in \mathbb{N}\} = \bigcup_{i=1}^{\infty} \{x_n\}$. So

$$\mu(\mathbb{Q} \cap (0,1) = \sum_{i=1}^{\infty} \mu(\{x_n\}) = \sum_{i=1}^{\infty} 0 = 0 \tag{2.10}$$

by countable additivity. Note that $\mathbb{1}_{\mathbb{Q}}$ is not a Riemann-integrable function.

Proposition 78. *The following hold.*

1. *Let A be a Lebesgue-measurable set of $\mathbb{R}^d$. Then*

$$\int_{\mathbb{R}}^{d} \mathbb{1}_A(x)\,\mathrm{d}x = \mu(A) \tag{2.11}$$

2. *Let f be Lebesgue-integrable. Then $-\infty < f(x) < \infty$ a.e. $x \in A$. This is clear from*

$$\int_A |f(x)|\,\mathrm{d}x < \infty \implies |f(x)| < \infty \tag{2.12}$$

so

$$\int_A f(x)\,\mathrm{d}x = \int f_+(x)\,\mathrm{d}x - \int f_-(x)\,\mathrm{d}x \tag{2.13}$$

so $|f(x)| = f_+(x) + f_-(x)$, $f_+, f_- \geq 0$. *Hence*

$$\int_A |f(x)|\,\mathrm{d}x = \int_A f_+(x)\,\mathrm{d}x - \int_A f_-(x)\,\mathrm{d}x \tag{2.14}$$

so it is clear that

$$\left|\int_A f_+(x)\,\mathrm{d}x - \int_A f_-(x)\,\mathrm{d}x\right| \leq \int_A f_+(x)\,\mathrm{d}x - \int_A f_-(x)\,\mathrm{d}x \tag{2.15}$$

3. *Let* f *be Lebesgue-integrable. Then*

$$\left|\int_A f(x)\,\mathrm{d}x\right| \leq \int_A |f(x)|\,\mathrm{d}x \leq \tag{2.16}$$

4. *If* f, g *Lebesgue-integrable on* A *and* $\alpha, \beta \in \mathbb{R}$ *then*

$$\int_A (\alpha f + \beta g)\,\mathrm{d}x = \alpha \int_A f(x)\,\mathrm{d}x + \beta \int_A g(x)\,\mathrm{d}x \tag{2.17}$$

5. *If* $f \geq 0$ *a.e. on* A *then*

$$\int_A f(x)\,\mathrm{d}x \geq 0. \tag{2.18}$$

6. *If* $f = 0$ *a.e. on* A *then*

$$\int_A f(x)\,\mathrm{d}x = 0. \tag{2.19}$$

7. *Let* $f \geq 0$ *a.e. on* A *and* $\int_A f(x)\,\mathrm{d}x = 0$. *Then* $f = 0$ *a.e. on* A.
8. *Let* f *be Lebesgue-integrable on* A *and for all* $B \subset A$, *we have* $\int_B f(x)\,\mathrm{d}x = 0$. *Then* $f(x) = 0$ *a.e. on* A.

Chapter 3

Function spaces

Definition 79. *For a Lebesgue-measurable set A of $\mathbb{R}^d$*

$$\mathcal{L}^1(A) = \{f : A \to \mathbb{R} : f \text{ Lebesgue-integrable}\} \tag{3.1}$$

forms a vector space. Define

$$f : \mathcal{L}^1(A) \to \int_A f(x)\,\mathrm{d}x \in \mathbb{R}. \tag{3.2}$$

This is a linear map. However, it does not induce a norm because the function $f(x)$

Definition 80. *Let $f \sim g$ iff $f = g$ a.e. on A iff there exists $B \subset A$ with $\mu(B) = 0$ such that $f(x) = g(x)$ for all $x \in A \setminus B$.*

Equivalence relation for almost everywhere

Proposition 81. *We have that $\sim$ is an equivalence relation.*

Definition 82. *Define the quotient space*

$$L^1(A) = \mathcal{L}^1(A)/\sim. \tag{3.3}$$

Hence

$$\tilde{f} \in L^1(A) \iff \tilde{f} = \{f : A \to \mathbb{R}\} \tag{3.4}$$

Lebesgue-measurable such that $f_1, f_2 \in \tilde{f}$ are such that $f_1 = f_2$ a.e. on A.

Proposition 83. *Suppose $\tilde{f} \in L^1(A)$ such that there exists $f \in \tilde{f}$ where f is continuous. Then f is unique.*

Unless otherwise stated, we always assume way take the continuous representative in the above class of functions.

For a generic $\tilde{f}$, we cannot evaluate $\tilde{f}$ at a given point. However, if $\tilde{f}$ contains a continuous representative f, we define $\tilde{f}(x) = f(x)$.

Define the linear form

$$f \in \mathcal{L}^1(A) \xrightarrow{I} \int_A f(x)\,\mathrm{d}x \in \mathbb{R}. \tag{3.5}$$

When $f \sim g$, we have that their integrals are identical. Hence, the map

$$f \in L^1(A) \xrightarrow{I} \int_A f(x)\,\mathrm{d}x \in \mathbb{R} \tag{3.6}$$

is well-defined, and defines a vector space.

Quotient by subspace

Definition 84. *If X is a vector space, M is a linear subspace of X, then we can define an equivalence relation $x \backsim y$ iff $x - y \in M$. Then if we define the space X/M by the quotient space $X/\backsim$, this is again a vector space.*

If we define the linear subspace M of $\mathcal{L}^1(A)$, then $f \in M$ iff $f(x) = 0$ a.e. in A. Then $L^1(A) = \mathcal{L}^1(A)/M$.

If $\tilde{f}, \tilde{g} \in L^1(A)$, then for $f \in \tilde{f}, g \in \tilde{g}$, then for $h = f + g$, $\tilde{h}$ does not depend on the choice of f, g, and hence defines the sum $\tilde{h} = \tilde{f} + \tilde{g}$.

The subspace M is the kernel of the map I. For a linear $I : X \to \mathbb{R}$, $M \subset \ker(I)$ one can define $\tilde{I} : X/M \to \mathbb{R}$ by $\tilde{x} \to \tilde{I}(\tilde{x}) = I(x)$ where x is any element of $\tilde{x}$.

L^1 norm

Definition 85. *Define*

$$\|f\|_{L^1} = \int_A |f(x)|\,\mathrm{d}x. \tag{3.7}$$

Proposition 86. *The $\|\cdot\|_{L^1}$ norm defines a norm on L^1.*

1. *Triangle inequality.*
2. $\|ax\|_{L^1} = a\|x\|_{L^1}$.
3. $f = 0$ a.e. $\iff$ $\|f\| = 0$.

Let A be an open set of $\mathbb{R}^d$, let $\tilde{f} \in L^1(A)$. Suppose that there exists $f \in \tilde{f}$ which is continuous on A, then we call it the *continuous representative*, it is unique, and we always take that continuous representative as representative of $\tilde{f}$.

For a continuous representative of $f \in L^1(A)$, $f(x)$ is well-defined. This definition of point evaluation extends to locally continuous functions on open sets.

Consider the Dirichlet problem

$$-\Delta u = u \qquad x \in M \qquad \Delta u = \sum_{k=1}^{d} \frac{\partial^2 u}{\partial {x_k}^2} \qquad u = g \text{ on } \delta M \tag{3.8}$$

for $u = g$ on δM. Then the boundary is a zero-measure set. It is not obvious how to combine the almost everywhere based definition of functions in L^1 with the need to prescribe boundary conditions – that this can always be done in appropriate function spaces is a nontrivial consequence of the *Trace Theorem.*

Definition 87. *Let $f : A \to \mathbb{C}$ for A Lebesgue-measurable and f Lebesgue-integrable. Define* Complex integration

$$\int_A f(x)\,\mathrm{d}x = \int_A \Re(f(x))\,\mathrm{d}x + i\int_A \Im(f(x))\,\mathrm{d}x. \tag{3.9}$$

3.1. Calculus with Lebesgue integrals

Theorem 88. *Let $(f_n)_{n\geq 1}$ for $f_n : A \to [0,\infty]$ with A, f_n Lebesgue-measurable. Suppose $n \leq m \implies f_n(x) \leq f_m(x)$ a.e. in A. Note that* Monotone Convergence Theorem

$f(x) = \lim_{n\to\infty} f_n(x)$ is defined in $[0,\infty]$ a.e. in A. Then f is Lebesgue-measurable and

$$\lim_{n\to\infty}\int_A f_n(x)\,\mathrm{d}x = \int_A \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \int_A f(x)\,\mathrm{d}x \tag{3.10}$$

so either one side of the identity is finite so that both are finite and equal, or both are $+\infty$. Furthermore, if both sides are finite then $\lim_{n\to\infty}\|f_n - f\|_{L^1} = 0$, so

$$\lim_{n\to\infty}\int_A |f_n(x) - f(x)|\,\mathrm{d}x = 0. \tag{3.11}$$

Fatou's Lemma

Lemma 89. *Take a sequence $(f_n)_{n\geq 1} : A \to [0,\infty]$ of Lebesgue-measurable functions. Then*

$$\int_A (\liminf_{n\to\infty} f_n)(x)\,\mathrm{d}x \leq \liminf_{n\to\infty}\int_A f_n(x)\,\mathrm{d}x. \tag{3.12}$$

Exercise 90. *Suppose we have f_1, f_2, then $\liminf\{f_1, f_2\} = \min\{f_1, f_2\}$. Draw a picture showing that*

$$\int_{\mathbb{R}} \liminf\{f_1, f_2\}(x)\,\mathrm{d}x \leq \liminf\left\{\int_{\mathbb{R}} f_1(x)\,\mathrm{d}x, \int_{\mathbb{R}} f_2(x)\,\mathrm{d}x\right\}. \tag{3.13}$$

Dominated Convergence Theorem

Theorem 91. *Let $(f_n)_{n\geq 1}$ be a sequence for $f_n : A \to \mathbb{R}$ (or $\mathbb{C}$) Lebesgue-integrable for all $n \geq 1$. Suppose the following.*

1. *$f_n(x) \to f(x)$ a.e. $x \in A$. (i.e. there exists Lebesgue-measurable $B \subseteq A$ with $\mu(B) = 0$ and $f_n(x) \to f(x)$ for all $x \in A \setminus B$.)*
2. *There exists $g : A \to [0,\infty]$, $g \in L^1(A)$ (i.e. $\int_A g(x)\,\mathrm{d}x < \infty$) such that $|f_n(x)| \leq g(x)$ a.e. $x \in A$ for all $n \geq 1$. Note that g is independent of n. Note also that $\int_A |f_n(x)|\,\mathrm{d}x \leq \int_A g(x)\,\mathrm{d}x < \infty$ does not imply $|f_n(x)| \leq g(x)$.*

Then the following hold.

1. *$f \in L^1(A)$.*
2. *$\int_A f_n(x)\,\mathrm{d}x \xrightarrow{n\to\infty} \int_A f(x)\,\mathrm{d}x$.*
3. *$||f_n - f||_{L^1} = \int_A |f_n(x) - f(x)|\,\mathrm{d}x \xrightarrow{n\to\infty} 0$.*

Proposition 92. *Let $(f_n)_{n\geq 1}$ Lebesgue-measurable on A with values in $\mathbb{R}$ or $\mathbb{C}$. Assume that*

$$\sum_{n=1}^{\infty}\int_A |f_n(x)|\,\mathrm{d}x < \infty. \tag{3.14}$$

Then the following hold.

1. *$\sum_{i=1}^{\infty} f_n(x)$ is absolutely convergent. (i.e. there exists $B \subseteq A$ with $\mu(B) = 0$ such that $\sum_{n=1}^{\infty}|f_n(x)|$ converges for all $x \in A \setminus B$.*
2. *$f(x) = \sum_{n=1}^{\infty} f_n(x)$ is defined a.e. $x \in A$, $f \in L^1(A)$, and $\int_A f(x)\,\mathrm{d}x = \sum_{n=1}^{\infty}\int_A f_n(x)\,\mathrm{d}x$. (Note that $\sum_{n=1}^{\infty}\int_A f_n(x)\,\mathrm{d}x$ is absolutely convergent because $\sum_{n=1}^{\infty}|\int_A f_n(x)\,\mathrm{d}x| \leq \sum_{n=1}^{\infty}\int_A |f_n(x)|\,\mathrm{d}x < \infty$ by assumption.)*

Proposition 93. *Let $f(t,x)$ be a function taking values in $I \times A$ where I is an interval of $\mathbb{R}$ and A is a Lebesgue-measurable set of $\mathbb{R}^d$. Suppose the following.*

1. *$t \to f(t,x)$ is continuous on I, a.e. $x \in A$.*
2. *$x \to f(t,x)$ is in $L^1(A)$ for all $t \in I$ (not almost all).*
3. *There exists a function $g \in L^1(A)$ (which is constant in t) such that $|f(t,x)| \leq g(x)$ for all $t \in I$ (not almost all), and a.e. $x \in A$.*

Then the function $t \to \int_A f(t,x)\,\mathrm{d}x$ is continuous.

Proof. Let $F(t) = \int_A f(t,x)\,\mathrm{d}x$. For all $t_0 \in I$, F is continuous at t_0. Recall that the function F is continuous at t_0 if and only if for any sequence of points $t_n \in I$, such that $t_n \to t_0$, we have $F(t_n) \to F(t_0)$. Take such a sequence t_n, so that $F(t_n) = \int_A f(t_n,x)\,\mathrm{d}x$. Define $f_n : x \to f(t_n,x)$. We have the following.

1. $f_n(x) \to f_\infty(x) = f(x,t_0)$ a.e. $x \in A$ by continuity of f w.r.t. t.
2. $|f_n(x)| = |f(t_n,x)| \leq g(x)$ for $g \in L^1(A)$ by assumption.

Hence, we can apply dominated convergence theorem and get $f_\infty \in L^1(A)$ so $\int_A f_\infty(x)\,\mathrm{d}x = \int_A f(t_0,x)\,\mathrm{d}x$ is well-defined and $\int_A f_n(x) \xrightarrow{n\to\infty} \int_A f_\infty(x)\,\mathrm{d}x$, so $\int_A f(t_n, x) \xrightarrow{n\to\infty} \int_A f_\infty(t_0, x)\,\mathrm{d}x$. Hence, F is continuous at the point t_0. ■

Proposition 94. *If $f_n, f \in L^1(A)$ and $\|f_n - f\|_{L^1(A)} \xrightarrow{n\to\infty} 0$ then there exists a subsequence $(f_{n_k})_{k\geq 1}$ such that $f_{n_k} \xrightarrow{n\to\infty} f$ a.e. $x \in A$, and $\|f_{n_k} - f\|_{L^1(A)} \xrightarrow{n\to\infty} 0$.*

Remark 95. *We cannot eliminate the subsequence from the previous proposition.*

Example 96. *Let $f : \mathbb{R} \to [0, \infty)$ such that $\int_{\mathbb{R}} f(x)\,\mathrm{d}x = c$, $0 < c < \infty$. For $\alpha > 0$, $n \in \mathbb{N}$, $n \geq 1$,*

$$I_n^\alpha = n \int_{\mathbb{R}} \ln\left[1 + \left(\frac{f(x)}{n}\right)^\alpha\right] \mathrm{d}x \in [0, \infty]. \tag{3.15}$$

Goal: deduce the limit $\lim_{n\to\infty} I_n^\alpha$.

- *First consider $\alpha \geq 1$. We can see $\left(\frac{f(x)}{n}\right)^\alpha \xrightarrow{n\to\infty} 0$, so $\ln\left[1 + \left(\frac{f(x)}{n}\right)^\alpha\right]$ is approximately $\left(\frac{f(x)}{n}\right)^\alpha$ as $n \to \infty$. So, I_n^α should be approximately $n \int_{\mathbb{R}} \left(\frac{f(x)}{n}\right)^\alpha \mathrm{d}x$, which is approximately $\frac{1}{n^{\alpha-1}} \int_{\mathbb{R}} f(x)^\alpha \,\mathrm{d}x$. This reasoning cannot work: we cannot assume this integral is finite, so these observations tell us nothing.*

- *Suppose now A is bounded Lebesgue-measurable in $\mathbb{R}^d$ with $\mu(A) < \infty$, $f : A \to \mathbb{R}$, then $\int_A |f(x)|^\alpha \,\mathrm{d}x < \infty \implies \int_A |f(x)|^\beta \,\mathrm{d}x$, $1 \leq \beta < \alpha$.*

- *For $\alpha = \infty$, if we know that $|f(x)| \leq c$, a.e. $x \in A$, so $\int_A |f(x)|\,\mathrm{d}x \leq c \int_A \mathrm{d}x = c\mu(A) < \infty$.*

- *For $\alpha \geq 1$, consider $g_n(x) n \ln\left[1 + \left(\frac{f(x)}{n}\right)^\alpha\right]$. Look at $u \to \Phi(u) = \frac{1}{u}\ln(1 + u^\alpha)$ for $u \in (0, \infty)$. We want to show that Φ is bounded: there exists c such that $|\Phi(u)| \leq c$ for all $u \in (0, \infty)$. Φ is continuous, so $\Phi(u) \xrightarrow{u\to 0} 0$ if $\alpha > 1$ and 1 if $\alpha = 1$. When $u \to 0$, $\ln(1 + u^\alpha) = u^\alpha + o(u^\alpha)$ so $\frac{1}{u}\ln(1 + u^\alpha) = u^{\alpha-1} + o(u^{\alpha-1})$. So, $\Phi(u) \xrightarrow{u\to 0} 0$, and* *as $u \to \infty$, $1 + u^\alpha$ is approximately u^α, $\ln(1 + u^\alpha)$ is approximately $\ln(u^\alpha) = \alpha \ln(u)$, so $\frac{1}{u} \ln(1 + u^\alpha)$ is approximately $\alpha \ln(u)/u$ which goes to 0. By definition of a limit, taking $\varepsilon = 1$, there exists a $\delta > 0$ such that for all $u \in (0, \delta)$ the functions $|\Phi(u) - 0| \leq 1$ and $|\Phi(u) - 1| \leq 1$, so $|\Phi(u)| < 2$. On the other hand, using the limit at ∞, there also exists an $\delta' > 0$ such that for all $u > \delta'$ we also have $|\Phi(u)| \leq 1 \leq 2$. Since Φ is continuous, it is bounded on $[\delta, \delta']$, so there exists a c such that $|\Phi(u)| \leq c$ for all $u \in [\delta, \delta']$. We thus conclude Φ is bounded.*

- *Next, we deduce that g_n is dominated by an integrable function independent of n. Take $g_n(x) = \Phi(f(x)/n)f(x)$. Note that $0 \leq g_n(x) \leq Cf(x)$. For fixed $x \in \mathbb{R}$, $g_n(x)$ is approximately $n\left(\frac{f(x)}{n}\right)^\alpha = \frac{f(x)^\alpha}{n^{\alpha-1}}$. If $\alpha > 1$, $g_n(x) \xrightarrow{n\to\infty} 0$ a.e. in $\mathbb{R}$. If $\alpha = 1$, $g_n(x) \xrightarrow{n\to\infty} f(x)$ a.e. in $\mathbb{R}$. So, the wrong reasoning above gives us the right answer here.*
- *Finally, consider $0 < \alpha < 1$. We want to show that $g_n(x) \xrightarrow{n\to\infty} \infty$. For a fixed x, $g_n(x)$ is approximately $n^{1-\alpha} f(x)^\alpha$ which diverges as $n \to \infty$. Deduce that $I_n^\alpha \xrightarrow{n\to\infty} \infty$ using Fatou's Lemma, which says $\int_A \liminf_{n\to\infty} g_n(x)\,\mathrm{d}x \leq \liminf_{n\to\infty} \int_A g_n(x)\,\mathrm{d}x$. But $\liminf_{n\to\infty} g_n = \lim_{n\to\infty} g_n = \infty$. Since $f_n \geq 0$, $\liminf_{n\to\infty} I_n^\alpha = \infty$, so $I_n^\alpha \xrightarrow{n\to\infty} \infty$.*

To see why convergence of a function does not imply convergence of its integral, consider a symmetric bounded function $f(x)$ centered at 0 whose integral is 1. Then taking $f_n(x) = \frac{1}{n} f(\frac{x}{n})$, we can see that the integral of each f_n is 1, but $f_n \xrightarrow{n\to\infty} 0$ pointwise.

Chapter 4

Banach Spaces

The study of Banach spaces focuses on the following topics.

- Spaces of differentiable functions.
- Spaces of integrable functions.
- Approximating integrable functions with differentiable functions.
- Duality. Consider a functional $f \xrightarrow{T_\phi} \int_X \phi(x) f(x)\,\mathrm{d}x$ where f is an integrable function. We identify T_ϕ with $\phi(x)$.

Definition 97. *Let E be a vector space over $\mathbb{R}$. A norm is a function $\|\cdot\| : E \to [0,\infty)$ satisfying the following.* Normed vector space

1. $\|x+y\| \le \|x\| + \|y\|$ *for all* $x, y \in E$.
2. $\|x\| = 0 \iff x = 0$.
3. $\|\lambda x\| = |\lambda|\|x\|$ *for all* $x \in E$ *and all* $\lambda \in \mathbb{R}$.

The pair $(E, \|\cdot\|)$ is called a normed vector space.

If E is over $\mathbb{C}$, then the same definition holds except where $|\lambda|$ is replaced with the complex modulus.

Proposition 98. *A normed vector space is endowed with a natural distance $d(x, y) = \|x - y\|$, so every normed vector space is a metric space.*

Definition 99. *Let E be a vector space. Suppose we have two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. They are* EQUIVALENT *iff there exists $c, c' > 0$ such that $c\|x\|_1 \leq \|x\|_2 \leq c'\|x\|_1$ for all $x \in E$. The metrics induced by two equivalent norms induce the same topology.*

Proposition 100. *If x_n is a sequence in E, and is convergent in a norm, then it is convergent in all equivalent norms, and the limits are the same.*

Definition 101. *A vector space E is finite-dimensional iff every element can be written as a finite combination of linearly independent basis vectors.*

Theorem 102. *Let E be a finite-dimensional space. Then all norms on E are equivalent.*

Example 103. *Consider $\mathbb{R}^d$, and let $x_k \in \mathbb{R}$, $k = 1, .., d$. Let*

$$\|x\|_p = \begin{cases} \left(\sum_{k=1}^d |x_k|^p\right)^{\frac{1}{p}} & 1 \leq p < \infty \\ \max_k\{|x_k|\} & p = \infty. \end{cases} \tag{4.1}$$

All of these norms are equivalent, and in fact every finite-dimensional normed vector space over $\mathbb{R}$ is isometric to $\mathbb{R}^d$ for some M. This also holds for $\mathbb{C}$ if $|\cdot|$ is replaced by the complex modulus.

Proposition 104. *For $\mathbb{R}^d$ and $\mathbb{C}^d$, all of the metrics induced by the above norms are complete.*

Proposition 105. *Let $K \subset E$ be a finite-dimensional normed vector space. Then we have K compact if and only if K closed and bounded. Note that the forward implication holds in any metric space, but the reverse implication can fail.*

Theorem 106. *Let E be a normed vector space. Then E is finite dimensional if and only if the closed unit ball* Riesz

$$\overline{B}(0,1) = \{x \in E : \|x\| \leq 1\} \tag{4.2}$$

is compact.

If $(x_n)_{n\geq 1}$ is a sequence in $\overline{B}(0,c)$ compact, then by the Bolzano-Weierstrass Theorem there exists a subsequence $(x_{n_k})_{k\neq 1}$ and $x \in E$ such that $x_{n_k} \xrightarrow{k\to\infty} x$.

For infinite-dimensional spaces, this result makes talking about convergence much more difficult than for finite-dimensional spaces. Typical approaches are either adding additional conditions on top of compactness to recover the Bolzano-Weierstrass theorem, or introducing weaker topologies in which it holds.

Definition 107. *A normed vector space $(E, \|\cdot\|)$ is a* Banach space *iff it is complete with respect to the metric induced by $\|\cdot\|$.* Banach space

4.1. Spaces of continuous and differentiable functions

Definition 108. *Let $(f_n)_{n\geq 1}$ be a sequence with $f_n : M \to \mathbb{R}$ for an open set $M \subseteq \mathbb{R}^d$ (or $\mathbb{C}^d$). Define the following.* Pointwise convergence, uniform convergence

1. *$f_n \xrightarrow{n\to\infty} f$ pointwise $\iff$ for all $x \in M$, $f_n(x) \xrightarrow{n\to\infty} f(x)$.*
2. *$f_n \xrightarrow{n\to\infty} f$ uniformly $\iff$ for all $\varepsilon > 0$ there exists n_0 such that for all $n \geq n_0$ we have $|f_n(x) - f(x)| \leq \varepsilon$ for all $x \in M$.*

This definition extends to functions defined on $\overline{M}$ where M is open.

Proposition 109. *$f_n \xrightarrow{n\to\infty} f$ uniformly $\implies$ $f_n \xrightarrow{n\to\infty} f$ pointwise – but not the other way around.*

Theorem 110. *Suppose $(f_n)_{n\geq 1}$ is a sequence of functions $f_n : M \to \mathbb{R}$. Assume the following.*

1. *f_n is continuous on M for all $n \geq 1$.*
2. *There exists $f : M \to \mathbb{R}$ such that $f_n \xrightarrow{n\to\infty} f$ uniformly.*

Then f is continuous.

Proof. Take any $x \in M$. We show that f is continuous at x. Consider $|f(x+h) - f(x)|$, $h \in \mathbb{R}^d$, such that $x + h \in M$. There exists $r > 0$ such that $B(x, r) \subseteq M$. So we can take $|h| < r$. We have

$$|f(x+h)-f(x)| \leq |f(x+h)-f_n(x+h)|+|f_n(x+h)-f_n(x)|+|f_n(x)-f(x)| \tag{4.3}$$

By uniform convergence of f_n to f, for every $\varepsilon > 0$ there exists an n_0 such that for all $n \geq n_0$, and all y, $|f_n(y) - f(y)| \leq \varepsilon$. Take n_0 such that the right-hand-side of the inequality is $\varepsilon/3$. Taking $n > n_0$, we get

$$|f(x+h) - f(x)| \leq \frac{2}{3}\varepsilon + |f_n(x+h) - f_n(x)|. \tag{4.4}$$

Since f_n is continuous, there exists a $\delta > 0$ such that for all h with $|h| < \delta$ we have $|f_n(x + h - f_n(x)| \leq \varepsilon/3$. We conclude that for all x and for all ε there exists a δ such that for all

$$|f(x+h) - f(x)| \leq \varepsilon \tag{4.5}$$

which gives continuity of f. ■

The above theorem extends to $\overline{M}$ if f is continuous on $\overline{M}$ iff. Note that f is continuous on $\overline{M}$ iff there exists an open set $\overline{M} \subset \widetilde{M}$ and there exists $\tilde{f} : \widetilde{M} \to \mathbb{R}$ and $f(x) = \tilde{f}(x)$ for all $x \in \overline{M}$.

Suppose M is an open set. Consider $f_n : \overline{M} \to \mathbb{R}$ (or $\mathbb{C}$). Let

$$B(\overline{M}) = \{f : \overline{M} \to \mathbb{R} : f \text{ bounded}\} \tag{4.6}$$

be the set of everywhere-bounded functions, i.e. $f \in B(\overline{M})$ iff there exists $c > 0$ such that $|f(x)| \leq c$ for all $x \in \overline{M}$. Define

$$\|f\|_\infty = \sup_{x\in\overline{M}} |f(x)| \tag{4.7}$$

which makes $B(\overline{M})$ a normed vector space. If we consider a sequence $(f_n)_{n\geq 1}$, with $f_n \in B(\overline{M})$, then $f_n \xrightarrow{n\to\infty} f$ uniformly means for all $\varepsilon > 0$ there exists an n_0 such that for all $n \geq n_0$, $|f_n(x) - f(x)| < \varepsilon$ for all $x \in \overline{M}$. Hence

$$|f(x)| \leq |f(x) - f_n(x)| + |f_n(x)| \leq \varepsilon + \|f_n\|_\infty \tag{4.8}$$

so $f \in B(\overline{M})$. Hence, our assumptions hold if and only if for all $\varepsilon > 0$, there exists $n_0 > 0$ such that for all $n \geq n_0$, $\sup_{x\in\overline{M}} |f_n(x) - f(x)| \leq \varepsilon$ and hence $\|f_n - f\|_\infty \xrightarrow{n\to\infty} 0$.

Proposition 111. *The space of bounded functions* $(B(\overline{M}), \|\cdot\|_\infty)$ *is a Banach space.*

Proof. Start with a sequence $(f_n)_{n\geq 1}$ with $f_n \in B(\overline{M})$, and suppose f_n is Cauchy for $\|\cdot\|_\infty$. Hence, for all $\varepsilon > 0$ there exists an n_0 such that for all n and all $m \geq n_0$, we have for all $x \in \overline{M}$ that $\|f_n(x) - f_m(x)\| \leq \varepsilon$. For all $x \in \overline{M}$, f_n is Cauchy in $\mathbb{R}$. Hence, there exists an $f(x)$ such that $f_n(x) \xrightarrow{n\to\infty} f(x)$ in $\mathbb{R}$. Since Cauchy sequences are bounded, $(f_n)_{n\geq 1}$ is bounded in $B(\overline{M})$, so there exists a $c > 0$ such that $\|f_n\|_\infty \leq c$ for all $n \geq 1$, which by previous statement holds for all $x \in \overline{M}$. Since taking $x \in \overline{M}$, $n \to \infty$, we get $|f(x)| \leq c$ for all $x \in \overline{M}$, we conclude $f \in B(\overline{M})$. Using that f_n is Cauchy in $(B(\overline{M}), \|\cdot\|_\infty)$, we conclude for all x and all $\varepsilon > 0$ there exists n_0 such that for all $n, m \geq n_0$, $|f_n(x) - f_m(x)| \leq \varepsilon$. Letting $m \to \infty$ we conclude for all x that $|f_n(x) - f(x)| \leq \varepsilon$ which means that $f_n \xrightarrow{n\to\infty} f$ uniformly. Since every Cauchy sequence converges, we conclude $B(\overline{M})$ is a Banach space. ∎

Remark 112. *From here on, assume M is a bounded open set.*

Definition 113. *Define* Continuous functions

$$C^0(\overline{M}) = \{f : \overline{M} \to \mathbb{R} \text{ s.t. } f \text{ continuous}\}. \tag{4.9}$$

We have $C^0(\overline{M}) \subseteq B(\overline{M})$. We know $(f_n)_{n\geq 1}$ and $f_n \xrightarrow{n\to\infty} f$ uniformly implies $\|f_n - f\| \xrightarrow{n\to\infty} 0$. Then $f \in C^0(\overline{M})$, which means $C^0(\overline{M})$ is a closed subspace of $B(\overline{M})$. So, $C^0(\overline{M})$ is complete, hence $(C_0(\overline{M}), \|\cdot\|_\infty)$ is a Banach space.

Differentiable functions

Definition 114. *Let $\widetilde{M}$ be an open set containing $\overline{M}$. Define*

$$C^0(\widetilde{M}) = \{f : \overline{M} \to \mathbb{R} \text{ s.t. } f \text{ } k \text{ times continuously differentiable}\} \tag{4.10}$$

and define

$$C^0(\overline{M}) = \{\tilde{f}|_{\overline{M}} : \tilde{f} \in C^0(\widetilde{M})\} \tag{4.11}$$

noting that this definition is independent of the chosen open set $\widetilde{M}$.

Notation for derivative

Remark 115. *The notation $\partial_{x_i} f$ denotes the partial derivative of f with respect to x_i.*

Multi-index

Definition 116. *Let $\alpha = (\alpha_1, \alpha_2, .., \alpha_d) \in \mathbb{N}^d$. Call α a multi-index, let $|\alpha| = \alpha_1 + .. + \alpha_d$ be its length, and define*

$$\partial^\alpha f(x) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} .. \partial x_d^{\alpha_d}}. \tag{4.12}$$

The previous definition for a space of k times differentiable functions is equivalent to saying that

$$C^k(\overline{M}) = \left\{ f \in C^k(M) \text{ s.t. } \begin{array}{l} \text{all partial derivatives of } f \text{ up to order } k \\ \text{extend continuously to a function on } C^0(\overline{M}) \end{array} \right\} \tag{4.13}$$

where the condition is taken to mean for any multi-index $\alpha \in \mathbb{N}^d$, with $|\alpha| \leq k$, the function $\partial^\alpha f(x)$ extends to a continuous function on $\overline{M}$.

C^k norm

Definition 117. *Define*

$$\|f\|_{k,\infty} = \max_{\substack{\alpha \in \mathbb{N}^d \\ |\alpha| \leq k}} \|\partial^\alpha f\|_\infty \tag{4.14}$$

which is a norm on $C^k(\overline{M})$.

Taylor's formula

Theorem 118. *Let M be an open subset of $\mathbb{R}^d$. Let $f : M \to \mathbb{R}$ such that $f \in C^1(M)$. Let $x, y \in M$ such that $[x, y] \subseteq M$ where this set is defined as $[x, y] = \{tx + (1-t)y, t \in [0,1]\}$. Then*

$$f(y) = f(x) + \int_0^1 \sum_{j=1}^{d} \partial_{x_j} f(x + t(y-x))(y_j - x_j)\,\mathrm{d}t. \tag{4.15}$$

Proof. We know from the fundamental theorem of calculus that

$$\phi(1) - \phi(0) = \int_0^1 \phi'(t)\,\mathrm{d}t \tag{4.16}$$

so define

$$\phi(t) = f((1-t)x + ty). \tag{4.17}$$

We have $\phi(0) = f(x)$, $\phi(1) = f(y)$, and

$$\phi'(t) = \sum_{j=1}^{d} \partial_{x_j}(x + t(y-x)) \frac{\partial(1-t)x_j + ty_j}{\partial t} \tag{4.18}$$

which gives the formula. ■

Theorem 119. *Let M be an open subset of $\mathbb{R}^d$ which is star-shaped about $f_0 \in M$. This means that for all $x \in M$, $[x_0, x] \subseteq M$ where this set is defined as $[x_0, x] = \{tx_0 + (1-t)x, t \in [0,1]\}$. Let $(f_n)_{n\geq 1}$, $f_n : M \to \mathbb{R}$, $f_n \in C^1(M)$ for all $n \geq 1$. Assume the following.*

1. *The sequence $(f_n(x_0))_{n\geq 1}$ converges in $\mathbb{R}$.*
2. *The sequence $(\partial_{x_i} f_n)_{n\geq 1}$ converges uniformly to a function $g_i : M \to \mathbb{R}$.*

Then f_n converges uniformly to a function $f : M \to \mathbb{R}$ such that $f \in C^1(M)$ and $g_i = \partial_{x_i} f$.

Proof. We only prove the case where M is bounded. Use Taylor's Formula for f_n. For all x, and $[x_0, x] \subseteq M$, we have

$$f_n(x) = f_n(x_0) + \int_0^1 \left[\sum_{j=1}^{d} \frac{\partial f_n}{\partial x_j}(x + t(x - x_{0j}))(x_i - x_{0j}) \right] \mathrm{d}t. \tag{4.19}$$

We know there exists $a \in \mathbb{R}$ such that $f_n(x_0) \xrightarrow{n\to\infty} a$, so for all $\varepsilon > 0$ there exists $n_0 > 0$ such that for all $n \geq n_0$, $|f_n(x_0) - a| \leq \varepsilon/2$. Since $\frac{\partial f_n}{\partial x_j} \xrightarrow{n\to\infty} g_j$ uniformly, with $\frac{\partial f_n}{\partial x_j} \in C^0(M)$, we can define

$$f(x) = a + \int_0^1 \left[\sum_{j=1}^d g_j(x + t(x - x_{0j}))(x_i - x_{0j})\right] \mathrm{d}t. \tag{4.20}$$

Uniform convergence means that for all $\varepsilon > 0$ there exists $n_1 > 0$ such that for all $n \geq n_1$, $\left|\frac{\partial f_n}{\partial x_j} - g_j(x)\right| \leq \varepsilon$. Define the diameter of M to be $\operatorname{diam}(M) = \max\{|x - y| : x, y \in M\}$. Taking the difference between the above quantities,

$$|f_n(x) - f(x)| \leq |f_n(x_0) - a| + \int_0^1 \sum_{j=1}^d \left|\left[\frac{\partial f_n}{\partial x_j} - g_j\right](x + t(x - x_{0j}))\right| \left|x_j - x_{0j}\right| \mathrm{d}t. \tag{4.21}$$

Note that $\left|x_j - x_{0j}\right| \leq L$ because the diameter of M is finite. For any x and any $\varepsilon > 0$ there exists $n \geq \max\{n_0, n_1\}$ such that

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2} + \int_0^1 \sum_{j=1}^d \frac{\varepsilon}{2dL} L \,\mathrm{d}t \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leq \varepsilon \tag{4.22}$$

hence $f_n \xrightarrow{n\to\infty}$ uniformly. We still need to prove that g_j is the partial derivative of f. Take $x \in M$, take $B(x, r) \subseteq M$, which is star-shaped about x. Using the previous step, write

$$f(y) = f(x) + \int_0^1 \sum_{j=1}^d g_j(x + t(y - x))(y_j - x_j) \,\mathrm{d}t \tag{4.23}$$

and repeat the previous proof. By definition

$$\frac{\partial f}{\partial x_1} = \lim_{h\to 0} \frac{f(x + h(1, 0, .., 0)) - f(x)}{h} \tag{4.24}$$

and

$$f(x + h(1, 0, .., 0)) - f(x) = \int_0^1 g_1(x + th(1, 0, .., 0))h \,\mathrm{d}t \tag{4.25}$$

and by dividing both sides by h, taking the limit to zero, and noting $\int_0^1 \mathrm{d}t = 1$, so we get $g_1(x) = \frac{\partial f}{\partial x_j}(x)$ which completes the argument. ■

The result about convergence of sequences of $C^1(M)$ functions, with star-shaped about x_0 domain M, stated previously, can be extended with M replaced by $\overline{M}$.

Proposition 120. *Let M be a bounded open set of $\mathbb{R}^d$. Then $C^k(\overline{M})$, equipped with the norm*

$$\|\cdot\|_{k,\infty} = \max_{\substack{\alpha\in\mathbb{N}^d \\ |\alpha|\leq k}}\left\{\|\partial^\alpha\cdot\|_\infty\right\} \tag{4.26}$$

is a Banach space.

Proof. For $k=0$, we already have the argument. For $k>1$, the argument is by induction. Here, we handle the case $k=1$. Take $(f_n)_{n\geq 1}$, $f_n : \overline{M} \to \mathbb{R}$, $f_n \in C^1(\overline{M})$, Cauchy for $\|\cdot\|_{1,\infty}$. This means that for every $\varepsilon >$, there exists an n_0 such that for all $n,m \geq n_0$, $\|f_n - f_m\|_{1,\infty} \leq \varepsilon$. This means that

$$\|f_n - f_m\|_\infty + \sum_{j=1}^{d}\left\|\partial_{x_j}(f_n - f_m)\right\|_\infty \leq \varepsilon. \tag{4.27}$$

The sequence $(\partial_{x_j} f_n)_{n\geq 1}$ is a sequence in $C^0(\overline{M})$. Clearly, we have

$$\|\partial_{x_k}(f_n - f_m)\|_\infty \leq \|f_n - f_m\|_\infty + \sum_{j=1}^{d}\left\|\partial_{x_j}(f_n - f_m)\right\|_\infty \leq \varepsilon \tag{4.28}$$

because the middle expression simply adds positive terms to the first expression. Hence, $(\partial_{x_k} f_n)_{n\geq 1}$ is Cauchy, and is convergent for all $k = 1,..,d$. Equivalently, $\partial_{x_k} f_n$ is uniformly convergent. Call its limit g_k. We also have that $(f_n)_{n\geq 1}$ is convergent in $C^0(\overline{M})$, so it is also uniformly convergent. Call its limit f. Both of these converge at any point, therefore by previous result since every ball is star-shaped, they converge over every ball, and moreover $g_k = \partial_{x_k} f$ over every ball. So, we conclude for all $k=1,..,d$, $\|\partial_{x_k} f_n - \partial_{x_k} f\|_\infty \xrightarrow{n\to\infty} 0$. But if we combine this with $\|f_n - f\|_\infty \xrightarrow{n\to\infty} 0$, we get that the sum converges, and $\|f_n - f\|_{1,\infty} \xrightarrow{n\to\infty} 0$. Hence, $C^k(\overline{M})$ is a Banach space. ■

4.2. Lebesgue spaces

Definition 121. *$L^1(M)$ it the space of of equivalence classes of Lebesgue-integrable functions under the equivalence relation $f \sim g$ iff $f = g$ a.e.*

Proposition 122. *The mapping $\|\cdot\|_{L^1(M)} : L^1(M) \to [0, \infty)$, $f \mapsto \int_M |f(x)| \,\mathrm{d}x$ is a norm. The pair $(L^1(M), \|\cdot\|_{L^1(M)})$ is a normed vector space.*

Definition 123. *Define $L^p(M)$, $1 \le p \le \infty$ by*

$$L^p(M) = \{f : M \to \mathbb{R} : f \text{ Lebesgue-measurable}, |f|^p \in L^1(M)\}. \tag{4.29}$$

and define $\|\cdot\|_{L^p(M)}$ by

$$\|f\|_{L^p(M)} = \left[\int_M |f(x)|^p \,\mathrm{d}x\right]^{1/p}. \tag{4.30}$$

Definition 124. *Define $L^\infty(M)$ by*

$$L^\infty(M) = \{f : M \to \mathbb{R} : f \text{ Lebesgue-measurable}, \exists c > 0 \text{ s.t. } |f| \le c \text{ a.e.}\}. \tag{4.31}$$

and

$$\|f\|_{L^p(M)} = \inf\{c > 0 \text{ s.t. } |f(x)| \le c \text{ a.e.} x \in M\}. \tag{4.32}$$

The above definitions are slight abuse of notation, because we refer to classes a.e. equivalent Lebesgue-measurable functions simple as functions.

Example 125. *The function $1/p$ is in L^p for all $p > 1$ but is not in L^1.*

Hölder's Inequality

Proposition 126. *Let $pin[1, \infty]$. Let the conjugate exponent q of p be defined by $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^P(M)$ and $g \in L^q(M)$, then $fg \in L^1(M)$, and*

$$\|fg\|_{L^1(M)} \le \|f\|_{L^p(M)} \|g\|_{L^q(M)}. \tag{4.33}$$

Proof. Exercise – follows from a convexity inequality. ■

Proposition 127. *For every $p \in [1, \infty]$ we have* Minkowski's ineqaulity

$$\|f+g\|_{L^p(M)} \le \|f\|_{L^p(M)} + \|g\|_{L^p(M)}. \tag{4.34}$$

Proof. Exercise – again follows from convexity inequalities. ■

Corollary 128. *The pair $(L^p(M), \|\cdot\|_{L^p(M)})$ is a normed vector space.*

Proof. Using Minkowski's inequality, it follows that for every p, $\|\cdot\|_{L^p(M)}$ defines a norm. ■

Theorem 129. *For all $p \in [1, \infty]$, the pair $(L^p(M), \|\cdot\|_{L^p(M)})$ is complete, and hence a Banach space.* Fischer-Riesz

Definition 130. *Let $f : M \to \mathbb{R}$ be continuous. Note that the set $\{x \in M : f(x) = 0\}$ is closed, because $\{0\}$ is closed, and the preimage of a closed set under a continuous map is closed. Hence, its complement is open. Define* Support

$$\operatorname{supp}(f) = \overline{\{x \in M : f(x) \neq 0\}} \tag{4.35}$$

which is the closure of an open set and hence closed.

Definition 131. *Let $f \in L^p_{\text{loc}}(M)$ for an open set M of $\mathbb{R}^d$ iff for all compact sets $K \subseteq M$, then $f\mathbb{1}_K \in L^P(M)$.*

Example 132. *The function $1/\sqrt{1+x^2}$ is not in L^1 because it behaves like $1/x$ at ∞. On the other hand, since $0 \le f(x) \le 1$, it is in L^1_{loc}, because*

$$\int_a^b |f(x)| \,\mathrm{d}x = \left\| f\mathbb{1}_{[a,b]} \right\|_{L^1(M)} \le \int_a^b 1 \,\mathrm{d}x. \tag{4.36}$$

Example 133. *Take $D = (0, 1)$ and $f(x) = 1/x$, which is not integrable* *at 0. Thus $f \notin L^1((0,1))$. If we let $[a,b] \subseteq (0,1)$, then*

$$\left\| f\mathbb{1}_{[a,b]} \right\|_{L^1((0,1))} = \int_a^b \frac{1}{x}\,\mathrm{d}x = \ln(b) - \ln(a) < \infty \tag{4.37}$$

so $f \in L^1_{\mathrm{loc}}((0,1))$.

Definition 134. *Let M be an open set of $\mathbb{R}^d$. Then $f \in C^0_c(M)$ iff f is continuous on M, i.e. $f \in C^0(M)$, and $\operatorname{supp}(f)$ is compact in M.*

Proposition 135. *We have that $f \in L^p_{\mathrm{loc}}(M)$ implies $f \in L^1_{\mathrm{loc}}$ for all $p \in [1,\infty]$.*

Proof. We have that $f \in L^p_{\mathrm{loc}}$ holds iff for all K compact, $f\mathbb{1}_K \in L^p(M)$. We know that $\mathbb{1}_K^p = \mathbb{1}_K$, so

$$\int_M (f\mathbb{1}_K)\,\mathrm{d}x = \int_M |f(x)|^p \mathbb{1}_K\,\mathrm{d}x = \int_K |f(x)|^p\,\mathrm{d}x < \infty. \tag{4.38}$$

Then using Hölder's inequality

$$\begin{aligned}
\int_M |(f\mathbb{1}_K)(x)|\,\mathrm{d}x &= \int_K |f(x)|\,\mathrm{d}x && (4.39)\\
&= \int_K |f(x)|\mathbb{1}(x)\,\mathrm{d}x && (4.40)\\
&\leq \left[\int_K |f(x)|^p\,\mathrm{d}x\right]^{1/p}\left[\int_K \mathbb{1}(x)\,\mathrm{d}x\right]^{1/q} && (4.41)\\
&\leq \left[\int_K |f(x)|^p\,\mathrm{d}x\right]^{1/p}\mu(K)^{1/q} && (4.42)\\
&< \infty && (4.43)
\end{aligned}$$

because the Lebesgue measure of K is finite. For the case $p = \infty$, we can write

$$\int_K |f(x)|\,\mathrm{d}x \leq \|f\|_{L^\infty(K)}\mu(K) < \infty \tag{4.44}$$

which gives the result. ∎

Proposition 136. *We have that $f \in C^0(M)$ implies $f \in L^p_{\mathrm{loc}}(M)$ for all $p \in [1,\infty]$.*

Proof. We have that $f \in C^0(M)$ implies for all $K \subseteq M$ compact, f is bounded on K, so $|f(x)| \le C_K$ for all $x \in K$. Hence

$$\int_K |f(x)|^p \,\mathrm{d}x \le C_K^p \mu(K) < \infty \tag{4.45}$$

which gives the result. ■

4.3. Approximation of integrable functions by smooth functions

Definition 137. *Let $f, g : \mathbb{R}^d \to \mathbb{R}$. Define* Convolution

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,\mathrm{d}y = \int_{\mathbb{R}^d} f(z)g(x-z)\,\mathrm{d}z \tag{4.46}$$

noting that $f * g = g * f$.

Proposition 138. *The following hold.*

1. *Let $f \in C_c^0(\mathbb{R}^d$, take $g \in L^1_{\text{loc}}(\mathbb{R}^d)$. Then $f * g$ is well-defined for every $x \in \mathbb{R}^d$, and $f * g \in C^0(\mathbb{R}^d)$.*
2. *If $f \in C_c^k(\mathbb{R}^d) = C^k(\mathbb{R}^d) \cap C_c^0(\mathbb{R}^d)$ and $g \in L^1_{\text{loc}}(\mathbb{R}^d)$, then $f * g \in C^k(\mathbb{R}^d)$, and for $\alpha \in \mathbb{N}^d$ with $|\alpha| \le k$, $\partial^\alpha(f * g) = \partial^\alpha f * g$.*
3. *If $f \in C_c^\infty(\mathbb{R}^d) = \bigcap_{k\in\mathbb{N}} C_c^k(\mathbb{R}^d)$ and $g \in L^1_{\text{loc}}$, then $f * g \in C^\infty(\mathbb{R}^d)$.*

Proof. Since $\operatorname{supp}(f)$ is compact, there exists $R > 0$ such that $\operatorname{supp}(f) \subseteq \bar{B}(0, R) = \{x \in \mathbb{R}^d : |x| \le R\}$. Fix $x \in \mathbb{R}^d$. Let y such that $y \in \operatorname{supp}(f)$, so $x - y \le R$. Then $|y| - |x| \le |x - y| \le R$ implies $|y| \le R + |x|$. Since $f \in C_c^0(M)$ we have

$$\sup_{x\in M} |f(x)| = \sup_{x\in K} |f(x)| = \|f\|_{\infty,D} < \infty. \tag{4.47}$$

Then

$$\|f(x-y)g(y)\| \le |g(y)| \|f\|_{\infty,D} \mathbb{1}_{\overline{B}(0,R+|x|)}(y) \tag{4.48}$$

so since $g\mathbb{1}_{\bar{B}(0,R+|x|)} \in L^1(\mathbb{R}^d)$ the norm is finite, and $y \mapsto f(x-y)g(y)$ is in $L^1(\mathbb{R}^d)$ and thus $f * g$ is finite and well-defined. Next, we show it is continuous. For all $x_0 \in \mathbb{R}^d$, we need to show that $f * g$ is continuous at x. Let $(x_n)_{n\geq 1}$, $x_n \in \mathbb{R}^d$ be a sequence such that $x_n \xrightarrow{n\to\infty} x_0$. We need to show that $(f * g)(x_n) \xrightarrow{n\to\infty} (f * g)(x_0)$. First note that f is continuous, $f(x_n - y) \xrightarrow{n\to\infty} f(x_0 - y)$ pointwise. Hence, $f(x_n - y)g(y) \xrightarrow{n\to\infty} f(x_0 - y)g(y)$ almost everywhere. We aim to show the result by dominated convergence. Let

$$|h_n(y)| = |f(x_n - y)g(y)| \leq |g(y)|\|f\|_{\infty,K}\mathbb{1}_{\bar{B}(0,R+|x_n|)} \leq |g(y)|\|f\|_{\infty,K}\mathbb{1}_{\bar{B}(0,R+m)} \tag{4.49}$$

where M is any constant such that $|x_n| \leq M$, which exists because every convergent sequence is bounded. But the function $|g(y)|\|f\|_{\infty,K}\mathbb{1}_{\bar{B}(0,R+m)}$ is integrable and independent of n, so we can apply the dominated convergence theorem, and conclude

$$(f * g)(x_n) = \int_{\mathbb{R}^d} f(x_n - y)g(y)\,\mathrm{d}y \xrightarrow{n\to\infty} \int_{\mathbb{R}^d} f(x_0 - y)g(y)\,\mathrm{d}y = (f * g)(x_0) \tag{4.50}$$

which gives the result. The proof of differentiability is very similar to the proof of continuity, and follows by writing the limit definition of the derivative. ■

Here, and in all of these notes, the notation $(x_n)_{n\geq 1}$ denotes a countably-indexed sequence.

Mollifier

Definition 139. *A sequence of mollifiers $(\rho_n)_{n\geq 1}$ is any sequence with the following properties.*

1. $\rho_n \in C_c^\infty(\mathbb{R}^d)$.
2. $\operatorname{supp}(\rho_n) \subseteq \bar{B}(0, 1/n)$.
3. $f_n \geq 0$ *on* $\mathbb{R}^d$.
4. $\int_{\mathbb{R}^d} \rho_n(x)\,\mathrm{d}x = 1$.

Example 140. *Take*

$$\rho(x) = \begin{cases} c\exp\left\{\frac{1}{|x|^2-1}\right\} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases} \quad c = \left[\int_{|x|<1} \exp\left\{\frac{1}{|x|^2-1}\right\}\mathrm{d}x\right]^{-1}. \tag{4.51}$$

On $\bar{B}(0,r)$, $r<1$, $f(x)$ is C^∞. For any x_0 such that $|x_0|=1$, take $x_n<1$ such that $x_n \xrightarrow{n\to\infty} x_0$. Hence, $|x_n|^2-1 \nearrow 0$, so $(|x_n|^2-1)^{-1}\to -\infty$ and $\rho(x_n)\to 0$. This shows that ρ is continuous everywhere, we now check its partial derivatives, which are

$$\partial_{x_k}\rho(x) = \begin{cases} c\exp\left\{\frac{1}{|x|^2-1}\right\}\frac{-2x_k}{(|x|^2-1)^2} & |x|<1 \\ 0 & |x|\geq 1 \end{cases} \tag{4.52}$$

which also goes to zero, showing that the derivatives are continuous everywhere, hence the function is C^1. Repeating this shows that $\rho\in C_c^\infty(\mathbb{R}^d)$. We can take

$$\rho_n(x) = n^d\rho(nx) \tag{4.53}$$

and which we can see immediately is a mollifier.

Proposition 141. *We have $\|f*g\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)}$.* Young's convolutional inequality

Proof.

$$\begin{aligned} \|f*g\|_{L^1(\mathbb{R}^d)} &= \int_{\mathbb{R}^d}\left|\int f(x-y)g(y)\,\mathrm{d}y\right|\mathrm{d}x & (4.54)\\ &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(x-y)||g(y)|\,\mathrm{d}y\,\mathrm{d}x & (4.55)\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(x-y)|\,\mathrm{d}x|g(y)|\,\mathrm{d}y & (4.56)\\ &= \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)} & (4.57)\end{aligned}$$

using Fubini's Theorem. ■

Proposition 142. *The following hold.*

1. *For $f\in C^0(\mathbb{R}^d)$, we have $\rho_n*f\in C^\infty(\mathbb{R}^d)$ and $\rho_n*f\to f$ uniformly on all compact sets $K\subseteq\mathbb{R}^d$, or equivalently $\|\rho_n*f-f\|_{\infty,K}$.*
2. *For $f\in L^p(\mathbb{R}^d)$ with $p\in[1,\infty)$, we have $\rho_n*f\in C^\infty(\mathbb{R}^d)$ and $||\rho_n*f-f||_{L_p(\mathbb{R}^d)} \xrightarrow{n\to\infty} 0$.*
3. *For an open set $D\subseteq\mathbb{R}^d$, $C_c^\infty(M)$ is dense in $L^p(M)$ for all $p\in[1,\infty)$.*

Proof. Let $K \subseteq \mathbb{R}^d$ be a compact set. Let $R > 0$ such that K is contained in the closed ball $\bar{B}(0, R)$. We prove the first part of the proposition: $\rho_n * f \xrightarrow{n\to\infty} f$ uniformly on $\bar{B}(0, R)$. For $x \in \bar{B}(0, R)$ we have

$$\begin{aligned}
(\rho_n * f)(x) - f(x) &= \int_{\mathbb{R}^d} \rho_n(x-y) f(y)\,\mathrm{d}y - f(x) && (4.58)\\
&= \int_{\mathbb{R}^d} \rho_n(x-y) f(y)\,\mathrm{d}y - f(x)\int_{\mathbb{R}^d} \rho_n(x-y)\,\mathrm{d}y && (4.59)\\
&= \int_{\mathbb{R}^d} \rho_n(x-y) f(y)\,\mathrm{d}y - \int_{\mathbb{R}^d} f(x)\rho_n(x-y)\,\mathrm{d}y && (4.60)\\
&= \int_{\mathbb{R}^d} \rho_n(x-y)\Big[f(y) - f(x)\Big]\,\mathrm{d}y. && (4.61)
\end{aligned}$$

Taking absolute values, we get

$$|(\rho_n * f)(x) - f(x)| \le \int_{\mathbb{R}^d} \rho_n(x-y)|f(y) - f(x)|\,\mathrm{d}y. \tag{4.62}$$

We have $x - y \in \operatorname{supp}(\rho_n) = \bar{B}(0, 1/n) \subseteq \bar{B}(0, 1)$. We also have $|y| - |x| \le |x - y| \le 1$ so $|y| \le 1 + |x|$. Hence, $y \in \bar{B}(0, R+1)$. We have f is uniformly continuous on $\bar{B}(0, R+1)$, because $\bar{B}(0, R+1)$ is compact and a continuous function on a compact set is uniformly compact. Let $\varepsilon > 0$. Then there exists a δ such that for all $x \in \bar{B}(0, R)$ and all $y \in \bar{B}(0, R+1)$, $|x-y| < \delta$, we have $|f(x) - f(y)| < \varepsilon$. Take $N \ge 1/\delta$, so that for all $n \ge N$, $1/n \le 1/N \le \delta$. Then $|x - y| \le 1/n$ implies $|f(x) - f(y)| < \varepsilon$. Hence

$$\begin{aligned}
|(\rho_n * f)(x) - f(x)| &\le \int_{\{y:|y-x|<1/n\}} \rho_n(x-y)|f(y) - f(x)|\,\mathrm{d}y && (4.63)\\
&\le \varepsilon \int_{\{y:|y-x|<1/n\}} \rho_n(x-y)\,\mathrm{d}y && (4.64)\\
&\le \varepsilon \int_{\mathbb{R}^d} \rho_n(x-y)\,\mathrm{d}y && (4.65)\\
&\le \varepsilon. && (4.66)
\end{aligned}$$

We have shown that for every $\varepsilon > 0$, there exists an N such that for all $n > N$, and all $x \in \bar{B}(0, R)$, we have$|(\rho_n * f)(x) - f(x)| \le \varepsilon$. Hence, $\rho_n * f \xrightarrow{n\to\infty} f$ uniformly. We now show the second part of the proposition. We assume a measure-theoretic result: for every $\varepsilon > 0$ and all $f \in L^1(\mathbb{R}^d)$ there exists a $g \in C_c^0(\mathbb{R}^d$ such that $\|f - g\|_{L^1(\mathbb{R}^d)} \le \varepsilon$, i.e. that $C_c^0(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$. Then for $\varepsilon > 0$,

$$\|\rho_n * f - f\|_{L^1(\mathbb{R}^d)} \leq \|\rho_n * f - \rho_n * g\|_{L^1(\mathbb{R}^d)} + \|\rho_n * g - g\|_{L^1(\mathbb{R}^d)} + \|g - f\|_{L^1(\mathbb{R}^d)}. \tag{4.67}$$

We have by the result that there exists a g such that

$$\|g - f\|_{L^1(\mathbb{R}^d)} < \frac{\varepsilon}{2}. \tag{4.68}$$

Using Young's Inequality $\|f * g\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}$ we have

$$\|\rho_n * (f - g)\|_{L^1(\mathbb{R}^d)} \leq \|\rho_n\|_{L^1(\mathbb{R}^d)} \|f - g\|_{L^1(\mathbb{R}^d)}. \tag{4.69}$$

But $\|\rho_n\|_{L^1(\mathbb{R}^d)} = 1$, so we have

$$\|\rho_n * f - f\|_{L^1(\mathbb{R}^d)} \leq \frac{1}{2}\varepsilon + \|\rho_n * g - g\|_{L^1(\mathbb{R}^d)} \tag{4.70}$$

for $g \in C_c^0(\mathbb{R}^d)$. We thus have $\|\rho_n * g - g\| \xrightarrow{n\to\infty} 0$ uniformly on all compact sets, and hence pointwise. We know that $\rho_n * g \xrightarrow{n\to\infty} g$ pointwise, so it suffices to show that there exists an $h \in L^1(\mathbb{R}^d)$ such that for all $n \geq 1$

$$|\rho_n * g)(x)| \leq h(x) \text{ a.e. } x \in \mathbb{R}^d. \tag{4.71}$$

We have

$$(\rho_n * g)(x) = \int_{\mathbb{R}^d} \rho_n(y) g(x - y) \, \mathrm{d}y. \tag{4.72}$$

Consider $|\rho_n(y)g(x-y)|$. We know $\operatorname{supp}(g) \subseteq \bar{B}(0, R)$ and $\operatorname{supp}(\rho_n) \subseteq \bar{B}(0, 1/n)$. For $|\rho_n(y)g(x-y)| \neq 0$, we need $|y| \leq 1/n$ and $|x - y| \leq R$. Using

$$|x| = |x - y + y| \leq |x - y| + |y| \leq R + \frac{1}{n} \leq R + 1 \tag{4.73}$$

we have

$$|\rho_n(y)g(x-y)| \leq \|g\|_\infty \rho_n(y) \leq \|g\|_\infty \rho_n(y) \mathbb{1}_{\bar{B}(0,R+1)}(x). \tag{4.74}$$

This gives

$$|(\rho_n * g)(x)| \leq \int_{\mathbb{R}^d} |\rho_n(y)g(x-y)| \, \mathrm{d}y \leq \|g\|_\infty \mathbb{1}_{\bar{B}(0,R+1)}(x) \int_{\mathbb{R}^d} \rho_n(y) \, \mathrm{d}y \leq \|g\|_\infty \mathbb{1}_{\bar{B}(0,R+1)}(x) \tag{4.75}$$

where $\|g\|_\infty \mathbb{1}_{\bar{B}(0,R+1)}(x)$ is an integrable function independent of n - take this function to be $h(x)$. Finally, we conclude $\|\rho_n * g - g\|_{L^1(\mathbb{R}^d)} \xrightarrow{n\to\infty} 0$ and

$$\limsup \|\rho_n * f - f\|_{L^1(\mathbb{R}^d)} \leq \frac{1}{2}\varepsilon \tag{4.76}$$

so $\|\rho_n * f - f\|_{L^1(\mathbb{R}^d)} \xrightarrow{n\to\infty} 0$. ■

Remark 143. *$C_c^\infty(M)$ is* NOT *dense in $L^\infty(M)$.*

Space of bounded continuous functions

Definition 144. *Let*

$$C_b^0(\mathbb{R}^d) = \Big\{ f : \mathbb{R}^d \to \mathbb{R} \text{ (or } \mathbb{C}) : f \text{ bounded continuous on } \mathbb{R}^d \Big\}. \tag{4.77}$$

If we let $\|f\| = \sup_{x\in\mathbb{R}^d} |f(x)| < \infty$, then $(C_b^0, \|\cdot\|_\infty)$ is a Banach space.

Space of limit-zero continuous functions

Definition 145. *Let*

$$C_0^0(\mathbb{R}^d) = \Big\{ f : \mathbb{R}^d \to \mathbb{R} \text{ (or } \mathbb{C}) : f \text{ continuous on } \mathbb{R}^d \text{ and } \lim_{n\to\infty} |f(x)| = 0 \Big\}. \tag{4.78}$$

We have that $C_0^0(\mathbb{R}^d)$ is a closed strict subspace of $C_b^0(\mathbb{R}^d)$, and that $(C_0^0, \|\cdot\|_\infty)$ is a Banach space.

Remark 146. *We have $C_c^\infty(\mathbb{R}^d) \subset C_0^0(\mathbb{R}^d)$ – the former is a dense subspace, and the latter is its completion in $\|\cdot\|_\infty$. So, $C_c^\infty(\mathbb{R}^d)$ is not dense in $L^\infty(\mathbb{R}^d)$.*

Proposition 147. *Let $u \in L^1_{\text{loc}}(M)$ such that, for all $\phi \in C_c^\infty(M)$,*

$$\int_M (u\phi)(x)\,\mathrm{d}x. \tag{4.79}$$

Then $u = 0$ a.e. on M.

4.4. Continuous maps between Banach spaces

Proposition 148. *Let E, F be normed vector spaces. Let $T : E \to F$ be a linear map. The following statements are equivalent.*

1. *T is continuous on E.*
2. *T is continuous at $x = 0$.*
3. *T is a bounded linear operator, in the sense that for all x*

$$\|T\|_{\mathrm{op}} = \|T\|_{\mathcal{L}(E,F)} = \sup_{\substack{x \in E \\ \|x\|_E \leq 1}} \|Tx\|_F = \sup_{\substack{x \in E \\ x \neq 0}} \frac{\|Tx\|_F}{\|x\|_E} \tag{4.80}$$

is bounded.

Note that

$$c = \|T\|_{\mathcal{L}(E,F)} = \sup_{x \neq 0} \frac{\|Tx\|_F}{\|x\|_E} \qquad \text{satisfies} \qquad \frac{\|Tx\|_F}{\|x\|_E} \leq c \tag{4.81}$$

for all $x \in E$. This implies that there exists a $C > 0$ such that

$$\|Tx\|_F \leq C\|x\|_E. \tag{4.82}$$

The latter property implies $T \in \mathcal{L}(E, F)$ and $\|T\|_{\mathcal{L}(E,F)} \leq C$.

Definition 149. *Let E, F be normed vector spaces. Define*

$$\mathcal{L}(E, F) = \{T : E \to F : T \text{ bounded linear}\}. \tag{4.83}$$

Then $(\mathcal{L}(E, F), \|\cdot\|_{\mathcal{L}(E,F)})$ is a normed vector space.

Proposition 150. *Let F be a Banach space. Then $\mathcal{L}(E, F)$ is a Banach space.*

Example 151. *Let E be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on E. The property that there exists a $c > 0$ such that for all x,*

$$\|x\|_2 \leq c\|x\|_1 \tag{4.84}$$

is equivalent to the property

$$\mathrm{Id} : (E, \|\cdot\|_1) \to (E, \|\cdot\|_2) \ \textit{ is continuous} . \tag{4.85}$$

Note that $\|x\|_2 \le c\|x\|_1$ does not imply that there exists a $c' > 0$ such that $\|x\|_1 \le c\|x\|_2$. More generally, if $T \in \mathcal{L}(E, F)$ and T is bijective, then T^{-1} is not necessarily continuous. However, the *Open Mapping Theorem* says that if both E, F are Banach spaces, and $T \in \mathcal{L}(E, F)$ is bijective, then $T^{-1} \in \mathcal{L}(F, E)$.

Theorem 152. *If $(E, \|\cdot\|_1)$ and $(E, \|\cdot\|_2)$ are Banach spaces, and there exists a $c > 0$ such that*

$$\|x\|_2 \le c\|x\|_1 \tag{4.86}$$

then there exists a $c' > 0$ such that

$$\|x\|_1 \le c'\|x\|_2. \tag{4.87}$$

Equivalence of norms

Definition 153. *If we have*

$$\frac{1}{c'}\|x\|_1 \le \|x\|_2 \le c\|x\|_1 \tag{4.88}$$

then we say that $\|x\|_2$ and $c\|x\|_1$ are EQUIVALENT.

4.5. Linear forms over Banach spaces

Linear form

Definition 154. *Let E be a normed vector space over $\mathbb{R}$ (or $\mathbb{C}$). Let $\phi : E \to \mathbb{R}$ (or $\mathbb{C}$). Then we say the following.*

1. *If ϕ is linear, we say ϕ is a* LINEAR FORM.
2. *If ϕ is continuous (or bounded), we say that ϕ is a* CONTINUOUS LINEAR FORM.
3. *The space $E' = \mathcal{L}(E, \mathbb{R})$ (or $\mathbb{C}$) is a called the* DUAL SPACE. *It is a Banach space even if E is not a Banach space.*

4. *For $\phi \in E'$, $x \in E$, then $\phi(x) \in \mathbb{R}$ (or $\mathbb{C}$). Call $\langle \phi \mid x \rangle_{E',E}$ a* DUALITY PRODUCT *between a continuous linear form and a vector $x \in E$.*
5. *If $\phi \in \mathcal{L}(E, \mathbb{R}) = E'$ (or $\mathbb{C}$), then $\|\phi\|_{E'} = \sup_{\substack{x \in E \\ \|x\| \le 1}} |\langle \phi \mid x \rangle| = \sup_{x \neq 0} \frac{|\langle \phi \mid x \rangle|}{\|x\|} < \infty$.*

THEOREM 155. *Let E be a normed vector space, and F a linear subspace of E. Assume $\bar{F} \neq E$. Then there exists a $\phi \in E'$ such that the following hold.*

1. *$\phi \neq 0$.*
2. *$\langle \phi \mid x \rangle = 0$ for all $x \in F$.*

Proof. Follows as a corollary of the Hahn-Banach Theorem. ∎

The way we use this is to show that $\bar{F} = E$. We assume it isn't true, so that there exists a $\phi \in E'$, $\phi \neq 0$, $\phi|_F = 0$, and show a contradiction.

THEOREM 156. *Let E be a normed vector space. Then for all x,*

$$\|x\| = \sup_{\substack{\phi \in E \\ \|\phi\|_E \le 1}} |\langle \phi \mid x \rangle|. \tag{4.89}$$

COROLLARY 157. *For all $x \in E$, $x \neq 0$, there exists a $\phi \in E'$ such that*

$$\langle \phi \mid x \rangle \neq 0. \tag{4.90}$$

EXAMPLE 158. *Take $E = L^p(M)$, $1 < p < \infty$. Let $g \in L^q(M)$ where $\frac{1}{p} + \frac{1}{q} = 1$. Let*

$$T_g : L^p(M) \to \mathbb{R} \text{ (or } \mathbb{C}) \qquad T_g : f \mapsto \int_M f(x) g(x) \, \mathrm{d}x. \tag{4.91}$$

Then

$$\left| \int_M f(x) g(x) \, \mathrm{d}x \right| \le \int_M |f(x)||g(x)| \, \mathrm{d}x \le \left[\int_M |f(x)| \, \mathrm{d}x \right]^{\frac{1}{p}} \left[\int_M |g(x)| \, \mathrm{d}x \right]^{\frac{1}{q}} = \|f\|_{L^p(M)} \|g\|_{L^q(M)} \le \infty \tag{4.92}$$

and T_g is well-defined on $L^p(M)$, is linear by linearity of integrals, and takes values in $\mathbb{R}$ (or $\mathbb{C}$) so T_g is a linear form. Moreover, T_g is bounded, in the sense that there exists $c = \|g\|_{L^q(M)}$ such that $|T_g(f)| \leq c\|f\|_{L^p(M)}$, so T_g is a continuous linear form. This means $T_g \in L^p(M)'$. Moreover, clearly $\|T_g\|_{L^p(M)'} \leq \|g\|_{L^q(M)}$, and it can further be shown that

$$\|T_g\|_{L^p(M)'} = \|g\|_{L^q(M)}. \tag{4.93}$$

If we consider

$$T : L^q(M) \to L^p(M)' \qquad\qquad T : g \mapsto T_g. \tag{4.94}$$

and recall that $T_g \in L^p(M)'$ is such that for all $f \in L^p(M)$, $\langle T_g \mid f\rangle_{L^p(M)',L^p(M)} = \int_M f(x)g(x)\,\mathrm{d}x$. Then

$$\|Tg\|_{L^p(M)'} = \|g\|_{L^q(M)} \tag{4.95}$$

is an isometry. Isometries are well-behaved because they are injective – for a linear map, to check injectivity it suffices to check the kernel, so if $T_g = 0$ then $\|Tg\|_{L^p(M)'} = 0$ so $\|g\|_{L^q(M)} = 0$ and $g = 0$. So, the only element of the kernel is 0 and T is injective. Moreover, as we prove now, T is bijective.

Isometry

Definition 159. *A function $f : X \to Y$ between normed vector spaces X, Y is an* isometry *iff, for any $x \in X$, $\|x\|_X = \|f(x)\|_Y$.*

The above definition does not imply that for every y, there exists an x such that $f(x) = y$, i.e. the mapping is injective but need not be surjective and therefore not bijective.

Reisz Representation Theorem

Theorem 160. *Let $\frac{1}{p} + \frac{1}{q} = 1$. Define $T_g \in L^p(M)'$ such that for all $f \in L^p(M)$,*

$$\langle T_g \mid f\rangle_{L^p(M)',L^p(M)} = \int_M f(x)g(x)\,\mathrm{d}x \tag{4.96}$$

and

$$T : L^q(M) \to L^p(M)' \qquad\qquad T : g \mapsto T_g. \tag{4.97}$$

Then T is bijective, so there is a bijective isometry between $L^p(M)'$ and $L^q(M)$, and thus we may identify the two spaces and write $L^p(M)' \cong L^q(M)$.

This means that we can somehow identify a function $g \in L^q(M)$ with the way g operates on other functions via integration.

Theorem 161. *Define $T_g \in L^1(M)'$ such that for all $f \in L^1(M)$,*

$$\langle T_g \mid f \rangle_{L^1(M)', L^1(M)} = \int_M f(x) g(x) \,\mathrm{d}x \tag{4.98}$$

and

$$T : L^\infty(M) \to L^1(M)' \qquad\qquad T : g \mapsto T_g. \tag{4.99}$$

Then T is bijective, so there is a bijective isometry between $L^1(M)'$ and $L^\infty(M)$, and thus we may identify the two spaces and write $L^1(M)' \cong L^\infty(M)$.

Theorem 162.

Define $T_g \in L^\infty(M)'$ such that for all $f \in L^\infty(M)$,

$$\langle T_g \mid f \rangle_{L^\infty(M)', L^\infty(M)} = \int_M f(x) g(x) \,\mathrm{d}x \tag{4.100}$$

and

$$T : L^1(M) \to L^\infty(M)' \qquad\qquad T : g \mapsto T_g. \tag{4.101}$$

Then T is injective, but it is NOT SURJECTIVE*. Thus $L^1(M)$ is a strict subspace of $L^\infty(M)$, so there exist $\phi \in L^\infty(M)'$ such that there exist no $g \in L^1(M)$ such that*

$$\langle \phi \mid f \rangle_{L^\infty(M), L^\infty(M)} = \int_M f(x) g(x) \,\mathrm{d}x. \tag{4.102}$$

Remark 163. *The above shows that $L^\infty(M)$ is the dual of $L^1(M)$, but $L^1(M)$ is not the dual of $L^\infty(M)$. We can embed $L^1(M)$ in an appropriate space of measures with finite total variation, which is the dual space of $L^\infty(M)$, but these are more general than function in $L^1(M)$ and we lose some properties of $L^1(M)$.*

Chapter 5

Hilbert Spaces

Definition 164. *Let H be a vector space over $\mathbb{R}$. An inner product is a map* Inner Product Space

$$H \times H \to \mathbb{R} \qquad (u, v) \mapsto \langle u, v\rangle. \tag{5.1}$$

satisfying the following.

1. *Bilinear: for all $u, v, w \in H$ and $\alpha, \beta \in \mathbb{R}$, we have $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$ and $\langle u, \alpha v + \beta w\rangle = \alpha\langle u, v\rangle + \beta\langle v, w\rangle$.*
2. *Symmetric: for all $u, v \in H$, we have $\langle u, v\rangle = \langle v, u\rangle$.*
3. *Positive definite: $\langle u, v\rangle \geq 0$, and $\langle u, v\rangle = 0 \iff u = 0$.*

The function

$$\|u\| = \sqrt{\langle u, u\rangle} \tag{5.2}$$

defines a norm. The pair $(H, \|\cdot\|)$ is a normed vector space, and we say $(H, \langle\cdot,\cdot\rangle)$ is an inner product space.

Definition 165. *An inner product space H is a* Hilbert space *iff $(H, \|\cdot\|)$, where $\|\cdot\|$ is the norm induced by the inner product, is complete as a normed vector space.* Hilbert space

Example 166. *Take $\mathbb{R}^d$, $\langle x, y\rangle = \sum_{i=1}^d x_i y_i$. This is a Hilbert space.*

Proposition 167. *Let H be a real inner product space. Then the following hold for all $u, v \in H$.*

1. *Cauchy-Schwarz Inequality:* $|\langle u, v\rangle| \leq \|u\|\|v\|$.
2. *Parallelogram Identity:* $\left\|\frac{u+v}{2}\right\|^2 + \left\|\frac{u-v}{2}\right\|^2 = \frac{1}{2}\left[\|u\|^2 + \|v\|^2\right]$.
3. *Polarization Identity:* $\langle u, v\rangle = \frac{1}{2}\left[\|u+v\|^2 - \|u\|^2\|v\|^2\right]$.

Proof. {missed} ■

Complex Hilbert space

Definition 168. *A complex inner product is a map*

$$H \times H \to \mathbb{C} \qquad (u, v) \mapsto \langle u, v\rangle. \tag{5.3}$$

satisfying the following. Let $\bar{\cdot}$ and be the complex conjugate and $|\cdot|$ be the complex modulus.

1. *Sesquilinear: for all $u, v, w \in H$ and $\alpha, \beta \in \mathbb{C}$, we have $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$ and $\langle u, \alpha v + \beta w\rangle = \overline{\alpha}\langle u, v\rangle + \overline{\beta}\langle v, w\rangle$.*
2. *Hermitian: for all $u, v \in H$, we have $\langle u, v\rangle = \overline{\langle v, u\rangle}$. Note that $\langle u, v\rangle = \overline{\langle v, u\rangle} \implies \langle v, u\rangle \in \mathbb{R}$.*
3. *Positive definite: $\langle u, u\rangle \geq 0$, and $\langle u, u\rangle = 0 \iff u = 0$. Note that $\langle u, u\rangle \in \mathbb{R}$.*

The function

$$\|u\| = \sqrt{\langle u, u\rangle} \tag{5.4}$$

defines a norm. The pair $(H, \|\cdot\|)$ is a normed vector space, and we say $(H, \langle\cdot,\cdot\rangle)$ is a COMPLEX INNER PRODUCT SPACE. *If $\|\cdot\|$ is complete, call $(H, \|\cdot\|)$ a* COMPLEX HILBERT SPACE.

Example 169. *Take $\mathbb{C}^d$, $\langle x, y\rangle = \sum_{i=1}^d x_i \overline{y}_i$. This is a complex Hilbert space.*

Proposition 170. *Let H be a complex inner product space. Then the following hold for all $u, v \in H$.*

1. *Cauchy-Schwarz Inequality:* $|\langle u, v\rangle| \leq \|u\|\|v\|$.
2. *Parallelogram Identity:* $\left\|\frac{u+v}{2}\right\|^2 + \left\|\frac{u-v}{2}\right\|^2 = \frac{1}{2}\left[\|u\|^2 + \|v\|^2\right]$.
3. *Polarization Identity:* $\Re(\langle u, v\rangle) = \frac{1}{2}\left[\|u+v\|^2 - \|u\|^2\|v\|^2\right]$ *and* $\Im(\langle u, v\rangle) = \frac{1}{2}\left[\|u+iv\|^2 - \|u\|^2 - \|v\|^2\right]$.

Example 171. *Let M be an open set in $\mathbb{R}^d$. Then*

$$L^2(M) = \left\{f : D \to \mathbb{R} : f \text{ Lebesgue-measurable}, \int_M |f(x)|^2 \,\mathrm{d}x < \infty\right\} \tag{5.5}$$

with the inner product

$$\langle f, g\rangle = \int_M f(x)g(x)\,\mathrm{d}x \tag{5.6}$$

is a Hilbert space.

Example 172. *Let M be an open set in $\mathbb{R}^d$. Then*

$$L^2(M) = \left\{f : M \to \mathbb{C} : f \text{ Lebesgue-measurable}, \int_M |f(x)|^2 \,\mathrm{d}x < \infty\right\} \tag{5.7}$$

with the inner product

$$\langle f, g\rangle = \int_M f(x)\overline{g(x)}\,\mathrm{d}x \tag{5.8}$$

is a Hilbert space.

Remark 173. *$L^p(M)$, with $p \neq 2$, cannot be given a Hilbert space structure, in the sense that there is no inner product on $L^p(M)$ such that $\|f\|_{L^p(M)} = \sqrt{\langle f, f\rangle}$. To see this, consider $L^p((0,1))$, and check that the parallelogram identity cannot be satisfied with $f = \mathbb{1}_{(0,1/2)}$ and $g = \mathbb{1}_{(1/2,1)}$.*

5.1. The projection theorem

Recall that A is convex iff the line segment $[x, y]$, defined as the set

$$[x, y] = \{z \in H : z = (1-t)x + ty, \forall t \in [0,1]\} \tag{5.9}$$

is contained in A for all $x, y \in A$.

Intuitively, the projection theorem says that for any non-empty closed convex subset A of a Hilbert space H, there is a unique point $\xi \in A$ for which the distance between x and ξ is minimal. Furthermore, the inner product between $x - \xi$ and any other vector is negative, i.e. visually they behave like the sides of an obtuse triangle. Stated differently, in the case of a smooth boundary in finite dimension where the notion of a tangent hyperplane makes sense, the set A is located on the side away from x of the hyperplane tangent to A at ξ. Finally, taking the projection of two points can only bring them closer together, because they move closer to the set A from where they started – though it need not do so, the distance can stay the same.

Exercise 174. *Draw the above three properties.*

Projection theorem

Theorem 175. *Let H be a Hilbert space. Let $A \subseteq H$ be a non-empty closed convex subset of H. Then*

1. *For all $x \in H$, there exists a unique $\xi \in A$ such that $\|x - \xi\|_H = \min_{y \in A} \|x - y\|_H$. We call $\xi = \mathcal{P}_A x$ be the projection of x on A. Note that $\mathcal{P}_A a = a$ for all $a \in A$.*
2. *$\xi = \mathcal{P}_A x$ iff $\langle x - \xi, \eta - \xi \rangle \leq 0$ for all $\eta \in A$.*
3. *The map $\mathcal{P}_a : H \to H$ is a contraction, so $\|\mathcal{P}_A x - \mathcal{P}_A y\|_H \leq \|x - y\|_H$.*
4. *$\mathcal{P}_A$ is continuous.*

Proof. We prove each property. Take $x \in H$. Consider the set $\{\|x - y\|, y \in A\} \subseteq [0, \infty)$. This set has a lower bound, namely 0. Let $d = \inf_{y \in A} \|x - y\|$. Then M exists and $d \geq 0$. Recall that for any set $S \subset \mathbb{R}$ bounded from below, letting $s = \inf S$, there exists a sequence $(s_n)_{n=1}^{\infty}$, such that $s_n \in S$ and $s_n \xrightarrow{n \to \infty} s$. In particular, taking $n \in \mathbb{N}$, $s + 1/n > s$ so $s + 1/n$ cannot be a lower bound, so there exists an $s_n \in S$ such that $s \leq s_n < s + 1/n$, and we can let $(s_n)_{n=1}^{\infty}$ be the sequence so defined. Hence, there exists a sequence $(y_n)_{n=1}^{\infty}$ such that $y_n \in A$ and $\|x - y_n\| \xrightarrow{n \to \infty} d$, which we call the minimizing sequence. Since we do not yet know the limit of this sequence, we aim to show it is Cauchy. From the parallelogram identity, we have

$$\left\|x - \frac{y_n + y_m}{2}\right\|^2 + \left\|\frac{y_n - y_m}{2}\right\|^2 = \frac{1}{2}\Big[\|x - y_n\|^2 + \|x - y_m\|^2\Big] \tag{5.10}$$

because $\frac{y_n+y_m}{2} \in A$ since A is convex. Let $d_n = \|x - y_n\|^2$, $d_m = \|x - y_m\|^2$. For $z \in A$, $\|x - z\|^2 \geq d^2$ because M is a lower bound. Therefore

$$d^2 + \left\|\frac{y_n - y_m}{2}\right\|^2 \leq \frac{1}{2}\Big[d_n^2 + d_m^2\Big] \tag{5.11}$$

so

$$\|y_n - y_m\|^2 \leq 2\Big[d_n^2 + d_m^2 - 2d^2\Big]. \tag{5.12}$$

We know that for any $\varepsilon > 0$ there exists an n_0 such that for all $n > n_0$, $|d_n - d| < \varepsilon$, and $|d_m - d| < \varepsilon$. So, $\|y_n - y_m\|^2 \leq 4\varepsilon$ and thus $\|y_n - y_m\| \leq 2\sqrt{\varepsilon}$ and thus $(y_n)_{n=1}^{\infty}$ is Cauchy, and y_n converges to $\xi \in H$. But A is closed, so for $\xi \in A$, we have

$$\|x - \xi\| = \inf_{y\in A}\|x - y\| = \min_{y\in A}\|x - y\|. \tag{5.13}$$

Part 1 follows. We now aim to show that $\langle x - \xi, \eta - \xi\rangle \geq 0$ for all $\eta \in A$. Take $\eta \in A$, $\xi \in A$, then $[\xi, \eta] \subseteq A$. For $z \in [\xi, \eta]$, with $z \neq \xi$, we have by convexity that there exists $t \in [0, 1]$ such that $z = (1 - t)\xi + t\eta$. Since $z \neq \xi$, $t \neq 0$. So

$$\|x - \xi\| \leq \|x - z\| = \|x - ((1 - t)\xi + t\eta)\| = \|(x - \xi) - t(\eta - \xi)\| \tag{5.14}$$

hence

$$\|x - \xi\|^2 \leq \|(x - \xi) - t(\eta - \xi)\|^2 \leq \|x - \xi\|^2 - 2t\langle x - \xi, \eta - \xi\rangle + t^2\|\eta - \xi\|^2 \tag{5.15}$$

and

$$2t\langle x - \xi, \eta - \xi\rangle \leq t^2\|\eta - \xi\|^2. \tag{5.16}$$

which since $t > 0$ we can rewrite as

$$\langle x - \xi, \eta - \xi\rangle \leq \frac{t}{2}\|\eta - \xi\|^2 \tag{5.17}$$

Letting $t \to 0$ gives the forward direction. For the reverse direction, suppose $\xi \in A$ such that $\langle x-\xi, \eta-\xi\rangle \le 0$ for all $\eta \in A$. We aim to show $\|x-\xi\|^2 - \|x-\eta\|^2 \le 0$. We have

$$\begin{aligned}
\|x-\xi\|^2 - \|x-\eta\|^2 &= \|x-\xi\|^2 - \|(x-\xi)-(\eta-\xi)\|^2 && (5.18)\\
&= \|x-\xi\|^2 - \Big[\|x-\xi\|^2 + \|\eta-\xi\|^2 - 2\langle x-\xi, \eta-\xi\rangle\Big] && (5.19)\\
&= 2\langle x-\xi, \eta-\xi\rangle - \|\eta-\xi\|^2 && (5.20)\\
&\le 0 && (5.21)
\end{aligned}$$

which gives the reverse direction. Next, we aim to show the uniqueness claim. Take $\xi_1, \xi_2 \in A$. Suppose ξ_1 and ξ_2 realize $\min_{y\in A}\|x-y\|^2$. Then

$$\langle x-\xi_1, \eta-\xi_1\rangle \le 0 \qquad \langle x-\xi_2, \eta-\xi_2\rangle \le 0 \tag{5.22}$$

for all $\eta \in A$. Taking $\eta = \xi_2$ in the 1st inequality. So

$$\langle x-\xi_1, \xi_2-\xi_1\rangle \le 0 \qquad \langle x-\xi_2, \xi_1-\xi_2\rangle \le 0. \tag{5.23}$$

But we have

$$\langle x-\xi_2, \xi_1-\xi_2\rangle \le 0 \iff \langle \xi_2-x, \xi_2-\xi_1\rangle \le 0 \tag{5.24}$$

so adding the inequalities we get

$$\|\xi_2-\xi_1\|^2 = \langle \xi_2-\xi_1, \xi_2-\xi_1\rangle \le 0 \tag{5.25}$$

which implies $\xi_2 = \xi_1$. Finally, we aim to show the contraction property. Take $x_1, x_2 \in H$. If $\|x_1-\xi_1\| = \min_{y\in A}\|x_1-y\|$, and $\|x_2-\xi_2\| = \min_{y\in A}\|x_2-y\|$. To show $\|\xi_1-\xi_2\| \le \|x_1-x_2\|$. By the characterization introduced before,

$$\langle x-\xi_1, \eta-\xi_1\rangle \le 0 \qquad \langle x-\xi_2, \eta-\xi_2\rangle \le 0 \tag{5.26}$$

for all $\eta \in A$ so taking $\eta = \xi_2$ in the first inequality and moving a -1 between the elements in the inner product, we get

$$\langle x-\xi_1, \xi_2-\xi_1\rangle \le 0 \qquad \langle -x-\}\xi_2, \xi_2-\xi_1\rangle \le 0. \tag{5.27}$$

Adding these up, we get

$$\langle x_1 - \xi_1 - x_2 + \xi_2, \xi_2 - \xi_1 \rangle \leq 0 \tag{5.28}$$

therefore

$$-\langle x_2 - x_1, \xi_2 - \xi_1 \rangle + \|\xi_2 - \xi_1\|^2 \tag{5.29}$$

and rearranging terms

$$\|\xi_2 - \xi_1\|^2 \leq \langle x_2 - x_1, \xi_2 - \xi_1 \rangle \leq \|x_2 - x_1\| \|\xi_2 - \xi_1\| \tag{5.30}$$

so

$$\|\xi_2 - \xi_1\| \leq \|x_2 - x_1\| \tag{5.31}$$

which gives the result. ■

We have $\xi = \mathcal{P}_M x$ iff $\langle x - \xi, \eta - \xi \rangle \leq 0$ for all $\eta \in M$. To see this, take $\xi \in M, \eta \in M$ where M is a vector space. We have $\{\eta - \xi, \eta \in M\} = \{\zeta : \zeta \in M\}$. This is equivalent to $\langle x - \xi, \zeta \rangle \leq 0$ for all $\zeta \in M$, which is itself equivalent to $\langle x - \xi, \zeta \rangle = 0$ for all $\zeta \in M$, and the statement of the projection theorem simplifies.

Definition 176. *We say that a vector $x \in H$ is orthogonal to a vector subspace $M \subseteq H$, written $x \perp M$, iff for all $y \in M$, $\langle x, y \rangle = 0$.*

Define the set

$$M^\perp = \{a \in X : a \perp M\}. \tag{5.32}$$

Then $M^\perp$ is closed, because $a \in M^\perp$ holds iff for all $y \in M$, $\langle x, y \rangle = 0$, which holds iff for all $y \in M$, $a = T_y^{-1}(\{0\})$, which holds iff $a \in \bigcap_{y \in M} T_y^{-1}(\{0\})$. But T_y is continuous, so the preimage of a closed set is closed, and since intersections are closed sets are closed, so $M^\perp$ is closed.

Corollary 177. *Let M be a closed vector subspace of H. Consider the projection $\mathcal{P}_M$ of H on M. Recall that M is a non-empty closed convex subset of H. Then the map $\mathcal{P}_M : H \to H$ is a continuous linear map. Call $\mathcal{P}_M$ the* ORTHOGONAL PROJECTION *of H onto M. Furthermore, $\xi = \mathcal{P}_m x$ is equivalent to $x \in M$ with $x - \xi \perp M$. For all $x \in H$, we have*

$$x = \mathcal{P}_M x + \mathcal{P}_{M^\perp} x \qquad \|x\|^2 = \|\mathcal{P}_M x\|^2 + \|\mathcal{P}_{M^\perp} x\|^2. \tag{5.33}$$

We have

$$\begin{aligned} y &= x - (\mathcal{P}_M x + \mathcal{P}_{M^\perp} x) && (5.34)\\ &= (x - \mathcal{P}_M x) - \mathcal{P}_{M^\perp} x && (5.35)\\ &= (x - \mathcal{P}_{M^\perp} x) - \mathcal{P}_M x && (5.36)\end{aligned}$$

so $y \in M \cap M^\perp = \{0\}$ so $y = 0$.

Remark 178. *For an arbitrary vector subspace M of H, $(M^\perp)^\perp = \overline{M}$.*

5.2. The Riesz-Fischer Representation Theorem

Recall that the dual norm is

$$\|\phi\|_{H'} = \sup_{\substack{x \in H\\ x \neq 0}} \frac{|\langle \phi \mid x\rangle|}{\|x\|_H}. \tag{5.37}$$

Riesz Representation Theorem

Theorem 179. *For every $\phi \in H'$, there exists a vector $\xi \in H$ such that for all $x \in H$, $\langle \phi \mid x\rangle_{H',H} = \langle \xi, x\rangle$. Moreover, $\|\xi\|_H = \|\phi\|_{H'}$. Thus the map $H' \to H$, $\phi \mapsto \xi$ is a linear bijective isometry of Banach spaces.*

Proof. Suppose $\phi = 0$. Obviously $\xi = 0$ is the unique solution – to check uniqueness, suppose $\xi_1, \xi_2 \in H$ such that $\langle \phi \mid x\rangle = \langle \xi_1, x\rangle = \langle \xi_2, x\rangle$ for all $x \in H$. Subtracting the two, we get $\langle \xi_1 - \xi_2, x\rangle = 0$ for all $x \in H$. So $\|\xi_1 - \xi_2\|^2 = 0$ which implies $\xi_1 = \xi_2$. Hence, suppose $\phi \neq 0$. Then there exists $y \in H$ such that

$$\langle \phi \mid y\rangle \neq 0. \tag{5.38}$$

Let

$$M = \ker(\phi) = \{x \in H : \langle \phi \mid x\rangle = 0\} \tag{5.39}$$

which is a subspace of H. Recall that ϕ is a continuous linear form, and the functions $x \mapsto \langle \phi \mid x\rangle$ and $x \mapsto \phi(x)$ are by definition identical. The set $\{0\}$ is closed with respect to the standard topology on $\mathbb{R}$, the preimage

$\phi^{-1}(\{0\})$ is closed. Hence, M is closed. So, the orthogonal projection $\mathcal{P}_M$ is well-defined. But by definition $y \notin M$. We have

$$z = \mathcal{P}_M y \tag{5.40}$$

if and only if $z \in M$ and $y - z \perp M$. Define $u = y - z = y - \mathcal{P}_M y$. We know $u \perp m$. We are going to look for a ξ which is proportional to u. Consider

$$\langle \phi \mid u \rangle = \langle \phi \mid y \rangle - \langle \phi \mid \mathcal{P}_M y \rangle. \tag{5.41}$$

But $\langle \phi \mid \mathcal{P}_M y \rangle = 0$ because $\mathcal{P}_M y \in \ker(y)$. So

$$\langle \phi \mid u \rangle = \langle \phi \mid y \rangle. \tag{5.42}$$

For $x \in H$ consider

$$x - \frac{\langle \phi \mid x \rangle u}{\langle \phi \mid u \rangle} \tag{5.43}$$

so

$$\left\langle \phi \;\middle|\; x - \frac{\langle \phi \mid x \rangle u}{\langle \phi \mid u \rangle} \right\rangle = \langle \phi \mid x \rangle - \frac{\langle \phi \mid x \rangle \langle \phi \mid u \rangle}{\langle \phi \mid u \rangle} = 0. \tag{5.44}$$

Hence

$$\mathcal{P}_M \left[x - \frac{\langle \phi \mid x \rangle}{\langle \phi \mid u \rangle} u \right] = x - \frac{\langle \phi \mid x \rangle}{\langle \phi \mid u \rangle} u \tag{5.45}$$

but also by linearity of $\mathcal{P}_M$,

$$\mathcal{P}_M \left[x - \frac{\langle \phi \mid x \rangle}{\langle \phi \mid u \rangle} u \right] = \mathcal{P}_M x - \frac{\langle \phi \mid x \rangle}{\langle \phi \mid u \rangle} \mathcal{P}_M u = \mathcal{P}_M x \tag{5.46}$$

since $u \in M^\perp$ so $\langle u, v \rangle = 0$ for all $v \in M$, and since $w = \mathcal{P}_M u$ so $w \in M$, $u - w \in M^\perp$ hence $w \in M^\perp$ thus $w = 0$ since $M \cap M^\perp = \{0\}$. Combining the two expressions, we get

$$x = \mathcal{P}_M x + \frac{\langle \phi \mid x \rangle}{\langle \phi \mid u \rangle} u. \tag{5.47}$$

Plugging this in to $\langle x, u \rangle$ we get

$$\langle x, u \rangle = \langle \mathcal{P}_M x, u \rangle + \frac{\langle \phi \mid x \rangle}{\langle \phi \mid u \rangle} \langle u, u \rangle = \frac{\langle \phi \mid x \rangle}{\langle \phi \mid u \rangle} \langle u, u \rangle \tag{5.48}$$

since $\langle \mathcal{P}_M x, u \rangle = 0$ by definition of M. We can rewrite this as

$$\langle \phi \mid x \rangle = \left\langle x, \frac{\langle \phi \mid u \rangle}{\|u\|_H^2} u \right\rangle \tag{5.49}$$

so define

$$\xi = \frac{\langle \phi \mid u \rangle}{\|u\|_H^2}. \tag{5.50}$$

We still need to show that $\|\xi\|_H = \|\phi\|_{H'}$. We know

$$\langle \phi \mid x \rangle = \langle \xi, x \rangle. \tag{5.51}$$

Note that if we were working with the complex case, we'd need to instead write $\langle x, \xi \rangle$ to handle complex conjugates correctly. By Cauchy-Schwarz,

$$|\langle \phi \mid x \rangle| = |\langle \xi, x \rangle| \leq \|\xi\|_H \|x\|_H. \tag{5.52}$$

Recall that a linear form ϕ is continuous iff there exists a $c > 0$ such that $|\langle \phi \mid x \rangle| \leq c\|x\|$. Furthermore, $\|\phi\|_{H'} \leq c$, because $\sup_{\|x\| \leq 1} |\langle \phi \mid x \rangle|$. Going back, in our case we have $c = \|\xi\|$, so

$$\|\phi\|_{H'} \leq \|\xi\|_H. \tag{5.53}$$

On the other hand, take $x = \xi$ so

$$|\langle \phi \mid \xi \rangle| = \|\xi\|^2. \tag{5.54}$$

We know ϕ is a continuous linear form, so

$$|\langle \phi \mid \xi \rangle| \leq \|\phi\|_{H'} \|\xi\|_H \tag{5.55}$$

thus

$$\|\xi\|_H \leq \|\phi\|_{H'} \tag{5.56}$$

and combining the two expressions

$$\|\xi\|_H = \|\phi\|_{H'} \tag{5.57}$$

which gives the result. ■

5.3. Lax-Milgram Theorem

Theorem 180. *Let H be a Hilbert space over $\mathbb{R}$. Let $a : H \times H \to \mathbb{R}$ be bilinear. Suppose the following.* Lax-Milgram

1. *a is continuous on $H \times H$: for all $x, y \in H$ there exists $c > 0$ such that $|a(x,y)| \leq c\|x\|\|y\|$.*

2. *a is coercive on H: for all $x \in H$ there exists $\alpha > 0$ such that $a(x,x) \geq \alpha\|x\|^2$.*

Let $\phi \in H'$, i.e. $\phi : H \to \mathbb{R}$ linear such that for $k > 0$, $|\phi(x)| < k\|x\|$ for all x. Then there exists a unique $\xi \in H$ such that for all $y \in H$, $a(\xi, y) = \langle \phi \mid y \rangle_{H',H}$, and the map $H' \to H$, $\phi \mapsto \psi$, is linear and continuous. Furthermore, if $\alpha > 0$ is the coercivity constant of a, then $\|\xi\|_H \leq \frac{1}{\alpha}\|\phi\|_{H'}$.

Proof. For the left-hand-side, fix $x \in H$. Consider the mapping

$$H \xrightarrow{a_x} \mathbb{R} \qquad\qquad y \xmapsto{a_x} a(x,y). \tag{5.58}$$

We have

$$|a(x,y)| \leq c_1\|x\|\|y\| \leq c_2\|y\| \tag{5.59}$$

so a_x is a continuous linear form on H. By the Reisz Representation Theorem, there exists a $\xi \in H$ such that

$$a(x,y) = a_x(y) = \langle a_x \mid y \rangle_{H',H} = \langle \xi_x, y \rangle. \tag{5.60}$$

Now, consider the mapping

$$H \to H \qquad\qquad x \mapsto \xi_x \tag{5.61}$$

which is linear so

$$\begin{aligned}
\langle \xi_{x_1} + \xi_{x_2}, y \rangle &= \langle \xi_{x_1}, y \rangle + \langle \xi_{x_2}, y \rangle && (5.62)\\
&= a(x_1, y) + a(x_2, y) && (5.63)\\
&= a(x_1 + x_2, y) && (5.64)\\
&= \langle \xi_{x_1+x_2}, y \rangle && (5.65)
\end{aligned}$$

for all $y \in H$, so we conclude $\xi_{x_1} + \xi_{x_2} = \xi_{x_1+x_2}$. We thus define a linear map A by

$$\langle Ax, y \rangle = a(x,y) \tag{5.66}$$

for all $x, y \in H$. By the Reisz Representation Theorem and definition of the dual norm, we know

$$\|Ax\|_H = \|\xi_x\|_H = \|a_x\|_{H'} = \sup_{y \neq 0} \frac{|a(x,y)|}{\|y\|} \leq \sup_{y \neq 0} \frac{c_1 \|x\| \|y\|}{\|y\|} \leq \sup_{y \neq 0} c_1 \|x\| \tag{5.67}$$

so A is a continuous linear map $H \to H$. We conclude that there exists a unique continuous linear map $A : H \to H$ such that $\langle Ax, y \rangle = a(x,y)$ for all $x, y \in H$. For the right-hand-side, we have $\phi \in H'$. By the Riesz Representation Theorem, there exists a unique $\xi \in H$ such that

$$\langle \phi \mid y \rangle_{H',H} = \langle \xi, y \rangle \tag{5.68}$$

for all $y \in H$ and we have $\|\phi\|_{H'} = \|\xi\|_H$. We have thus shown that

$$a(x,y) = \langle \phi \mid y \rangle \qquad \forall y \in H \tag{5.69}$$

is equivalent to

$$\langle Ax, y \rangle = \langle \xi, y \rangle \qquad \forall y \in H \tag{5.70}$$

which is equivalent to $\langle Ax - \xi, y \rangle = \langle 0, y \rangle$ and therefore $Ax = \xi$. We thus need to show that $Ax = \xi$ has a unique solution. We show that A is injective and surjective to conclude it is bijective. We first show it is injective, by showing that

$$\ker(A) = \{x \in H : Ax = 0\} \tag{5.71}$$

is equal to $\{0\}$ and using linearity. Take $x \in \ker(A)$ so $Ax = 0$. Then

$$0 = \langle Ax, x \rangle = a(x,x) \geq \alpha \|x\|^2 \tag{5.72}$$

by coercivity implying $\|x\| = 0$ so $x = 0$, concluding that A is injective. We now aim to show that A is surjective. The image is

$$\mathrm{Im}(A) = \{y \in H : \exists x \in H \text{ s.t. } y = Ax\}. \tag{5.73}$$

We show that $\mathrm{Im}(A)$ is closed in H, and that it is dense in H, to conclude that $\mathrm{Im}(A) = H$. We first show it is closed in H. Suppose $(y_n)_{n \geq 1}$ is a sequence $y_n \in \mathrm{Im}(A)$ such that there exists $y \in H$ and $y_n \xrightarrow{n \to \infty} y$. We aim to use the characterization that a set is closed iff then it contains the limit of all such sequences. We have $y_n \in \mathrm{Im}(A)$ iff there exists $x_n \in H$ such that $y_n = Ax_n$. Our goal is to show $(x_n)_{n\geq 1}$ converges by showing it is Cauchy. Using coercivity, write

$$\begin{aligned} \alpha\|x_n - x_m\|^2 &\leq A(x_n - x_m, x_n - x_m) && (5.74)\\ &= \langle A(x_n - x_m), x_n - x_m\rangle && (5.75)\\ &= \langle y_n - y_m, x_n - x_m\rangle && (5.76)\\ &\leq \|y_n - y_m\|\|x_n - x_m\| && (5.77) \end{aligned}$$

using Cauchy-Schwarz so

$$\|x_n - x_m\| \leq \frac{1}{\alpha}\|y_n - y_m\| \tag{5.78}$$

but $(y_n)_{n\geq 1}$ is convergent so it is Cauchy so $(x_n)_{n\geq 1}$ so there exists an $x \in H$ such that $x_n \xrightarrow{n\to\infty} x$. But A is continuous, so $Ax_n = y_n \xrightarrow{n\to\infty} Ax$. Finally, we show $\mathrm{Im}(A)$ is dense. Proceeding by contradiction, suppose $\overline{\mathrm{Im}(A)} \neq H$. Since $\overline{\mathrm{Im}(A)} \subseteq H$, there exists a point $x \in H$ such that $x \notin \overline{\mathrm{Im}(A)}$. Clearly, $x \neq 0$, because $0 \in \mathrm{Im}(A)$. $\overline{\mathrm{Im}(A)}$ is closed, so we can consider orthogonal projection $\mathcal{P}_{\overline{\mathrm{Im}(A)}}$. Then the vector

$$\eta = x - \mathcal{P}_{\overline{\mathrm{Im}(A)}}x \neq 0. \tag{5.79}$$

But we also have

$$\eta \perp \overline{\mathrm{Im}(A)} \tag{5.80}$$

by the characterization of the projection. Thus

$$\langle \eta, y\rangle = 0 \qquad \forall y \in \overline{\mathrm{Im}(A)}. \tag{5.81}$$

In particular, $\langle \eta, y\rangle = 0$ for all $y \in A$. This is equivalent to

$$\langle \eta, Ax\rangle = 0 \qquad \forall x \in H. \tag{5.82}$$

Take $x = \eta$, getting

$$\langle \eta, A\eta\rangle = 0. \tag{5.83}$$

But

$$0 \leq \alpha\|\eta\|^2 \leq a(\eta, \eta) = \langle \eta, A\eta\rangle = 0 \tag{5.84}$$

so $\|\eta\| = 0$ and $\eta = 0$, giving the necessary contradiction. We conclude A is surjective, and thus that it is bijective. Therefore, for any $\xi \in H$, there exists a unique $x \in H$ such that $Ax = \xi$. Finally, we show $\|x\| \leq \frac{1}{\alpha}\|\phi\|_{H'}$. Using coercivity, write

$$\langle Ax, x\rangle = \langle \xi, x\rangle \tag{5.85}$$

and using Cauchy-Schwarz

$$\alpha\|x\|^2 = a(x,x) = \langle \xi, x\rangle \leq \|\xi\|\|x\| \tag{5.86}$$

so using the Reisz Representation Theorem

$$\|x\| \leq \frac{1}{\alpha}\|\xi\| = \frac{1}{\alpha}\|\phi\|_{H'} \tag{5.87}$$

which completes the proof. ■

Remark 181. *Bilinear and multilinear forms are continuous iff they are bounded – for the bilinear case, in the sense that for all $x, y \in H$ there exists $c > 0$ such that $|a(x,y)| \leq \|x\|\|y\|$. The multilinear case is analogous. Note that this characterization is not necessarily true for maps that are not bilinear of multilinear.*

Remark 182. *Bilinear forms are not necessarily symmetric, and so not necessarily inner products.*

Example 183. *Consider an interval $(0,1)$, and we want to find a function $u : (0,1) \to \mathbb{R}$ such that $-u'' + u = f$ where f is a given function on $(0,1)$, and $u(0) = u(1) = 0$. Then for an appropriate notion of reasonable, we let can let H be a space of reasonable functions, with $v \in H$ satisfying $v(0) = v(1) = 0$. Multiply the given equation by v, integrate, then apply integration by parts to obtain*

$$-\int_0^1 u''(x)v(x)\,\mathrm{d}x + \int_0^1 u(x)v(x)\,\mathrm{d}x = \int_0^1 f(x)v(x)\,\mathrm{d}x \quad \forall v \in H \tag{5.88}$$

$$\int_0^1 u'(x)v'(x)\,\mathrm{d}x - u'v\Big|_0^1 + \int_0^1 u(x)v(x)\,\mathrm{d}x = \int_0^1 f(x)v(x)\,\mathrm{d}x \quad \forall v \in H \tag{5.89}$$

and note that $u'v|_0^1 = 0$ *due to the boundary conditions on* u *and* v*, obtaining*

$$\underbrace{\int_0^1 u'(x)v'(x)\,\mathrm{d}x + \int_0^1 u(x)v(x)\,\mathrm{d}x}_{a(u,v)} = \underbrace{\int_0^1 f(x)v(x)\,\mathrm{d}x}_{\langle f \mid v\rangle} \quad \forall v \in H \tag{5.90}$$

which we call the VARIATIONAL FORMULATION *of the problem. The Lax-Milgram Theorem then gives us conditions for existence, uniqueness, and stability of the problem.*

PROPOSITION 184. *Suppose that* a *and* ϕ *satisfy the assumptions of the Lax-Milgram Theorem and suppose that* a *is symmetric so that* $a(x, y) = a(y, x)$ *for all* $x, y \in H$*. Then the unique solution* x_0 *to*

$$a(x, y) = \langle \phi \mid y\rangle \qquad \forall y \in H \tag{5.91}$$

is also the unique solution to the optimization problem

$$x_0 = \arg\min_{x\in H} J(x) \qquad J(x) = \frac{1}{2}a(x, x) + \langle \phi \mid x\rangle \tag{5.92}$$

so that $J(x_0) < J(x)$ *for all* $x \in H$ *such that* $x \neq x_0$*.*

Proof. If a is symmetric, bilinear, and coercive, then it is an inner product on H. So we have two inner products $\langle\cdot,\cdot\rangle_H$ and $\langle\cdot,\cdot\rangle_a$ on H. Consider $\|\cdot\|_a$. We have

$$\alpha\|x\|_H^2 \leq a(x, x) \leq c\|x\|_H^2 \tag{5.93}$$

by coercivity and continuity so

$$\sqrt{\alpha}\|x\|_H \leq \|x\|_a \leq \sqrt{c}\|x\|_H \tag{5.94}$$

and $\|\cdot\|_a$ is equivalent to a Banach norm, thus it is complete, and $(H, \langle\cdot,\cdot\rangle_a)$ is a Hilbert space. By the Reisz Representation Theorem, there exists a vector $g \in H$ such that $\langle \phi \mid x\rangle_{H',H} = \langle g, x\rangle_a$. Hence ξ, the solution of the problem

$$a(\xi, y) = \langle \phi \mid y\rangle \qquad \forall g \in H. \tag{5.95}$$

Note that this exactly gives the statement of the Lax-Milgram Theorem, which in this scenario is much easier to show than in the general because $a(\cdot,\cdot)$ is symmetric. Proceeding, we have

$$\begin{aligned} & a(\xi,y) = a(g,y) && \forall y \in H && (5.96)\\ \implies & a(\xi - g, y) = 0 && \forall y \in H && (5.97)\\ \implies & a(g-\xi, y-\xi) = 0 && \forall y \in H && (5.98)\\ \implies & a(g-\xi, y-\xi) \le 0 && \forall y \in H. && (5.99)\end{aligned}$$

This characterizes ξ as the projection of g onto H in the sense

$$\|g-\xi\|_a = \min_{x\in H}\|g-x\|_a \qquad \xi = \arg\min_{x\in H}\|g-x\|_a \quad (5.100)$$

$$\implies \quad \|g-\xi\|_a^2 = \min_{x\in H}\|g-x\|_a^2 \quad (5.101)$$

$$\implies \quad \langle g-\xi, g-\xi\rangle = \min_{x\in H}\langle g-x, g-x\rangle \quad . \quad (5.102)$$

But by symmetry we can rewrite this as

$$a(g-\xi,g-\xi) = a(g,g)+a(\xi,\xi)-a(y,x)-a(x,y) = a(g,g)+a(\xi,\xi)-2a(y,\xi) \quad (5.103)$$

and drop the term $a(g,g)$ which does not depend on ξ so

$$\begin{aligned} \xi &= \arg\min_{x\in H} a(x,x) - 2a(g,x) && (5.104)\\ &= \arg\min_{x\in H} \frac{1}{2}a(x,x) - \langle \phi \mid x\rangle && (5.105)\\ &= J(x) && (5.106)\end{aligned}$$

which gives the result.

■

Example 185. *Note that for $H=\mathbb{R}$, any bilinear form $a(x,y)=\alpha xy$ for some α. Here, a is coercive if $\alpha>0$, so for $\langle\phi \mid y\rangle = \xi y$ with $\xi\in\mathbb{R}$ we have $\alpha xy = \xi y$ for all y and $x=\frac{\xi}{\alpha}$. Here, $J(x)=\frac{1}{2}\alpha x^2 - \xi x$ so $J'(x)=\alpha x-\xi$ and $J'(x)=0$ iff $x=x_0=\frac{\xi}{\alpha}$, illustrating the above proposition.*

Remark 186. *Note that symmetry in the quadratic optimization problem characterization is essential.*

5.4. Hilbert sums and Hilbert bases

Definition 187. *Let X be a Banach space. We say X is separable iff there exists a set $Y \subseteq X$ such that the following hold.* Separable

1. *Y is countable.*
2. *$\overline{Y} = X$.*

Example 188. *$L^2(M)$ is separable. $L^\infty(M)$ is not.*

Remark 189. *Assume henceforth that H is a separable Hilbert space.*

Definition 190. *Let* Direct sum

$$\bigoplus_{n=1}^{\infty} H_n = \left\{ x \in H : \exists (x_n)_{n\geq 1} \text{ s.t. } x_n \in H_n, x_n \neq 0 \text{ finitely many times}, x = \sum_{n\geq 1} x_n \right\}. \tag{5.107}$$

Note that the sum is necessarily finite.

Theorem 191. *Let H be a separable Hilbert space. Let $(H_n)_{n\geq 1}$ be a finite or countable family of closed subspaces of H such that $H_n \perp H_{n'}$ for all $n \neq n'$, in the sense that for any $x \in H_n$, $y \in H_{n'}$, $\langle x, y\rangle = 0$ if $n \neq n'$. Assume $H = \overline{\bigoplus_{n=1}^{\infty} H_n}$. Then the following hold.*

1. *For all $x \in H$, define $x_n = \mathcal{P}_{H_n} x$. Then $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$ and $x = \sum_{n=1}^{\infty} x_n$.*
2. *Suppose $(x_n)_{n\geq 1}$ is a sequence satisfying the following conditions.*
   (a) *For all $n \geq 1$, $x_n \in H_n$.*
   (b) *$\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$.*

   Then there is a unique $x \in H$ such that $x_n = \mathcal{P}_{H_n} x$.
3. *Parseval's Formula: for $x, y \in H$ with $x = \sum_{n=1}^{\infty} x_n$, $y = \sum_{n=1}^{\infty} y_n$, we have the following*

(a) $\langle x, y\rangle = \sum_{n=1}^{\infty}\langle x_n, y_n\rangle$.

(b) $\|x\|^2 = \sum_{n=1}^{\infty}\|x_n\|^2$.

Hilbert sum

Definition 192. *Define*

$$\bigoplus_{n\geq 1}^{\perp} H_n = \overline{\bigoplus_{n=1}^{\infty} H_n} = H. \tag{5.108}$$

Call this the direct Hilbert sum of H_n.

We can write

$$x - \sum_{n=1}^{p} x_n = \sum_{n=p+1}^{\infty} x_n \tag{5.109}$$

so

$$\left\| x - \sum_{n=1}^{p} x_n \right\|^2 = \left\| \sum_{n=p+1}^{\infty} x_n \right\|^2 = \sum_{n=p+1}^{\infty} \|x_n\|^2 \xrightarrow{p\to\infty} 0 \tag{5.110}$$

where the last equality holds by orthogonality.

Span

Definition 193. *Let H be a separable Hilbert space. Let S be a countable subset of H. Define*

$$\operatorname{span} S = \left\{ x : \exists N \in \mathbb{N}, \exists (a_1, .., a_N) \in \mathbb{R}^N, \exists (s_1, .., s_N) \in S^N \text{ s.t. } x = \sum_{n=1}^{N} a_n s_n \right\} \tag{5.111}$$

Definition 194. *Let H be a separable Hilbert space. Let S be a countable subset of H. Define the following.*

1. *S is complete iff $\overline{\operatorname{span} S} = H$.*
2. *S is orthonormal iff for all $x, y \in S$, $x \neq y \implies \langle x, y\rangle = 0$, and $\langle x, x\rangle = \|x\|^2 = 1$.*
3. *S is a Hilbert basis iff S is complete and orthonormal.*

Gram-Schmidt Orthogonalization

Theorem 195. *Let H be a separable Hilbert space. The following hold.*

1. *There exists a countable complete set* $S = \{b_n\}_{n\geq 1}$ *of linearly independent vectors* b_n.
2. *There exists a Hilbert basis* $\{e_n\}_{n\geq 1}$ *such that* $\operatorname{span}\{e_1, .., s_n\} = \operatorname{span}\{b_1, .., b_n\}$.

Proof. Without loss of generality, suppose $\dim H = \infty$. By separability there exists a countable set Σ such that $\overline{\Sigma} = H$. Write $\Sigma = \{\sigma_n\}_{n\geq 1}$ with each σ_n unique. Either $\sigma_1 = 0, \sigma_2 \neq 0$ or $\sigma_1 \neq 0$: in the former case take $b_1 = \sigma_2$, in the latter take $b_1 = \sigma_1$. Suppose now inductively that $\{b_1, .., b_p\}$ are linearly independent. Each b_k corresponds to σ_{n_k}, for $k = 1, .., p$. We know there exists an index $n > n_p$ such that σ_n is linearly independent from all of $\{b_1, .., b_p\}$, otherwise $\overline{\Sigma} = H$ and $p = \dim \overline{\Sigma} = \dim H$, a contradiction. For this n, take $b_{p+1} = \sigma_n$. This defines a set $\{b_k\}_{k=1}^{\infty}$. We have

$$\Sigma \subseteq \operatorname{span}\{b_k\}_{k=1}^{\infty} \tag{5.112}$$

by construction because between each b_k and b_{k+1}, all vectors are linearly dependent on some vector in $\{b_1, .., b_k\}$, and span by definition only includes finite sums. But, clearly

$$H = \overline{\Sigma} \subseteq \operatorname{span}\{b_k\}_{k=1}^{\infty} \subseteq \Sigma \subseteq H \tag{5.113}$$

so $\operatorname{span}\{b_k\}_{k=1}^{\infty} = H$ and thus $\operatorname{span}\{b_k\}_{k=1}^{\infty}$ is complete. We now orthonormalize b_k. Take

$$e_1 = \frac{b_1}{\|b_1\|}. \tag{5.114}$$

Proceeding inductively, suppose $\{e_1, .., e_p\}$ are orthonormal such that $\operatorname{span}\{e_1, .., e_p\} = \operatorname{span}\{b_1, .., b_p\}$. Let

$$c_{p+1} = b_{p+1} - \sum_{k=1}^{p} \langle b_{p+1}, e_k\rangle e_k \tag{5.115}$$

so $c_{p+1} \in \operatorname{span}\{b_1, .., b_{p+1}\}$. But

$$\langle c_{p+1}, e_h\rangle = 0 \qquad \forall h = 1, .., p \tag{5.116}$$

because

$$\langle c_{p+1}, e_h\rangle = \langle b_{p+1}, e_h\rangle - \sum_{k=1}^{p} \langle b_{p+1}, e_k\rangle\langle e_k, e_h\rangle = \langle b_{p+1}, e_h\rangle - \sum_{k=1}^{p} \langle b_{p+1}, e_k\rangle \mathbb{1}_{k=h} \tag{5.117}$$

and

$$\langle c_{p+1}, e_h \rangle = \langle b_{p+1}, e_h \rangle - \langle b_{p+1}, e_h \rangle = 0. \tag{5.118}$$

We have that $c_{p+1} \neq 0$ because it is a linear combination of $\{b_1, .., b_{p+1}\}$ which are linearly independent and the coefficient of b_{p+1} is 1. If $c_{p+1} = 0$, this coefficient would be 0, and it is not, so $c_{p+1} = 0$. Define

$$e_{p+1} = \frac{c_{p+1}}{\|c_{p+1}\|}. \tag{5.119}$$

Then $\{e_1, .., e_{p+1}\}$ is orthonormal and $\text{span}\{e_1, .., e_{p+1}\} = \text{span}\{b_1, .., b_{p+1}\}$ so $\overline{\text{span}\{e_1, .., e_{p+1}\}} = \overline{\text{span}\{b_1, .., b_{p+1}\}}$ and therefore $\overline{\text{span}\{e_k\}_{k=1}^{\infty}} = \overline{\text{span}\{b_k\}_{k=1}^{\infty}} = H$. ■

Example 196. *Consider $L^2(0,1)$. Let $\{x_k\}_{k=1}^{\infty}$ be a complete set. There exists a Hilbert basis $\{p_k(x)\}_{k=0}^{\infty}$ where $\text{span}\{p_0(x), .., p_n(x)\} = \text{span}\{1, x, .., x^n\}$ and p_k is a polynomial for all k. Using different inner products, such as*

$$\langle f, g \rangle_{\text{Hermite}} = \int_{\mathbb{R}} f(x) g(x) e^{-x^2} \, \mathrm{d}x \quad \langle f, g \rangle_{\text{Laguerre}} = \int_{\mathbb{R}} f(x) g(x) e^{-x} \, \mathrm{d}x \tag{5.120}$$

can use this to construct spaces of polynomials, here Hermite polynomials and Laguerre polynomials, but also others such as Bernstein polynomials. Note that here the Hermite polynomials are generated by the inner product with respect to a Gaussian measure.

Theorem 197. *Let $\{e_n\}_{n \geq 1}$ be a Hilbert space. Then the following hold.*

1. *For all $x \in H$, $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$, where the sum converges in H and $\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 < \infty$.*
2. *Let $u = (u_n)_{n \geq 1}$, with $u_n \in \mathbb{R}$ (or $\mathbb{C}$), such that $\sum_{n=1}^{\infty} |u_n|^2 < \infty$. Note that this is an ℓ^2 space.*

Then there exists an $x \in H$ such that the following hold.

1. *$\langle x, e_n \rangle = u_n$ and $x = \sum_{n=1}^{\infty} u_n e_n$ converges in H*
2. *$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2$.*

3. *For* $x = \sum_{n=1}^{\infty}\langle x, e_n\rangle e_n$ *and* $y = \sum_{n=1}^{\infty}\langle y, e_n\rangle e_n$*, we have* $\langle x, y\rangle = \sum_{n=1}^{\infty}\langle x, e_n\rangle\overline{\langle y, e_n\rangle}$ *where* $\overline{\cdot}$ *is the complex conjugate. We say that* $\langle x, e_n\rangle$ *is the* n*th Fourier coefficient of* x *in the basis* $(e_n)_{n\geq 1}$*.*

We conclude that is a non-unique isometry with H *and* ℓ^2*.*

Example 198. *Consider* $L^2((-\pi,\pi),\mathbb{C})$ *with*

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x. \tag{5.121}$$

Let $e_n(x) = e^{inx}$*, so that* $\{e_n\}_{n\in\mathbb{Z}}$ *is an basis. We can check orthonormality via*

$$\langle e_n, e_{n'}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\overline{e^{in'x}}\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-n')x}\,\mathrm{d}x = \begin{cases}1 & n = n'\\ 0 & n \neq n'\end{cases} \tag{5.122}$$

and completeness via Stone-Weierstrass Theorem. We conclude that $\{e_n\}_{n\in\mathbb{Z}}$ *is a Hilbert basis. We have*

$$\langle f, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,\mathrm{d}x = \hat{f}_n. \tag{5.123}$$

We have $f \in L^2((-\pi,\pi),\mathbb{C})$ *iff* $\sum_{n\in\mathbb{Z}}|\hat{f}_n|^2 < \infty$ *and* $f = \sum_{n\in\mathbb{Z}}\hat{f}_n e_n$ *means that* $\left\|f - \sum_{n\in\mathbb{Z}}\hat{f}_n e_n\right\|_{L^2} \xrightarrow{n\to\infty} 0$*. For other notions such as pointwise or uniform convergence, which do not involve the* L^2 *notion of almost everywhere, we need much stronger assumptions. Parseval's Identity says*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,\mathrm{d}x = \sum_{n\in\mathbb{Z}}\hat{f}_n\overline{\hat{g}_n} \qquad \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)|^2\,\mathrm{d}x = \sum_{n\in\mathbb{Z}}|\hat{f}_n|^2 \tag{5.124}$$

and we recover the classical integration results.

Chapter 6

Distribution Theory

6.1. Introduction

Let $f \in C^0(\mathbb{R})$. We view f not as a function not as a map between points, but using duality as linear map

$$C_c^\infty(\mathbb{R}) \xrightarrow{T_f} \mathbb{R} \qquad \phi \xmapsto{T_f} \int_M f(x)\phi(x)\,\mathrm{d}x \tag{6.1}$$

where $\phi \in C_c^\infty(\mathbb{R})$ is the test function. Since ϕ has compact support, the integral converges. The map T_f is a linear form on $C_c^\infty(\mathbb{R})$.

Suppose now that $f \in C^1(\mathbb{R})$. Then $f' \in C^0(\mathbb{R})$, so we can define a mapping

$$T_{f'}(\phi) = \int_M f'(x)\phi(x)\,\mathrm{d}x \tag{6.2}$$

and use integration parts to move the derivative over to ϕ. Using compact support,

$$T_{f'}(\phi) = \int_a^b f'(x)\phi(x)\,\mathrm{d}x = -\int_a^b f(x)\phi'(x)\,\mathrm{d}x + (f\phi)(b) - (f\phi)(a) = -\int_M f(x)\phi'(x)\,\mathrm{d}x \tag{6.3}$$

where $(f\phi)(b), (f\phi)(a)$ are the boundary terms which vanish because ϕ has compact support. Hence if $f \in C^1(\mathbb{R})$, we can consider making the above identification, with only the assumption $f \in C^0(\mathbb{R})$, for which the right-hand side of the formula above is well-defined because all derivatives are taken with respect to ϕ.

This gives the a notion of *distributional derivative*, which is defined for any sufficiently regular classical function, even if it is not differentiable. Understanding the space of test functions $C_c^\infty(\mathbb{R})$ is critical to seeing how to proceed.

6.2. Space of test functions

We begin by defining our space of test functions.

Definition 199. *Let $C_c^\infty(\mathbb{R}^d)$ be the set of infinitely differentiable functions with compact support.*

Remark 200. *Recall that for $f \in C^0(\mathbb{R}^d) = D(\mathbb{R}^d)$, $\operatorname{supp} f = \overline{\{x \in \mathbb{R}^d : f(x) \neq 0\}}$. Recall also that we can define a sequence of* mollifier *functions $\phi_m \in C_c^\infty(\mathbb{R}^d)$, such that*

$$\int_{\mathbb{R}^d} \phi_m(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} p^d \phi(px)\,\mathrm{d}x \tag{6.4}$$

where ϕ is a standard mollifier and the following hold.

1. $\phi_m \geq 0$.
2. $\operatorname{supp}\phi_m \subseteq \overline{B}(0, 1/p)$.
3. $\int_{\mathbb{R}^d} \phi_m(x)\,\mathrm{d}x = 1$.

We recall now that any sufficiently regular function can be approximated by mollifiers.

Proposition 201. *Let $f \in C^0(\mathbb{R}^d)$. Then, letting $*$ be convolution,*

$$f * \phi_m \in C^\infty(\mathbb{R}^d) \qquad\qquad f * \phi_m \xrightarrow{m\to\infty} f \tag{6.5}$$

where convergence is uniform on compact sets. Furthermore, if we also have $f \in C^\infty(\mathbb{R}^d)$, then

$$\partial^\alpha(f * \phi_m) = f * \partial^\alpha \phi_m \qquad \partial^\alpha(f * \phi_m) \xrightarrow{m\to\infty} \partial^\alpha f \tag{6.6}$$

uniformly on all compact sets. Finally, for $f \in L^p(\mathbb{R}^d)$, $1 \le p < \infty$,

$$f * \phi_m \in C^\infty(\mathbb{R}^d) \qquad \|f * \phi_m - f\|_{L^\infty(\mathbb{R}^d)} \xrightarrow{m\to\infty} 0 \tag{6.7}$$

where convergence is in $L^p(\mathbb{R}^d)$ norm.

A similar result holds for continuous functions differentiable finitely many times. Note also that the result is *not* uniform with respect the the order of the derivatives. Finally, note that the result is *not* true for $L^\infty(\mathbb{R}^d)$.

Corollary 202. *Let $D(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$. Then the following hold.*

1. *$D(\mathbb{R}^d)$ is dense in $C^0(\mathbb{R}^d)$ for the topology of uniform convergence on all compact sets of $\mathbb{R}^d$.*
2. *$D(\mathbb{R}^d)$ is dense in $C^\infty(\mathbb{R}^d)$ for the topology of uniform convergence of the function and all derivatives, both on all compact sets of $\mathbb{R}^d$.*
3. *$D(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$ for all $p \in [1, \infty)$.*

We now define a topology on $D(\mathbb{R}^d)$.

Definition 203. *Let $(f_n)_{n\ge1}$ be a sequence of functions in $D(\mathbb{R}^d)$. Then we say $f_n \xrightarrow{n\to\infty}$ in $D(\mathbb{R}^d)$ if and only if the following statements hold.*

1. *There is a compact set K such that $\operatorname{supp} f \subseteq K$, and $\operatorname{supp} f_n \subseteq K$ for all $n \ge 1$.*
2. *For all $\alpha \in \mathbb{N}^d$, $\partial^\alpha f_n \xrightarrow{n\to\infty} \partial^\alpha f$ uniformly on K, or equivalently by compactness of $\operatorname{supp} f_n$ uniformly on $\mathbb{R}^d$.*

This is not a nice topology – it is not a normed space, or even a Fréchet space.

6.3. Distributions

Definition 204. *A* distribution *on* $\mathbb{R}^d$ *is a continuous linear form on* $D(\mathbb{R}^d)$*. The space of distributions is denoted by* $D^*(\mathbb{R}^d)$ *where* $(\cdot)'$ *is the dual space. Hence,* $T \in D^*(\mathbb{R}^d)$ *iff the following hold.*

1. *T is a linear map $D(\mathbb{R}^d) \to \mathbb{R}$.*
2. *For all convergent sequences $(\phi_n)_{n\geq 1}$, $T(\phi_n) \xrightarrow{n\to\infty} T(\phi)$. This means that for all sequences $(\phi_n)_{n\geq 1}$ with $\phi_n \in D(\mathbb{R}^d)$ such that (i) there exists $\phi \in D(\mathbb{R}^d)$, and (ii) for all $\alpha \in \mathbb{N}^d$, $\partial^\alpha \phi_n \xrightarrow{n\to\infty} \partial^\alpha \phi$ uniformly on a compact set K, we have $T(\phi_n) \xrightarrow{n\to\infty} T(\phi)$.*

In general, this property is the only way we show that some object is a distribution.

Duality product for D

Definition 205. *If $T \in D^*(\mathbb{R}^d)$ and $\phi \in D(\mathbb{R}^d)$, then we write*

$$T(\phi) = \langle T \mid \phi \rangle_{D^*,D} = \langle T \mid \phi \rangle \tag{6.8}$$

for the duality product.

We can show that all sufficiently regular functions are distributions.

Example 206. *Take $f \in L^1_{loc}(\mathbb{R}^d)$. Note that this is a Fréchet space, but not a Banach space. Because* $\operatorname{supp}\phi \subseteq K$ *is compact, the integral*

$$\int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x = \int_K f(x)\phi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} (f\mathbb{1}_K)(x)\phi(x)\,\mathrm{d}x \tag{6.9}$$

is defined since $f\mathbb{1}_K \in L^1(\mathbb{R}^d)$ and $\phi \in D(\mathbb{R}) \subseteq L^\infty(\mathbb{R}^d)$. Consider

$$D(\mathbb{R}^d) \xrightarrow{T_f} \mathbb{R} \qquad\qquad \phi \xmapsto{T_f} \int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x. \tag{6.10}$$

This is by definition a linear map. We now check that it is continuous. Consider a sequence of functions $\phi_n \in D(\mathbb{R}^d)$ such that there exists $\phi \in D(\mathbb{R}^d)$ and the following hold.

1. *For all K compact,* $\operatorname{supp} \phi_n \subseteq K$ *for all* $n \geq 1$*, and* $\operatorname{supp} \phi \subseteq K$*.*
2. *For all $\alpha \in \mathbb{N}^d$, $\partial^\alpha \phi_n \xrightarrow{n\to\infty} \partial^\alpha \phi$ uniformly in K.*

We need to show that $\langle T_f \mid \phi_n \rangle \xrightarrow{n\to\infty} \langle T_f \mid \phi \rangle$. We aim to use the Dominated Convergence Theorem. We need to check two conditions.

1. *$f(x)\phi_n(x) \xrightarrow{n\to\infty} f(x)\phi(x)$ for a.e. $x \in \mathbb{R}^d$. This is trivially true because ϕ_n converges uniformly to ϕ and therefore the product, which is defined almost everywhere, converges pointwise.*
2. *There exists $g \in L^1(\mathbb{R}^d)$ such that $|f\phi_n| \leq g$ a.e. for all n in $\mathbb{R}^d$. Since $\phi_n \to \phi$ uniformly in $\mathbb{R}^d$, it also converges in $\|\cdot\|_{\infty,\mathbb{R}^d}$, so $(\phi_n)_{n\geq 1}$ is bounded for $\|\cdot\|_{\infty,\mathbb{R}^d}$. Thus there exists a constant c such that $\|\phi_n\|_{\infty,\mathbb{R}^d} \leq c$. This is clearly also true for ϕ. Furthermore, by compactness, $\phi_n \leq c\mathbb{1}_K(x)$, and thus $|(f\phi_n)(x)| \leq c\mathbb{1}_K(x)f(x)$ which is finite because $f \in L^1_{\text{loc}}(\mathbb{R}^d)$.*

We conclude that the DCT implies convergence. Thus T_f is a distribution.

A slightly different way to show that T_f as defined previously is to instead show

$$|\langle T_f \mid \phi_n \rangle - \langle T_f \mid \phi \rangle| = \left| \int_{\mathbb{R}^d} f(x)(\phi_n(x) - \phi(x)) \, \mathrm{d}x \right| \tag{6.11}$$

$$\leq \int_{\mathbb{R}^d} |f(x)||\phi_n(x) - \phi(x)| \, \mathrm{d}x \tag{6.12}$$

$$\leq \int_{\mathbb{R}^d} |f(x)|\|\phi_n - \phi\|_{\infty,\mathbb{R}^d} \mathbb{1}_K(x) \, \mathrm{d}x \tag{6.13}$$

$$\leq \|\phi_n - \phi\|_{\infty,\mathbb{R}^d} \int_K |f(x)| \, \mathrm{d}x \tag{6.14}$$

to conclude that the sequence converges.

Proposition 207. *The mapping*

$$L^1_{loc} \xrightarrow{\mathcal{T}} D^*(\mathbb{R}^d) \qquad f \xmapsto{\mathcal{T}} T_f \tag{6.15}$$

is injective, i.e. it is one-to-one.

Proof. The mapping is linear, so it suffices to check that its kernel is zero. Suppose that $T_f = 0$. Then for all $\phi \in D(\mathbb{R}^d)$,

$$\langle T_f \mid \phi \rangle = \int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x = 0 \tag{6.16}$$

which implies $f = 0$ a.e. and the proposition follows. ■

This means that we can identify T_f with f, and speak of locally integrable functions as distributions.

Not all distributions can be represented by functions. The next example shows one that cannot.

Example 208. *Let $x_0 \in \mathbb{R}^d$. Define the linear form*

$$D(\mathbb{R}^d) \xrightarrow{\delta_{x_0}} \mathbb{R} \qquad\qquad \phi \xmapsto{\delta_{x_0}} \phi(x_0). \tag{6.17}$$

Note that since $\phi \in D(\mathbb{R}^d)$, its pointwise value is well-defined, as no almost everywhere equivalence relation is used in its definition. To show this is a distribution, we need to show that it is continuous it is continuous with respect to the topology of $D(\mathbb{R}^d)$. We also show it cannot be represented by a function. We first show continuity. Take a sequence $(\phi_n)_{n\geq 1}$, $\phi_n \in D(\mathbb{R}^d)$ such that there exists $\phi \in D(\mathbb{R}^d)$ with

$$\phi_n \xrightarrow{n\to\infty} \phi \tag{6.18}$$

in $D(\mathbb{R}^d)$. We want to show that $\langle \delta_{x_0} \mid \phi_n \rangle \xrightarrow{n\to\infty} \langle \delta_{x_0} \mid \phi \rangle$. This is obviously true because $\phi_n \xrightarrow{n\to\infty} \phi$ uniformly which implies pointwise convergence, and $\langle \delta_{x_0} \mid \phi \rangle = \phi(x_0)$. We now want to show there does not exist $f \in L^1_{loc}(\mathbb{R}^d)$ such that $\delta_{x_0} = T_f$. Suppose for contradiction that such a function exists. Then for all $\phi \in D(\mathbb{R}^d)$ we have

$$\langle \delta_{x_0} \mid \phi \rangle = \langle T_f \mid \phi \rangle \tag{6.19}$$

for some $f \in L^1_{loc}(\mathbb{R}^d)$. Hence

$$\phi(x_0) = \int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x. \tag{6.20}$$

Take ϕ such that $x_0 \notin \operatorname{supp}\phi$ – for any x_0, such a ϕ exists. Then

$$\int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x = 0. \tag{6.21}$$

Then for the open set $D = \mathbb{R}^d \setminus \{x_0\}$, we have

$$\int_M f(x)\phi(x)\,\mathrm{d}x = 0 \qquad\qquad \forall\phi \in C_c^\infty(M) \tag{6.22}$$

hence

$$f(x) = 0 \qquad\qquad \text{a.e. } x \in M. \tag{6.23}$$

But $\{x_0\}$ is a set of measure zero, so

$$f(x) = 0 \qquad\qquad \text{a.e. } x \in \mathbb{R}^d. \tag{6.24}$$

We conclude that

$$\int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x = 0 \tag{6.25}$$

for all $\phi \in D(\mathbb{R}^d)$, now irrespective of whether or not $x_0 \in \operatorname{supp}\phi$. But this means that

$$\phi(x_0) = \int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x = 0 \tag{6.26}$$

for all ϕ, which is a contradiction, since any mollifier centered at x_0 will not satisfy $\phi(x_0) = 0$. Thus δ_{x_0} is not a function in the classical sense.

6.4. Convergence and differentiability

We now want to define a topology on $D^*(\mathbb{R}^d)$. Since the topology on $D(\mathbb{R}^d)$ is complicated, so is the topology on $D^*(\mathbb{R}^d)$. We thus define what it means for a sequence of distributions to converge.

Definition 209. *Let $(T_n)_{n\geq 1}$ be a sequence of distributions, and let T is a distribution. We say that*

$$T_n \xrightarrow{n\to\infty} T \tag{6.27}$$

iff for all ϕ

$$T_n(\phi) \xrightarrow{n\to\infty} T(\phi). \tag{6.28}$$

This is just pointwise convergence: it means that

$$\forall \phi \in D(\mathbb{R}^d), \forall \varepsilon > 0, \exists N \in N \text{ s.t. } \forall n \geq N, |\langle T_n \mid \phi\rangle - \langle T \mid \phi\rangle| < \varepsilon. \tag{6.29}$$

Note that pointwise convergence is a fairly weak requirement, so we speak of weak convergence.

Example 210. *Suppose* $(f_n)_{n\geq 1}$ *is a sequence with* $f_n \in L^1_{loc}(\mathbb{R}^d)$ *such that there exists* $f \in L^1_{loc}(\mathbb{R}^d)$ *and for all* $K \subseteq \mathbb{R}^d$ *compact, we have*

$$\|f_n - f\|_{L^1(K)} \xrightarrow{n\to\infty} 0. \tag{6.30}$$

Then for the distributions T_{f_n}, T_f *associated with these functions, we have*

$$T_{f_n} \to T_f \tag{6.31}$$

in $D^*(\mathbb{R}^d)$*. Note that this means that the injective map* $\mathcal{T}$ *defined previously also continuous. Hence the topology of distributions is weaker than the topology of* $L^1_{loc}(\mathbb{R}^d)$*. The converse isn't true: an open set in* $L^1_{loc}(\mathbb{R}^d)$ *is not necessarily an open* $D^*(\mathbb{R}^d)$*. To show this, we need to show that*

$$\langle T_{f_n} \mid \phi\rangle \xrightarrow{n\to\infty} \langle T_f \mid \phi\rangle \tag{6.32}$$

which means

$$\int_{\mathbb{R}^d} f_n(x)\phi(x)\,\mathrm{d}x \xrightarrow{n\to\infty} \int_{\mathbb{R}^d} f(x)\phi(x)\,\mathrm{d}x. \tag{6.33}$$

We have

$$\begin{aligned}
\left|\int_{\mathbb{R}^d} (f_n - f)(x)\phi(x)\,\mathrm{d}x\right| &\leq \int_{\mathbb{R}^d} |(f_n - f)(x)|\phi(x)\,\mathrm{d}x & (6.34)\\
&\leq \int_{\mathbb{R}^d} |(f_n - f)(x)|\|\phi\|_{\infty,\mathbb{R}^d}\mathbb{1}_K(x)\,\mathrm{d}x & (6.35)\\
&\leq \|\phi\|_{\infty,\mathbb{R}^d} \int_K |(f_n - f)(x)|\,\mathrm{d}x & (6.36)
\end{aligned}$$

where convergence of the integral follows from convergence of the functions as previously.

Remark 211. *Recall that a topology O_2 is weaker than O_1 iff $O_2 \subseteq O_1$, so any open set in O_2 is also an open set in O_1. This is, for sufficiently regular topologies, equivalent to the statement that the identity map i between two topologies*

$$(X, O_1) \xrightarrow{i} (X, O_2) \qquad\qquad x \overset{i}{\mapsto} x \tag{6.37}$$

is continuous.

Example 212. *Take a sequence of points $(x_n)_{n\geq 1}$ such that $x_n \xrightarrow{n\to\infty} x$ in $\mathbb{R}^d$. Then $\delta_{x_n} \xrightarrow{n\to\infty} \delta_x$ in $D^*(\mathbb{R}^d)$.*

Example 213. *Take $(\phi_m)_{m\geq 1}$ be a sequence of shrinking mollifiers centered at zero. Then $\phi_m \xrightarrow{m\to\infty} \delta_0$ in $D^*(\mathbb{R}^d)$. To show this, we need to check that for $f \in D(\mathbb{R}^d)$, we have*

$$\int_{\mathbb{R}^d} \phi_m(x) f(x)\,\mathrm{d}x \xrightarrow{m\to\infty} f(0). \tag{6.38}$$

But we have

$$\left|\int_{\mathbb{R}^d} \phi_m(x) f(x)\,\mathrm{d}x - f(0)\right| \leq \left|\int_{\mathbb{R}^d} \phi_m(x) f(x)\,\mathrm{d}x - f(0)\int_{\mathbb{R}^d} \phi_m(x)\,\mathrm{d}x\right| \tag{6.39}$$

$$\leq \left|\int_{\mathbb{R}^d} \phi_m(x) |f(x) - f(0)|\,\mathrm{d}x\right|. \tag{6.40}$$

since f is continuous at 0, meaning that for any ε there exists a η such that $x - 0 < \delta$ implies $|f(x) - f(0)| < \varepsilon$. For $M > 1/\eta$, and all $m > M$, we have $1/m < \eta$. Then we have

$$\left|\int_{\mathbb{R}^d} \phi_m(x)|f(x) - f(0)|\,\mathrm{d}x\right| = \left|\int_{\overline{B}(0,1/m)} \phi_m(x)|f(x) - f(0)|\,\mathrm{d}x\right| \leq \varepsilon \int_{\overline{B}(0,1/m)} \phi_m(x)\,\mathrm{d}x \leq \varepsilon \tag{6.41}$$

and we have shown that for all $\varepsilon > 0$ there exists an $M \in \mathbb{N}$ such that for all $m > M$

$$\left|\int_{\mathbb{R}^d} \phi_m(x) f(x)\,\mathrm{d}x - f(0)\right| < \varepsilon \tag{6.42}$$

which means

$$|\langle \phi_m \mid f\rangle - \langle \delta_0 \mid f\rangle| \le \varepsilon \tag{6.43}$$

as required, provided we identify the function $\phi_m \in L^1_{loc}(\mathbb{R}^d)$ with its distribution.

A similar density result is available for distributions.

Result 214. *$D(\mathbb{R}^d)$ is dense in $D^*(\mathbb{R}^d)$.*

This is proven using convolution of distributions, which is not presented here.

We now define differentiability.

Distributional derivative

Definition 215. *Let $T \in D^*(\mathbb{R}^d)$ and $i \in \{1, .., d\}$. Define*

$$D(\mathbb{R}^d) \xrightarrow{\partial_{x_i} T} \mathbb{R} \qquad\qquad \phi \xmapsto{\partial_{x_i} T} -\langle T \mid \partial_{x_i}\phi\rangle. \tag{6.44}$$

Observe that $\langle \partial_{x_i} T \mid \phi\rangle = -\langle T \mid \partial_{x_i}\phi\rangle$, and that if $T = T_f$ with $f \in C^1(\mathbb{R}^d)$, then

$$\begin{aligned}
-\langle T_f \mid \partial_{x_i}\phi\rangle &= -\int_{\mathbb{R}^d} f(x)\partial_{x_i}\phi(x)\,\mathrm{d}x && (6.45)\\
&= -\int_{B(0,R)} f(x)\partial_{x_i}\phi(x)\,\mathrm{d}x && (6.46)\\
&= \int_{B(0,R)} \partial_{x_i} f(x)\phi(x)\,\mathrm{d}x - \int_{\partial\overline{B}(0,R)} f(x)\phi(x)n_i(x)\,\mathrm{d}S(x) && (6.47)\\
&= \int_{\mathbb{R}^d} \partial_{x_i} f(x)\phi(x)\,\mathrm{d}x && (6.48)\\
&= \langle \partial_{x_i} T_f \mid \phi\rangle && (6.49)
\end{aligned}$$

where $n_i(x)$ is the unit vector pointing outward on the boundary. We have used that ϕ is compactly supported, that $C^1(\mathbb{R}^d) \subset C^0(\mathbb{R}^d) \subset L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and that the boundary is a smooth $(d-1)$-dimensional manifold, for which we can define a surface measure $\mathrm{d}S(x)$. $\mathrm{d}S(x)$ is defined by defining tubular neighborhood V^ε, for which we can define the Lebesgue measure $\lambda(V^\varepsilon)$, and then the surface measure $S(x) = \lambda(V^\varepsilon)/\varepsilon$.

Thus the definition generalizes Green's formula

$$\int_M f(x)\partial_{x_i}\phi(x)\,\mathrm{d}x = -\int_M \partial_{x_i} f(x)\phi(x)\,\mathrm{d}x + \int_{\partial M} f(x)\phi(x)n_i(x)\,\mathrm{d}S(x) \tag{6.50}$$

to distributions and distributional derivatives in a way that is consistent with classical functions.

It is obvious that

$$\partial_{x_i} T \tag{6.51}$$

is a linear form on $D(\mathbb{R}^d)$. We now need to show continuity. Suppose $\phi_n \to \phi$ in $D(\mathbb{R}^d)$. We need to show

$$\langle \partial_{x_i} T \mid \phi_n \rangle \xrightarrow{n\to\infty} \langle \partial_{x_i} T \mid \phi \rangle \tag{6.52}$$

which is equivalent to

$$-\langle T \mid \partial_{x_i} \phi_n \rangle \xrightarrow{n\to\infty} -\langle T \mid \partial_{x_i} \phi \rangle \tag{6.53}$$

which is true because the following hold.

1. There exists compact $K \subseteq \mathbb{R}^d$ such that $\operatorname{supp}\phi_n \subseteq K$ for all $n \geq 1$ and $\operatorname{supp}\phi \subseteq K$.
2. For all $\alpha \in \mathbb{N}^d$, $\partial^\alpha \phi_n \xrightarrow{n\to\infty} \partial^\alpha \phi$ uniformly on K (or on $\mathbb{R}^d$.

The first property is true because $\phi_n = 0$ on $\mathbb{R}^d \setminus K$ implies $\partial_{x_i}\phi_n = 0$ on $\mathbb{R}^d \setminus K$. The second property is true because $\partial^\alpha \partial_{x_i} \phi_n \xrightarrow{n\to\infty} \partial^\alpha \partial_{x_i} \phi$ uniformly on K, which is obvious, because there exists a multi-index such that $\partial^\alpha \partial_{x_i}$ is itself a partial derivative.

We can extend the definition inductively to arbitrary order.

Definition 216. *For any multi-index $\alpha \in \mathbb{N}^d$ with length $|\alpha|$, $T \in D^*(\mathbb{R}^d)$, and all $\phi \in D(\mathbb{R}^d)$, define*

$$\langle \partial^\alpha T \mid \phi \rangle = (-1)^{|\alpha|} \langle T \mid \partial^\alpha \phi \rangle. \tag{6.54}$$

As before, this extends the notion of a higher order derivative of a smooth function.

Example 217. *Consider δ_{x_0}, with $\langle \delta_{x_0} \mid \phi \rangle = \phi(x_0)$. Then we have*

$$\langle \partial^\alpha \delta_{x_0} \mid \phi \rangle = (-1)^\alpha \partial^\alpha \phi(x_0). \tag{6.55}$$

Note that all distributions have a distributional derivative, and all are infinitely distributionally differentiable. On the other hand, as will be subsequently shown, we cannot in general define multiplication of distributions.

Proposition 218. *Distributional differentiation is a continuous operation, in the sense that if $T_n \to T$ in $D^*(\mathbb{R}^d)$, we have*

$$\partial^\alpha T_n \xrightarrow{n\to\infty} \partial^\alpha T \tag{6.56}$$

where convergence is in the sense that $\langle \partial^\alpha T_n \mid \phi \rangle \xrightarrow{n\to\infty} \langle \partial^\alpha T \mid \phi \rangle$ for all ϕ.

Proof. We note that

$$\lim_{n\to\infty} \langle \partial^\alpha T_n \mid \phi \rangle = \lim_{n\to\infty} (-1)^{|\alpha|} \langle T_n \mid \partial^\alpha \phi \rangle = (-1)^{|\alpha|} \langle T \mid \partial^\alpha \phi \rangle = \langle \partial^\alpha T \mid \phi \rangle \tag{6.57}$$

because if $\phi \in D(\mathbb{R}^d)$ then $\partial^\alpha \phi \in D(\mathbb{R}^d)$. ■

Example 219. *Let's try to find all $T \in D^*(\mathbb{R})$ such that*

$$\frac{\mathrm{d}T}{\mathrm{d}x} = 0. \tag{6.58}$$

This is equivalent to, for all $\phi \in D(\mathbb{R})$, that

$$\left\langle \frac{\mathrm{d}T}{\mathrm{d}x} \;\middle|\; \phi \right\rangle = 0 \quad \Longrightarrow \quad -\left\langle T \;\middle|\; \frac{\mathrm{d}\phi}{\mathrm{d}x} \right\rangle = 0 \quad \Longrightarrow \quad -\langle T \mid \psi \rangle = 0 \tag{6.59}$$

for all $\psi \in D(\mathbb{R})$ for which there exists $\phi \in D(\mathbb{R})$ for which $\psi = \frac{\mathrm{d}\phi}{\mathrm{d}x}$. Let

$$S = \left\{ \psi \in D(\mathbb{R}) : \exists \phi \in D(\mathbb{R}) \text{ s.t. } \frac{\mathrm{d}\phi}{\mathrm{d}x} = \psi \right\}. \tag{6.60}$$

We have

$$\int_a^b \psi \,\mathrm{d}x = \int_a^b \frac{\mathrm{d}\phi}{\mathrm{d}x} \,\mathrm{d}x = \phi(b) - \phi(a) = 0 \tag{6.61}$$

where a, b are the limits of the support of ϕ. Hence every $\psi \in S$ has integral zero. Define

$$\phi(x) = \int_a^x \psi(y) \,\mathrm{d}y \tag{6.62}$$

which gives ϕ compactly supported because the integral of ψ is zero. Therefore

$$S = \Big\{\psi \in D(\mathbb{R}) : \int_{-\infty}^{\infty} \psi(x) = 0\Big\}. \tag{6.63}$$

For any ϕ, we can construct a ψ with integral zero, by writing

$$\psi = \phi - \Big(\int_{-\infty}^{\infty} \phi \,\mathrm{d}x\Big)\phi_1 \tag{6.64}$$

where ϕ_1 is a mollifier with integral one. Then using linearity, we get

$$\begin{aligned} &\langle T \mid \phi \rangle = 0 && (6.65)\\ \implies &\Big\langle T \;\Big|\; \phi - \Big(\int_{-\infty}^{\infty} \phi \,\mathrm{d}x\Big)\phi_1 \Big\rangle = 0 && (6.66)\\ \implies &\langle T \mid \phi \rangle = \langle T \mid \phi_1 \rangle \int_{-\infty}^{\infty} \phi \,\mathrm{d}x && (6.67)\\ \implies &\langle T \mid \phi \rangle = \langle C \mid \phi \rangle && (6.68) \end{aligned}$$

for some constant C, since $\langle T \mid \phi_1 \rangle$ and $\int_{-\infty}^{\infty} \phi \,\mathrm{d}x$ are both constants. This implies that $T = C$. Conversely, it is obvious that every constant function satisfies the differential equation.

For $f \in L^1_{\mathrm{loc}}, g \in C^\infty(\mathbb{R}^d)$, we have

$$\langle T_{fg} \mid \phi \rangle = \int_M f(x)g(x)\phi(x) \,\mathrm{d}x = \langle T_f \mid g\phi \rangle \tag{6.69}$$

which we'd like to generalize.

Definition 220. *Take $T \in D^*(\mathbb{R}^d), g \in C^\infty(\mathbb{R}^d)$, then we can define*

$$\langle gT \mid \phi \rangle = \langle T \mid g\phi \rangle \qquad \forall \phi \in D(\mathbb{R}^d). \tag{6.70}$$

Remark 221. *The above definition cannot be extended generically, which limits the applicability of the theory to non-linear partial differential equations. In particular, we cannot multiply two Dirac delta distributions centered at the same point.*

6.5. Convolution of distributions

Tensor product

Definition 222. *Let $T \in D^*(\mathbb{R}^{d_1})$, $S \in D^*(\mathbb{R}^{d_2})$. The* TENSOR PRODUCT *$S \otimes T$ is the distribution in $D^*(\mathbb{R}^{d_1+d_2})$ such that for all $\phi \in D(\mathbb{R}^{d_1})$, $\psi \in D\mathbb{R}^{d_2})$, we have*

$$\langle S \otimes T \mid \phi \otimes \psi \rangle = \langle S \mid \phi \rangle \langle T \mid \psi \rangle \tag{6.71}$$

where $\phi \otimes \psi(x_1, x_2) = \phi(x_1)\psi(x_2)$. Since the space generated by tensor products $\phi \otimes \psi$ is dense in $D(\mathbb{R}^{d_1+d_2})$ by the Stone-Weierstrass Theorem, this definition unambiguously gives a unique distribution in $D^*(\mathbb{R}^{d_1+d_2})$.

Definition 223. *Let $T \in D^*(\mathbb{R}^d)$, $S \in D^*(\mathbb{R}^d)$ such that T and S have compact support. Then the* CONVOLUTION *$T * S$ is the distribution in $D^*(\mathbb{R}^d)$ defined by*

$$\langle S \otimes T \mid \phi \rangle = \langle S \otimes T \mid \phi(x+y) \rangle \tag{6.72}$$

for all $\phi \in D(\mathbb{R}^d)$.

Note here that the map $(x, y) \mapsto \phi(x+y)$ is not in $D(\mathbb{R}^{2d})$ because it does not necessarily have compact support, so we define $\langle S \otimes T \mid \phi(x+y) \rangle$ formally as follows. Let $\psi \in D(\mathbb{R}^d)$ be any function such that $x \in \operatorname{supp} T \cup \operatorname{supp} S$ implies $\psi(x) = 1$. Then define

$$\langle S \otimes T \mid \phi(x+y) \rangle = \langle S \otimes T \mid \phi(x+y)\psi(x)\psi(y) \rangle. \tag{6.73}$$

This function now has compact support, and since the support of S and T are compact, its value does not depend on the choice of ψ.

Obviously, $T * S = S * T$.

One can relax the assumptions of T and S having both compact support and replace it by a weaker assumption of *adapter support*. The property we need is that $(\operatorname{supp} T) \times (\operatorname{supp} S)$ in $\mathbb{R}^d$ intersects and set $\tilde{K} = \{(x, y) \in \mathbb{R}^{2d}, x + y \in K, K \text{ compact}\}$ on a compact set. In particular, it is enough that S or T has compact support.

Proposition 224. *The following hold.*

1. *Let $T, S \in D^*(\mathbb{R}^d)$ with adapted support for convolution. Let $\alpha \in \mathbb{N}^d$. Then $\partial^\alpha (T * S) = (\partial^\alpha T) * S = T * (\partial^\alpha S)$.*
2. *Suppose in addition that $S \in C^k(\mathbb{R}^d)$, then $T * S \in C^k(\mathbb{R}^d)$ and $T*S(x) = \langle T \mid S(x - \cdot)\rangle$, $\partial^\alpha (T*S)(x) = \langle T \mid \partial^\alpha S(x - \cdot)\rangle$, for $|\alpha| \leq k$.*
3. *Let $\phi_p(x)$ be a sequence of mollifiers. Then $T * \phi_p \in C^\infty(\mathbb{R}^d)$ and $T * \phi_p \xrightarrow{p \to \infty} T$ in $D^*(\mathbb{R}^d)$.*
4. *The space $D(\mathbb{R}^d)$ is dense in $D^*(\mathbb{R}^d)$.*

Definition 225. *Let $g \in C^\infty(\mathbb{R}^d)$ and $T \in D^*(\mathbb{R}^d)$. Then $gT \in D^*(\mathbb{R}^d)$ is defined by*

$$\langle gT \mid \phi\rangle = \langle T \mid g\phi\rangle \tag{6.74}$$

for all $\phi \in D(\mathbb{R}^d)$.

6.6. Fourier transform and tempered distributions

Everything presented so far can be extended from distributions taking values in $\mathbb{R}$ to distributions taking values in $\mathbb{C}$, i.e. continuous *anti-linear* forms on $D(\mathbb{R}^d, \mathbb{C})$. In particular, for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d, \mathbb{C})$, we define $T_f \in D^*(\mathbb{R}^d, \mathbb{C})$ by

$$\langle T_f \mid \phi\rangle = \int_{\mathbb{R}^d} f(x)\overline{\phi(x)}\,\mathrm{d}x \tag{6.75}$$

for all $\phi \in D(\mathbb{R}^d, \mathbb{C})$. Similarly, $\langle \delta_{x_0} \mid \phi\rangle = \overline{\phi(x_0)}$. For this section, we work with complex-valued distributions.

Definition 226. *For $f \in L^1(\mathbb{R}^d)$, we define the* Fourier transform *$\widehat{f}$ or $\mathcal{F}f$ by*

$$\widehat{f}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) e^{-i\langle x \mid \xi\rangle} \,\mathrm{d}x. \tag{6.76}$$

We have $\widehat{f} \in C^0(\mathbb{R}^d)$ by the dominated convergence theorem. In fact, $widehat f \in C^0_0(\mathbb{R}^d)$, the subspace of continuous functions on $\mathbb{R}^d$, tending to zero at infinity, by the Riemann-Lebesgue Lemma. Note that $\left\|\widehat{f}\right\|_\infty \leq c\|f\|_1$.

The Fourier transform transforms convolutions into products, and derivatives into multiplication by polynomials. It is a very convenient tool to find analytic solutions of constant coefficient PDEs. To make it effective, we need to extend its definition to more general objects such as distributions.

The main obstacle to this is that for general distributions, we do not have any control of growth at infinity. For instance, $f(x) = e^{x^2}$ is a locally integrable function, and therefore defines a distribution. Since the Fourier transform involves an integral, some control at infinity as required. We thus introduce a subspace of distributions $\mathcal{S}'(\mathbb{R}^d)$, called the *tempered distributions*.

Definition 227. *The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is the space of $C^\infty(\mathbb{R}^d)$ functions such that for any $\alpha \in \mathbb{N}^d$ and all $p \in \mathbb{N}$, we have*

$$\sup_{x\in\mathbb{R}^d} |(1+|x|)^p||\partial^\alpha f(x)| < \infty. \tag{6.77}$$

This means that $f \in \mathcal{S}(\mathbb{R}^d)$ iff f and its derivatives of arbitrary order decay at infinity faster than the inverse of any polynomial.

Definition 228. *Let $(\phi_n)_{n\in\mathbb{N}}$ be a sequence of functions in $\mathcal{S}(\mathbb{R}^d)$ and let $\phi \in \mathcal{S}(\mathbb{R}^d)$. Then*

$$\phi_n \xrightarrow{n\to\infty} \phi \tag{6.78}$$

in $\mathcal{S}(\mathbb{R}^d)$ if and only if

$$\sup_{x\in\mathbb{R}^d} |(1+|x|)^p||\partial^\alpha f(x)| \to 0 \tag{6.79}$$

for all $p \in \mathbb{N}$ *and all* $\alpha \in \mathbb{N}^d$.

We are ready to introduce the Fourier transform.

Definition 229. *Let* $\phi \in \mathcal{S}(\mathbb{R}^d)$. *Then the* Fourier transform $\widehat{\phi}$ *or* $\mathcal{F}\phi$ *is defined by*

$$\widehat{\phi}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \phi(x) e^{-i\langle x \mid \xi \rangle} \, \mathrm{d}x. \tag{6.80}$$

Proposition 230. $\mathcal{F}$ *is a continuous linear map from* $\mathcal{S}(\mathbb{R}^d)$ *into* $\mathcal{S}(\mathbb{R}^d)$ *which satisfies the following for all* $\phi \in \mathcal{S}(\mathbb{R}^d)$.

1. $\widehat{\partial_x^\alpha \phi}(\xi) = i^{|\alpha|} \xi^\alpha \widehat{\phi}(\xi)$ *for all* $\alpha \in \mathbb{N}^d$.
2. $\widehat{\partial^\alpha \phi}(\xi) = i^{|\alpha|} \partial_\xi^\alpha \widehat{\phi}(\xi)$ *for all* $\alpha \in \mathbb{N}^d$.

Proof. The identities are obvious using Lebesgue's Theorem, so we only show continuity. We have

$$(1+|\xi|^2)^m = \partial_\xi^\alpha \widehat{\phi}(\xi) = \mathcal{F}((1-\Delta_x)^m((-ix)^\alpha \phi))(1-\Delta_x)^m((-ix)^\alpha \phi) \tag{6.81}$$

which is in $\mathcal{S}(\mathbb{R}^d)$. Thus $\mathcal{F}((1-\Delta_x)^m(-ix)^\alpha \phi)$ is bounded, which shows that $\widehat{\phi} \in \mathcal{S}(\mathbb{R}^d)$. Furthermore,

$$\sup_{\xi \in \mathbb{R}^d} \left| (1+|\xi|^2)^m \partial_\xi^\alpha \widehat{\phi}(\xi) \right| \le \left\| (1-\Delta_x)^m((-ix)^\alpha \phi) \right\|_1 \tag{6.82}$$

$$\le c \sup_{x \in \mathbb{R}^d} \left\{ (1+|x|)^{d+1} |(1-\Delta_x)^m((-ix)^\alpha \phi)| \right\}. \tag{6.83}$$

Therefore, if $(\phi_n)_{n\in\mathbb{N}}$ satisfies $\phi_n \xrightarrow{n\to\infty} 0$ in $\mathcal{S}(\mathbb{R}^d)$, we have $\widehat{\phi}_n \xrightarrow{n\to\infty} 0$ in $\mathcal{S}(\mathbb{R}^d)$ as well. Thus $\mathcal{F}$ is continuous on $\mathcal{S}(\mathbb{R}^d)$. ■

Proposition 231. *If* $\phi, \psi \in \mathcal{S}(\mathbb{R}^d)$ *then* $\phi * \psi \in \mathcal{S}(\mathbb{R}^d)$ *and* $\widehat{\phi * \psi}(\xi) = (2\pi)^{d/2} \widehat{\phi}(\xi) \widehat{\psi}(\xi)$.

We now define the Fourier transform for tempered distributions.

Definition 232. *A* tempered distribution *is a continuous linear form on $\mathcal{S}(\mathbb{R}^d)$. Thus, a linear map*

$$T : \mathbb{R}^\lceil \to \mathbb{R} \qquad\qquad T : \phi \mapsto \langle T \mid \phi \rangle \tag{6.84}$$

is a tempered distribution iff for any sequence of functions $(\phi_n)_{n\in\mathbb{N}}$ in $\mathcal{S}(\mathbb{R}^d)$ with $\phi \in \mathcal{S}(\mathbb{R}^d)$,

$$\phi_n \xrightarrow{n\to\infty} \phi \tag{6.85}$$

in $\mathcal{S}(\mathbb{R}^d)$ implies

$$\langle T \mid \phi_n \rangle \xrightarrow{n\to\infty} \langle T \mid \phi \rangle. \tag{6.86}$$

The definition for $\mathbb{C}$ is analogous, except that we require that it is antilinear, i.e. $\langle T \mid \lambda\phi \rangle = \overline{\lambda}\langle T \mid \phi \rangle$ for all $\phi \in \mathcal{S}(\mathbb{R}^d)$ and $\lambda \in \mathbb{C}$.

Example 233. *The following are tempered distributions.*

1. *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ such that there exist $p \in \mathbb{N}$, $R, C > 0$, and $|x| \geq R$ implies $|f(x)| \leq C(1+|x|)^p$. Then $T_f : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ defined by $T_f : \phi \mapsto \int_{\mathbb{R}} f(x)\overline{\phi}(x)\,\mathrm{d}x$ is a tempered distribution.*
2. *δ_{x_0} is a tempered distribution.*
3. *$VP(1/x)$ is a tempered distribution.*
4. *e^{x^2} is* not *a tempered distribution.*

Remark 234. *The topology of the tempered distribution space is Hausdorff, so if we have $T_n \xrightarrow{n\to\infty} T$, then the limit is unique.*

We have the following.

1. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence such that $f_n \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Then we have the following.
   (a) There exists $p \in \mathbb{N}$, $R, C > 0$, and $|x| \geq R$ implies $|f_n(x)| \leq C(1+|x|)^p$ for all $n \in \mathbb{N}$.

(b) For all K compact, if $f_n|_k \to f|k$ in $L^1(K)$, then $T_{f_n} \to T_f$ in $\mathcal{S}'(\mathbb{R}^d)$.

2. The sequence ϕ_p of mollifiers satisfies $\phi_p \to \delta$ in $\mathcal{S}'(\mathbb{R}^d)$.
3. Let $x_n \in \mathbb{R}^d$ with $x_n \xrightarrow{n\to\infty} x_0$. Then $\delta_{x_n} \xrightarrow{n\to\infty} \delta_{x_0}$ in $\mathcal{S}'(\mathbb{R}^d)$, and in $D^*(\mathbb{R}^d)$ as well.
4. The space $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathcal{S}'(\mathbb{R}^d)$.

Definition 235. *Let $g \in C^\infty(\mathbb{R}^d)$ such that for all $\alpha \in \mathbb{N}^d$ there exists $p_\alpha \in \mathbb{N}$ and $c_\alpha > 0$ with $|\partial^\alpha g(x)| \leq c_\alpha(1+|x|)^{p_\alpha}$. Then for all T in $\mathcal{S}'(\mathbb{R}^d)$, gT can be defined in $\mathcal{S}'(\mathbb{R}^d)$ by*

$$\langle gT \mid \phi\rangle = \langle T \mid g\phi\rangle \tag{6.87}$$

for all $\phi \in \mathcal{S}\mathbb{R}^d$.

With g satisfying this property, we have that $\phi \in \mathcal{S}(\mathbb{R}^d)$ implies $g\phi \in \mathcal{S}(\mathbb{R}^d)$.

For $\phi, \psi \in \mathcal{S}(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} \mathcal{F}\phi(\xi)\overline{\psi}(\xi)\,\mathrm{d}\xi = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \phi(x)e^{-i\langle x \mid \xi\rangle}\overline{\psi}(\xi)\,\mathrm{d}\xi \tag{6.88}$$

$$= \int_{\mathbb{R}^d} \phi(x)\overline{\int_{\mathbb{R}^d} \psi(\xi)e^{i\langle x \mid \xi\rangle}\,\mathrm{d}\xi}\,\mathrm{d}x \tag{6.89}$$

$$= \int_{\mathbb{R}^d} \phi(x)\overline{\overline{\mathcal{F}}\xi}(x)\,\mathrm{d}x \tag{6.90}$$

where the conjugate Fourier transform $\overline{\mathcal{F}}$ is defined by

$$\overline{\mathcal{F}}\psi(x) = \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} e^{i\langle x \mid \xi\rangle}\,\mathrm{d}\xi. \tag{6.91}$$

Therefore, using the $L^2(M)$ inner product, we have

$$\langle \mathcal{F}\phi, \psi\rangle = \left\langle \phi, \overline{\mathcal{F}}\right\rangle. \tag{6.92}$$

This motivates the following definition.

Definition 236. *Let $T \in \mathcal{S}'(\mathbb{R}^d)$. The Fourier transform $\widehat{T}$ or $\mathcal{F}T$ of T is the tempered distribution defined by*

$$\langle \mathcal{F}T \mid \phi \rangle = \left\langle T \,\middle|\, \overline{\mathcal{F}}\phi \right\rangle \tag{6.93}$$

for all $\phi \in \mathcal{S}(\mathbb{R}^d)$.

Proposition 237. *$\mathcal{F}$ is a continuous linear map from $\mathcal{S}'(\mathbb{R}^d)$ into $\mathcal{S}'(\mathbb{R}^d)$ which satisfies the following for all $T \in \mathcal{S}'(\mathbb{R}^d)$.*

1. *$\mathcal{F}(\partial_x T)(\xi) = i^{|\alpha|}\xi^\alpha \mathcal{F}T(\xi)$ for all $\alpha \in \mathbb{N}^d$.*
2. *$\mathcal{F}(x^\alpha T)(\xi) = i^{|\alpha|}\partial_\xi^\alpha (\mathcal{F}T)(\xi)$ for all $\alpha \in \mathbb{N}^d$.*

Moreover, $\mathcal{F}$ on $\mathcal{S}'(\mathbb{R}^d)$ is an extension of $\mathcal{F}$ on $\mathcal{S}(\mathbb{R}^d)$ as defined previously.

Fourier integration theorem

Theorem 238. *We have the following.*

1. *The continuous linear map $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is bijective. Its inverse is $\mathcal{F}^{-1} = \overline{\mathcal{F}}$.*
2. *For all $\phi \in \mathcal{S}(\mathbb{R}^d)$, the Plancharel identity*

   $$\int_{\mathbb{R}^d} |\phi(x)|^2 \,\mathrm{d}x = \int_{\mathbb{R}^d} |\widehat{\phi}(\xi)|^2 \,\mathrm{d}\xi \tag{6.94}$$

   holds.
3. *For all $\phi \in \mathcal{S}(\mathbb{R}^d)$, we have*

   $$\int_{\mathbb{R}^d} \phi(x)\overline{\psi(x)} \,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{\phi}(\xi)\overline{\widehat{\psi}(\xi)} \,\mathrm{d}\xi. \tag{6.95}$$

4. *The continuous linear map $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is bijective and its inverse is $\mathcal{F}^{-1} = \overline{\mathcal{F}}$.*
5. *The map $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ can be extended by continuity into a bijective isometry of $L^2(\mathbb{R}^d)$ whose inverse is $\overline{\mathcal{F}}$.*

Remark 239. *For $f \in L^2(\mathbb{R}^d)$, we can write*

$$\widehat{\phi}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) e^{-i\langle x,\xi\rangle} \,\mathrm{d}x \tag{6.96}$$

but if $f \in L^2(\mathbb{R}^d) \setminus L^1(\mathbb{R}^d)$ then this integral is NOT *defined as a Lebesgue integral. Instead, it is constructed via extension from $\mathcal{S}(\mathbb{R}^d)$ and is seen as the semi-convergent integral*

$$\widehat{\phi}(\xi) = \frac{1}{(2\pi)^{d/2}} \lim_{R\to\infty} \int_{|x|\le R} f(x) e^{-i\langle x,\xi\rangle}\,\mathrm{d}x. \tag{6.97}$$

One should thus be particularly careful if one needs to apply Lebesgue's Theorem to the Fourier transform of an $L^2(\mathbb{R}^d)$ function.

We conclude with some properties of the Fourier transform.

Proposition 240. *The following hold.*

1. *Let $S, T \in \mathcal{S}'(\mathbb{R}^d)$ with adapted support for convolution. Then $\widehat{S*T}(\xi) = (2\pi)^{d/2}\widehat{S}\widehat{T}$.*
2. *Let T be compactly supported. Then $\widehat{T}(\xi) = \frac{1}{(2\pi)^{d/2}}\left\langle T \,\middle|\, e^{i\langle\cdot,\xi\rangle}\right\rangle$ for all $\xi \in \mathbb{R}^d$.*
3. *$\widehat{\delta}_0 = \frac{1}{(2\pi)^{d/2}}$ and $\widehat{1} = (2\pi)^{d/2}\delta_0$.*
4. *Let τ_h be translation by the vector h. Then for $T \in \mathcal{S}'(\mathbb{R}^d)$, $\tau_h T \in \mathcal{S}'(\mathbb{R}^d)$ for all $h \in \mathbb{R}^d$. We have $\mathcal{F}(\tau_h T)(\xi) = e^{-i\langle h,\xi\rangle}\widehat{T}(\xi)$*
5. *Let H_λ be dilation by $\lambda > 0$. Then for $T \in \mathcal{S}'(\mathbb{R}^d), H_\lambda T \in \mathcal{S}'(\mathbb{R}^d)$ for all $\lambda > 0$. We have $\mathcal{F}(H_\lambda T)(\xi) = \frac{1}{\lambda^d}(H_{1/\lambda}\widehat{T})(\xi)$.*
6. *The un-normalized Gaussian density is a fixed point of the Fourier transform: $\mathcal{F}(e^{-|x|^2/2})(\xi) = e^{-|\xi|^2/2}$.*

These may be used to fid elementary solutions of the heat and wave equations.

Chapter 7

Sobolev spaces

Definition 241. *Let M be an open set of $\mathbb{R}^d$. Define*

$$H^1(M) = \left\{ f \in L^2(M) : \forall i \in \{1, .., d\}, \partial_{x_i} f \in L^2(M) \right\} \tag{7.1}$$

and equip it with the inner product

$$\langle \cdot, \cdot \rangle_{H^1(M)} = \langle \cdot, \cdot \rangle_{L^2(M)} + \sum_{i=1}^{d} \langle \partial_{x_i} \cdot, \partial_{x_i} \cdot \rangle_{L^2(M)} \tag{7.2}$$

$$\|\cdot\|^2_{H^1(M)} = \|\cdot\|^2_{L^2(M)} + \sum_{i=1}^{d} \|\partial_{x_i} \cdot\|^2_{L^2(M)} \tag{7.3}$$

Then we have that $H^1(M)$ is a separable Hilbert space.

Recall that $L^2(M)$ is a separable Hilbert space. In the above, we are using that $\partial_{x_i} f \in L^2(M)$, and hence the square of the distributional derivative makes sense. Note also that $H^1(M)$ is not closed with respect to the $L^2(M)$ topology.

Before proceeding we first recall some ideas from calculus. If we take $f \in L^1((a, b))$ and define

$$F(x) = \int_a^x f(y) \, \mathrm{d}y \tag{7.4}$$

then $F(x)$ is absolutely continuous, so for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all N, and all collections $\{(a_i, b_i)\}_{i=1}^N$ with $a_1 < b_1 < a_2 < b_2 < .. < a_N < b_N$ with $\sum_{i=1}^N (b_i - a_i) < \delta$, we have $\sum_{i=1}^N |F(b_i) - F(a_i)| < \varepsilon$. This implies but is not equivalent to uniform continuity. Furthermore, F is a.e. differentiable, so

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} \tag{7.5}$$

exists for a.e. $x \in (a, b)$. Moreover, $F'(x) = f(x)$, a.e. $x \in (a, b)$. Conversely, if F is absolutely continuous, then there exists a unique $f \in L^1((a, b))$ such that

$$F(x) = \int_a^x f(y)\,\mathrm{d}y + c. \tag{7.6}$$

The regularity requirements are important: there exist counterexamples like the devil's staircase that are continuous but not absolutely continuous for which the result fails to hold.

We now proceed to prove a simple Sobolev embedding theorem.

Proposition 242. *Take $D = (a, b)$ to be a finite interval. We have $H^1((a, b)) \hookrightarrow C^0([a, b])$, i.e. that $H^1((a, b)) \subseteq C^0([a, b])$, and the canonical one-to-one mapping*

$$(H^1((a,b)), \|\cdot\|_{H^1(M)}) \to (C^0([a,b]), \|\cdot\|_\infty) \qquad u \mapsto u \tag{7.7}$$

is continuous. Hence, there exists c depending only on $b - a$ such that for all $u \in H^1((a, b))$, we have

$$\|u\|_\infty \leq c\|u\|_{H^1(M)} \tag{7.8}$$

which we call a Sobolev embedding.

Proof. Take $u \in H^1((a, b))$, so that $u \in L^2((a, b))$ and the distributional derivative u' satisfies $u' \in L^2((a, b))$. We have

$$\|u\|_{H^1} = \|u\|_{L^2} + \|u'\|_{L^2} < \infty. \tag{7.9}$$

We also have $u' \in L^2((a, b)) \subseteq L^1((a, b))$. Define

$$v(x) = \int_a^x u'(y)\,\mathrm{d}y \tag{7.10}$$

which is well-defined because $u' \in L^1((a,b))$. We have $v \in C^0([a,b])$. We have $\|v\|_\infty \leq c\|u'\|_{L^2}$. We now need to show that v', when viewed as a distributional derivative, is equal to u. Take $\phi \in D((a,b)) = C_c^\infty$ to be an arbitrary test function. Then v' is defined by

$$\begin{aligned}
\langle v' \mid \phi \rangle = -\langle v \mid \phi' \rangle &= -\int_a^b v(x)\phi'(x)\,\mathrm{d}x && (7.11)\\
&= -\int_a^b \int_a^x u'(y)\,\mathrm{d}y\phi'(x)\,\mathrm{d}x && (7.12)\\
&= -\int_a^b u'(y)\int_y^b \phi'(x)\,\mathrm{d}x\,\mathrm{d}y && (7.13)\\
&= -\int_a^b u'(y)\Big[\phi(b) - \phi(y)\Big]\,\mathrm{d}y && (7.14)\\
&= \int_a^b u'(y)\phi(y)\,\mathrm{d}y && (7.15)\\
&= \langle u' \mid \phi \rangle && (7.16)
\end{aligned}$$

where we have used that $\phi \in C_c^\infty$ so we can apply the fundamental theorem of calculus to it directly, and its value on the boundary is zero. This implies that $v' \in L^2((a,b))$ and $v' = u'$, a.e. on (a,b). Hence, the distributional derivative $(v-u)' = 0$, and hence there exists a $c \in \mathbb{R}$ such that $v - u = c$ by previous result. This holds even though $u \in H^1((a,b))$, and $v, c \in C^0([a,b])$. But

$$u(x) = \int_a^x u'(y)\,\mathrm{d}y + c \tag{7.17}$$

so $c = u(a)$, and thus

$$u(x) = u(a) + \int_a^x u'(y)\,\mathrm{d}y. \tag{7.18}$$

This gives the necessary injection. We have

$$\begin{aligned}
\|u\|_\infty &\leq |u(a)| + \left|\int_a^x u'(y)\,\mathrm{d}y\right| && (7.19)\\
&\leq |u(a)| + \sqrt{b-a}\|u'\|_{L^2} && (7.20)
\end{aligned}$$

but also

$$|u(a)| \leq |u(x)| - \int_a^b u'(y)\,\mathrm{d}y \tag{7.21}$$

so

$$|u(a)|(b-a) \leq \int_a^b |u(x)|\,\mathrm{d}x + (b-a)^{3/2}\|u'\|_{L^2} \tag{7.22}$$

$$\leq \sqrt{b-a}\left[\int_a^b |u(x)|^2\,\mathrm{d}x\right]^{1/2} + (b-a)^{3/2}\|u'\|_{L^2} \tag{7.23}$$

hence

$$|u(a)| \leq \frac{1}{\sqrt{b-a}}\|u\|_{L^2} + \sqrt{b-a}\|u'\|_{L^2} \leq c_1\|u\|_{H^1}. \tag{7.24}$$

Combining the formulas, we get

$$\|u\|_\infty \leq c_2\|u\|_{H^1} \tag{7.25}$$

which gives the continuous embedding of H^1 into C^0. ■

Remark 243. *Note that the above is* ONLY TRUE IN DIMENSION ONE – IT IS FALSE IN HIGHER DIMENSIONS. *If $d \geq 2$, then for M and open set in $\mathbb{R}^d$, we have*

$$H^1(M) \not\subseteq C^0(\overline{M}). \tag{7.26}$$

The counterexample in $d = 2$ is

$$u(x) = \left(\ln\frac{1}{|x|}\right)^k \tag{7.27}$$

with $k < 1/2$. We can show that $u \in H^1(B(0,R))$ for $R < 1$, but $u \notin C^0(\overline{B}(0,R))$ which can be checked by computing derivatives.

In particular, the Navier-Stokes equations in 3D are only known to have local solutions, i.e. those that exist for a finite length of time. If we instead consider weak solutions in the distributional sense, then they don't necessarily satisfy an energy identity arising from physical principles – only an inequality. The resolution of these issues is a long-standing open question in mathematics.

7.1. Traces

Consider the equation

$$-\Delta u + u = f \tag{7.28}$$

with f given. Take $d = 1$, $M = (a, b)$. Note that the equation is linear, so if we solve the homogeneous form

$$-u'' + u = 0 \tag{7.29}$$

we can add it to any solution $Ae^x + Be^{-x}$ with $A, B \in \mathbb{R}$ of the original equation to obtain another solution. We are going to instead fix the solution on the boundaries, to obtain

$$-u'' + u = 0 \qquad u(a) = A \qquad u(b) = B \tag{7.30}$$

for which we have a unique solution.

On the other hand, in the L^2 setting, the measure of the boundary ∂M is zero, so we can change the value of u arbitrarily on the boundary. We are going to show that if u is in H^1, then u has a well-defined boundary on the domain. We begin with some results.

Theorem 244. *$D(\mathbb{R}^d)$ is dense in $H^1(\mathbb{R}^d)$.*

Proof. We introduce a mollifier sequence and truncate. Taking $u \in H^1(\mathbb{R}^d)$, we show

$$\rho_n * u \xrightarrow{n\to\infty} u \tag{7.31}$$

with $\rho_n * u \in C^\infty(\mathbb{R}^d)$ and $u \in H^1(\mathbb{R}^d)$. This follows similarly to previous arguments. ■

Let M be an open set of $\mathbb{R}^d$. Let

$$D(M) = \{\phi \in D(\mathbb{R}^d) : \operatorname{supp}\phi \subseteq M\}. \tag{7.32}$$

Note that because M is open, there is always a gap between $\operatorname{supp}\phi$ and the boundary outside M. Next, define

$$D(\overline{M}) = \{\,\phi|_{\overline{M}} : \phi \in D(\mathbb{R}^d)\} \tag{7.33}$$

where $\cdot|_{\overline{M}}$ is the restriction to $\overline{M}$. Here, there is no gap on the boundary – it is not even necessarily zero on the boundary.

Extension

Lemma 245. *Take $x \in \mathbb{R}^d$, and write $x = (x_1, .., x_d)$. Let*

$$\mathbb{R}^d_+ = \{x \in \mathbb{R}^d : x_d > 0\}. \tag{7.34}$$

There exists an extension operator $\mathcal{E} : H^1(\mathbb{R}^d_+) \to H^1(\mathbb{R}^d)$ satisfying the following properties.

1. *$\mathcal{E}$ is linear.*
2. *For all $u \in H^1(\mathbb{R}^d_+)$, $\mathcal{E}u|_{\mathbb{R}^d_+} = u$. Note that the restriction operator is obviously continuous.*
3. *There exists a constant c such that $\|\mathcal{E}u\|_{H^1(\mathbb{R}^d)} = c\|u\|_{H^1(\mathbb{R}^d_+)}$.*

Proof. We only sketch the argument. Define $\mathcal{E}u$

$$\mathcal{E}u(x_{-d}, x_d) = \begin{cases} u(x_{-d}, x_d) & x_d > 0 \\ u(x_{-d}, -x_d) & x_d < 0. \end{cases} \tag{7.35}$$

The norm is clearly controlled, and the extension is still in H^1. The derivative may no longer be smooth, but this is no issue for H^1. ■

Theorem 246. *$D(\overline{\mathbb{R}^d_+})$ is dense in $H^1(\mathbb{R}^d_+)$.*

Proof. We only sketch the argument. Take $u \in H^1(\mathbb{R}^d_+)$. Consider $\mathcal{E}u \in H^1(\mathbb{R}^d)$. Use the result that $D(\mathbb{R}^d)$ is dense in $H^1(\mathbb{R}^d)$ to construct $\phi \in D(\mathbb{R}^d)$ such that $\|\phi - \mathcal{E}u\|_{H^1(\mathbb{R}^d)} < \varepsilon$. We restrict the function to $\mathbb{R}^d_+$ which gives a function in $H^1(\mathbb{R}^d_+)$ which is within ε of u. ■

Note that it is important to take $D(\overline{\mathbb{R}^d_+})$ rather than $D(\mathbb{R}^d_+)$ which vanish on the boundary.

If we take $\phi \in D(\overline{\mathbb{R}^d_+})$, then the restriction $\phi|_{\partial M}$ is well-defined. We can show

$$\left\| \phi|_{\partial M} \right\|_{L^1(\partial M)} \leq c\|\phi\|_{H^1(\mathbb{R}^d_+)}. \tag{7.36}$$

Extension

Theorem 247. *Let E be a normed vector space. Let G be a Banach space. Let $F \subseteq E$ be a dense subspace of E. Let $T : F \to G$ be a linear* *map which is continuous for $\|\cdot\|_E$ and $\|\cdot\|_G$. Then there exists a unique continuous linear map $\tilde{T} : E \to G$ such that $\tilde{T}\Big|_F = T$.*

This is proven in the obvious way by constructing a Cauchy sequence in F which is dense in E, considering the image of that sequence in G which since T is continuous and G is Banach is also Cauchy and converges, and assigning to each point in E its limit of that sequence. For this to work, we need to prove that the limit does not depend on the choice of sequence, and that the resulting operator is linear and continuous.

Theorem 248. *The map*

$$D(\overline{\mathbb{R}^d_+}) \xrightarrow{\gamma} L^2(\partial\mathbb{R}^d_+) \qquad \phi \overset{\gamma}{\mapsto} \phi|_{\partial\mathbb{R}^d_+} \tag{7.37}$$

can be extended uniquely to a map

$$H(\mathbb{R}^d_+) \xrightarrow{\gamma} L^2(\partial\mathbb{R}^d_+) \tag{7.38}$$

which is continuous in the sense that there exists a $c > 0$ such that

$$\|\gamma u\|_{L^2(\partial\mathbb{R}^d_+)} \leq c\|u\|_{H^1(\mathbb{R}^d_+)}. \tag{7.39}$$

Note that we have $\gamma(H^1(\mathbb{R}^d_+)) \subset L^2(\partial\mathbb{R}^d_+)$.

Not all boundary conditions are reasonable. We need two.

Definition 249. *Let M be a connected open set.* Boundary

1. *∂M is* SMOOTH *if it is Lipschitz and piecewise C^1.*
2. *∂M is* REGULAR *if M is on one side of the boundary.*

In particular, smoothness means that we can define a measure on ∂M induced by the Lebesgue measure of $\mathbb{R}^d$, along with a normal vector a.e. on ∂M – this explicitly excludes domains with cusps or fractal structures. Regularity excludes domains where a vector is tangent to two different parts of the same domain.

To extend the previous results, we replace the domain with a piecewise rectangular approximation, and maps each rectangle to a subset of the half-plane considered previously. Then, we use the preimage to map the reflected functions back to the original domain. Next, using partitions of unity – a standard trick in differential geometry – we can make all results proven to $\mathbb{R}^d_+$ to M, provided the ∂M is smooth and regular. This gives the following results.

Theorem 250. *$D(\overline{M})$ is dense in $H^1(M)$.*

Theorem 251. *There exists a continuous linear map $\mathcal{E} : H^1(M) \to H^1(\mathbb{R}^d)$ such that the following hold.*

1. *Extension: $\mathcal{E}u|_M = u$.*
2. *Continuity: there exists a c such that $\|\mathcal{E}u\|_{H^1(\mathbb{R}^d)} \leq c\|u\|_{H^1(M)}$.*

Theorem 252. *The linear map*

$$D(\overline{M}) \xrightarrow{\gamma} L^2(\partial M) \qquad\qquad \phi \overset{\gamma}{\mapsto} \phi|_{\partial M} \tag{7.40}$$

is continuous from $(D(\overline{M}), \|\cdot\|_{H^1(M)})$ into $(L^2(\partial M), \|\cdot\|_{L^2(\partial M)}$ and extends uniquely into a continuous linear map

$$\gamma : H^1(M) \to L^2(\partial M) \tag{7.41}$$

called the TRACE MAP.

As before, the density result is not true if we replace $D(\overline{M})$ with $D(M)$. Similarly, $\mathrm{Im}(\gamma) \subset L^2(\partial M)$ is a strict subset.

Recall that for $u, v \in C^1(\overline{M})$, Green's formula is

$$\int_M u \frac{\partial v}{\partial x_k} \,\mathrm{d}x = -\int \frac{\partial u}{\partial x_k} v \,\mathrm{d}x + \int_{\partial M} u(x)v(x)\nu_k \,\mathrm{d}S(x) \tag{7.42}$$

where ν_k is the unit normal pointing outwards, and S is the Lebesgue measure on the boundary. Using the trace map, we can show the following.

THEOREM 253. *Let $u, v \in H^1(M)$, $k \in \{1, .., d\}$. Then*

$$\int_M u \frac{\partial v}{\partial x_k}\,\mathrm{d}x = -\int \frac{\partial u}{\partial x_k} v\,\mathrm{d}x + \int_{\partial M} \gamma_u(x)\gamma_v(x)\nu_k\,\mathrm{d}S(x) \tag{7.43}$$

where γ_u and γ_v are the trace map applied to u and v.

Proof. We sketch the argument. The result follows by density of $D(\overline{M})$ into $H^1(M)$. The formula is true for $u, v \in D(\overline{M})$, so we pass to the limit by introducing two approximating sequences u_n and v_n. ■

We next sketch a consequence of Green's formula. Take $\phi \in H^1(M)$. Take a vector field $A \in H^1(M)^d$, i.e. $(A_1(x), .., A_d(x)$ with $A_k \in H^1(M)$ for all k. Define

$$\operatorname{div} A = \sum_{k=1}^{d} \frac{\partial A_k}{\partial x_k} \qquad \nabla\phi = \left(\frac{\partial \phi}{\partial x_1}, .., \frac{\partial \phi}{\partial x_d}\right) \tag{7.44}$$

so $\operatorname{div} A \in L^2(M)$ and $(\nabla\phi)_k \in L^2(M)$ for all $k = 1, .., d$. We have

$$\int_M \phi(x)\operatorname{div} A(x)\,\mathrm{d}x = -\int_M \nabla\phi \cdot A\,\mathrm{d}dx + \int_{\partial M} \phi(A\nu)\,\mathrm{d}S(x) \tag{7.45}$$

$$= -\sum_{k=1}^{d} \int_M \frac{\partial \phi}{\partial x_k} A_k\,\mathrm{d}x + \sum_{k=1}^{d} \int_{\partial M} \phi(A_k\nu_k)\,\mathrm{d}S(x). \tag{7.46}$$

where ν is the outward-pointing normal vector. Note that we have omitted explicitly writing the trace map γ, with respect to which the integral is understood to have been defined implicitly.

7.2. THE SPACE $H_0^1(M)$

DEFINITION 254. *Take M an open subset of $\mathbb{R}^d$. Define*

$$H_0^1(M) = \overline{D(M)}^{H^1(M)} \tag{7.47}$$

where $(\cdot)^{H^1(M)}$ is closure with respect to the H^1 norm.

Recall that for $D \neq \mathbb{R}^d$, we have $D(\overline{M})$ is dense in M. Hence, we can expect that $H_0^1(M) \subset H^1(M)$. By density and continuity of the trace map γ, we have that

$$H_0^1(M) \subseteq \ker\gamma = \{\phi \in H^1(M) : \gamma(\phi) = 0\} \tag{7.48}$$

and we can show that this is actually an equality.

Theorem 255. *We have $H_0^1(M) = \ker\gamma$, where*

$$H^1(M) \to L^2(\partial M) \qquad\qquad f \mapsto \gamma f \tag{7.49}$$

is the trace map. Moreover, let $u \in H_0^1(M)$, and let $\overline{u}$ be the extension of u by zero outside M, i.e.

$$\overline{u} = \begin{cases} u & \text{a.e. } x \in M \\ 0 & \text{a.e. } x \in \mathbb{R}^d \setminus D. \end{cases} \tag{7.50}$$

Then $\overline{u} \in H^1(\mathbb{R}^d)$.

Proof. We sketch the argument. We show this for $\mathbb{R}^d_+$ and then by partitions show it for any sufficiently well-behaved M. Write $\overline{u} = \lim_{p\to\infty} u * \rho_p$ for a sequence of mollifiers ρ_p which is a $C^\infty(\mathbb{R}^d)$ function. The main difficulty is that $u * \rho_p$ is not necessarily zero on the boundary – but ρ_p has compact support, so there is a shrinking interval outside the boundary which is nonzero. To bypass this, we translate u and then smooth it out, to guarantee that support is inside $\mathbb{R}^d_+$. This gives the necessary sequence and taking limits completes the argument. ∎

We need a definition.

H^1 seminorm

Definition 256. *Define the seminorm*

$$|u|_{H^1(M)} = \|\nabla u\|_{L^2(M)^d} = \sum_{k=1}^{d} \int_M \left|\frac{\partial u}{\partial x_k}\right|^2 \mathrm{d}x. \tag{7.51}$$

Call $|\cdot|_{H^1(M)}$ the $H^1(M)$ seminorm.

Next, we show a fundamental result with application to many fields.

THEOREM 257. *Let M be a bounded open set of $\mathbb{R}^d$. Then there exists a constant $C > 0$ such that for any $u \in H_0^1(M)$ – the zero boundary is important – we have* Poincaré inequality

$$\sum_{k=1}^{d} \int_M \left| \frac{\partial u}{\partial x_k} \right|^2 \mathrm{d}x = |u|^2_{H^1(M)} \geq C\|u\|^2_{L^2(M)} = C \int_M |u|^2 \,\mathrm{d}x \tag{7.52}$$

Proof. We sketch the argument – write $u(x)$ as the integral of the derivatives, and use the Cauchy-Schwarz inequality. ■

Note that the constant function c does not belong in $H_0^1(M)$. Hence, the Poincaré inequality can't possibly hold for all functions in $H^1(M)$.

Note also that the Poincaré inequality follows if the domain is bounded with respect to only one choice of direction, i.e. squeezed on two sides by lines but possibly infinite everywhere else. Formally, this occurs if there exists an $\omega \in \mathbb{R}^d$, $|\omega| = 1$, and there exists $c > 0$, for which $M \subseteq \{x \in \mathbb{R}^d : |\langle x, \omega \rangle| \leq c\}$, because there exists a direction in which u can be integrated from $-\infty$ to x.

Note finally that the Poincaré also follows if we replace $H_0^1(M)$ with the function space $H^1_{\Gamma_0}(M)$ whose values are known on some part Γ_0 of the boundary with nonzero measure. Note that the trace map γ_{Γ_0} restricted to Γ_0 is again a continuous linear map, and $H^1_{\Gamma_0}(M) = \ker \gamma_{\Gamma_0}$ which is a closed subspace of $H^1(M)$.

This inequality has a number of consequences. We have thus shown that on $H_0^1(M)$, the $H^1(M)$ seminorm and $H^1(M)$ norm are equivalent, and hence the $H^1(M)$ seminorm is actually a norm. We can see this because for all $u \in H_0^1(M)$, there exists constants c_1, c_2 such that

$$c_1\|u\|_{H^1(M)} \leq |u|_{H^1(M)} \leq c_2 \leq \|u\|_{H^1(M)}. \tag{7.53}$$

The right-hand inequality is obvious because $\|\cdot\|_{H^1(M)} = |\cdot|_{H^1(M)} + \|\cdot\|_{L^2(M)}$. The left-hand inequality can be shown by noting

$$\|u\|^2_{H^1(M)} = |u|_{H^1(M)} + \|u\|_{L^2(M)} \leq |u|^2_{H^1(M)} + C^{-1}|u|^2_{H^1(M)} = (1+C^{-1})|u|_{H^1(M)}. \tag{7.54}$$

The optimal constant in the Poincaré inequality only depends on M. Knowing the Poincaré constant gives information about geometric properties of M. It is also related to the eigenvalues of the Laplacian on M – indeed it is possible to reconstruct information about M using all the eigenvalues of the Laplacian. This makes intuitive sense: it's akin to hitting a wall and noting that the sound it gives off – which is directly related to the eigenvalues of the Laplacian – is also related to whether or not the wall is solid or hollow, thick or thin, its shape, and other properties.

The Poincaré inequality is an example of a *functional inequality* – it tells us about the structure of the function class $H_0^1(M)$. Another famous example is given by the Heisenberg uncertainty principle – position is defined as the integral of the wave function, whereas momentum is defined as the integral of the Fourier transform of the wave function – a particular functional inequality involving the Fourier transform gives rise to the principle. Yet another example is given in probability: the Poincaré constant tells us something about how fast certain stochastic processes converge to their stationary distributions. This shows that the Poincaré inequality is a crossroads of sorts between different branches of mathematics.

7.3. The space $H^m(M)$

Definition 258. *Let M be an open domain of $\mathbb{R}^d$. Recall that $|\alpha|$ is the length of a multi-index. Let*

$$H^m(M) = \{u \in L^2(M) : \forall \alpha \in \mathbb{N}^d, |\alpha| \leq m, \partial^\alpha u \in L^2(M)\} \tag{7.55}$$

equipped with

$$\|u\|_{H^m(M)} = \sqrt{\sum_{\substack{\alpha \in \mathbb{N}^d \\ \alpha \leq m}} \|\partial^\alpha u\|_{L^2(M)}} \quad \langle u, v \rangle_{H^m(M)} = \sum_{\substack{\alpha \in \mathbb{N}^d \\ \alpha \leq m}} \langle \partial^\alpha u, \partial^\alpha v \rangle_{L^2(M)}. \tag{7.56}$$

7.4. Nonlinear transformations of distributional derivatives

Recall that in ordinary calculus the chain rule is $(g \circ f)' = (g' \circ f)f'$ where $\circ$ is function composition. We'd like to check that this still works.

Theorem 259. *Let $G : \mathbb{R} \to \mathbb{R}$ such that $G \in C^1(\mathbb{R})$. Write $\circ$ for composition of functions. Assume the following.*

1. *$G \in C^1(\mathbb{R})$.*
2. *$G(0) = 0$.*
3. *There exists $M > 0$ such that $|G'(t)| \leq M$ for all $t \in \mathbb{R}$.*

Then $G \circ u \in H^1(M)$ and for all $i \in \{1, .., d\}$, $\partial_{x_i}(G \circ u) = (G' \circ u)\partial_{x_i} u$.

Proof. We sketch the argument. Using density of $D(\overline{M})$ in $H^1(M)$, we show the result is bounded in $H^1(M)$, and extend by continuity to the full space. ∎

The conditions ensure that the function's slopes are between two lines intersecting at the origin. The first condition ensures differentiability of G. The second condition is needed to guarantee integrability at ∞. The third condition ensures that if u is in $L^2(M)$ then $G \circ u$ is also in $L^2(M)$ – by Hölder's inequality, since $\partial_{x_i} u \in L^2(M)$, so we have no choice but to take G' in $L^\infty(M)$.

7.5. Sobolev embeddings and compactness

We first recall some results about compactness in Banach spaces.

Result 260. *Let E be a finite-dimensional space over $\mathbb{R}$ (or $\mathbb{C}$). Let $K \subseteq E$. We say that K is* relatively compact *if $\overline{K}$ is compact. Then K is compact iff K is closed and bounded, and K is compact iff K is bounded.* Bolzano-Weierstrass Theorem for Banach spaces

Reisz

Theorem 261. *Let E be a normed vector space. Consider $\overline{B}(0,1) = \{x \in E : \|x\| \leq 1\}$. Then $\overline{B}(0,1)$ is compact iff E is finite-dimensional.*

Example 262. *Consider the special case of separable Hilbert space H. Let $\{e_n\}_{n\geq 1}$ be a complete orthonormal basis. If $\overline{B}(0,1)$ in H is compact, then we can find a convergent subsequence $\{e_{n_k}\}_{k\geq 1}$. This sequence is Cauchy, so $\|e_{n_k} - e_{n_l}\| \xrightarrow{k,l\to\infty} 0$. But*

$$\|e_{n_k} - e_{n_l}\| = \|e_{n_k}\|^2 + \|e_{n_l}\|^2 - 2\langle e_{n_k}, e_{n_l}\rangle = \|e_{n_k}\|^2 + \|e_{n_l}\|^2 = 2 \quad (7.57)$$

so the sequence cannot possibly converge. Thus we cannot extract a convergent subsequence, so $\overline{B}(0,1)$ is not compact.

Definition 263. *A set $A \subseteq C^1(\overline{B}(0,R))$ is bounded iff there exists $c > 0$ such that for*

$$\|f\|_{1,\infty} = \max_{i=1,..,d}\{\|f\|_\infty, \|\partial x_i f\|_\infty\} = \max_{i=1,..,d}\left\{\max_{x\in\overline{B}(0,R)} |f(x)|, \max_{x\in\overline{B}(0,R)} |\partial_{x_i} f(x)|\right\} \quad (7.58)$$

we have

$$\|f\|_{1,\infty} \leq c \quad (7.59)$$

for all $f \in A$.

Ascoli-Arzela

Theorem 264. *Any bounded set A of $C^1(\overline{B}(0,R))$ (or any compact set) is relatively compact in $C^0(\overline{B}(0,R))$ iff, letting $(f_n)_{n\geq 1}$ be a sequence of $C^1(\overline{B}(0,R))$ such that there exists $c > 0$ and $\|f_n\|_{1,\infty} \leq c$ for all $n \geq 1$, there exists a subsequence $(f_{n_k})_{k\geq 1}$ and $f \in C^0(\overline{B}(0,R))$ such that $\|f_{n_k} - f\|_{1,\infty} \xrightarrow{k\to\infty} 0$.*

This tells us that compactness in function space is related to regularity: we can pass to the subsequence provided the functions in question do not begin oscillating more and more as the sequence increases. This means we need bounds on derivatives to ensure compactness.

It's also crucial that we consider functions defined on a compact set – if we are not on a compact set, consider the sequence of mean-n Gaussian densities. This converges pointwise to zero, but doesn't converge in sup-norm.

If we consider the canonical one-to-one mapping

$$C^1(\overline{B}(0,R)) \xrightarrow{i} C^0(\overline{B}(0,R)) \qquad u \overset{i}{\mapsto} u \tag{7.60}$$

then Ascoli-Arzela says i is a continuous linear map that bounded sets into relatively compact sets. Such a map is called a compact map. We have the following definition.

Definition 265. *Let E, F be Banach spaces. Consider $T \in \mathcal{L}(E,F)$. Then T is compact iff T maps $\overline{B}(0,1) \in E$ into a relatively compact set of F, i.e. if $T(\overline{B}(0,1))$ is compact.*

We can show that compact maps are limits of maps whose range is finite-dimensional. We can also show the following

Theorem 266. *Let M be a bounded open set in $\mathbb{R}^d$. Then the canonical one-to-one map* Rellich-Kondrachov

$$H^1(M) \xrightarrow{i} L^2(M) \qquad u \overset{i}{\mapsto} u \tag{7.61}$$

is compact.

This means – and indeed is equivalent to – the statement that if $(f_n)_{n\geq 1}$ is a sequence with $f_n \in H^1(M)$ for all $n \geq 1$ and there exists $c > 0$ such that $\|f_n\|_{H^1(M)} \leq c$ for all $n \geq 1$, then there exists a subsequence $(f_{n_k})_{k\geq 1}$ and $f \in L^2(M)$ such that $\|f_{n_k} - f\|_{L^2(M)} \xrightarrow{k\to\infty} 0$.

The Rellich-Kondrachov Theorem can be used to prove Poincaré-like inequalities by {MISSED}.

Theorem 267. *Let $d \geq$. Let C_0^0 be the space of continuous functions tending to zero at infinity. Recall that $\hookrightarrow$ denotes a continuous one-to-one embedding.*

1. *Let $s > d/2$. Then* $(H^s(\mathbb{R}^d), \|\cdot\|_{H^s(\mathbb{R}^d)}) \hookrightarrow (C_0^0(\mathbb{R}^d), \|\cdot\|_\infty)$.
2. *Let $s < d/2$. Then* $(H^s(\mathbb{R}^d), \|\cdot\|_{H^s(\mathbb{R}^d)}) \hookrightarrow (L^p(\mathbb{R}^d), \|\cdot\|_{L^p(\mathbb{R}^d)})$ *for all p such that $2 \leq p \leq p^*$ with $p^* = \frac{2d}{d-2s}$.*
3. *Let $s = d/2$. Then* $(H^s(\mathbb{R}^d), \|\cdot\|_{H^s(\mathbb{R}^d)}) \hookrightarrow (L^p(\mathbb{R}^d), \|\cdot\|_{L^p(\mathbb{R}^d)})$, *for all p such that $2 \leq p < \infty$.*

Exercise 268. *Let $u \in H^1(M)$. Show that there exists $c > 0$ such that for all $\phi \in D$, we have*

$$\left| \int_M u \partial_{x_k} \phi(x)\, \mathrm{d}x \right| \leq c \|\phi\|_{L^2(M)}. \tag{7.62}$$

Proof. Using Green's formula, write

$$\int_M u \partial_{x_k} \phi(x)\, \mathrm{d}x = -\int_M \partial_{x_k} u \phi(x)\, \mathrm{d}x + \int_{\partial M} u \phi \nu_k \, \mathrm{d}S(x) = -\int_M \partial_{x_k} u \phi(x)\, \mathrm{d}x \tag{7.63}$$

where the boundary term disappears because ϕ has compact support. But then

$$\left| \int_M u \partial_{x_k} \phi(x)\, \mathrm{d}x \right| \leq \|\partial_{x_k} u\|_{L^2(M)} \|\phi\|_{L^2(M)} \tag{7.64}$$

by Cauchy-Schwarz so take $c = \max_{k=1,..,d} \|\partial_{x_k} u\|_{L^2(M)}$. ■

Exercise 269. *Let $u \in L^2(M)$ and suppose there exists a c such that for all $\phi \in D(M)$ we have*

$$\left| \int_M u \partial_{x_k} \phi(x)\, \mathrm{d}x \right| \leq c \|\phi\|_{L^2(M)}. \tag{7.65}$$

Show the following.

1. *Show that the distribution $\partial_{x_k} u$ satisfies* $|\langle \partial_{x_k} u \mid \phi \rangle_{D^*,D}| \leq c \|\phi\|_{L^2(M)}$
2. *Show that $\partial_{x_k} u$ is also a continuous linear form on D for the $L^2(M)$ topology.*
3. *Show that $\partial_{x_k} u$ can be extended uniquely to a continuous linear form on $L^2(M)$.*

4. *Show that* $\partial_{x_k} u \in L^2(M)$ *for all* $k \in \{1, .., d\}$.

Conclude that $u \in H^1(M)$.

Proof. By definition

$$\langle \partial_{x_k} u \mid \phi \rangle_{D^*,D} = -\langle u \mid \partial_{x_k}\phi \rangle_{D^*,D} = -\int_M u \partial_{x_k}\phi(x)\,\mathrm{d}x \tag{7.66}$$

and the first part follows by assumption. We have that $D(M) \subseteq L^2(M)$. Note that the $D(M)$ topology is stronger than the $L^2(M)$ topology. This implies that a continuous linear form in $D(M)$ is not necessarily continuous in $L^2(M)$. Since $\partial_{x_k} u$ is a linear map, to check continuity, it suffices to check boundedness of the map $\phi \mapsto \langle \partial_{x_k} u \mid \phi \rangle$. But by the previous part we immediately have

$$\frac{|\langle \partial_{x_k} u \mid \phi \rangle|}{\|\phi\|_{L^2(M)}} \leq c \tag{7.67}$$

so the linear form is bounded and therefore continuous, giving the second part. Next, we use the Extension Theorem. Taking $F = L^2(M)$, $E = D(M)$, $G = \mathbb{R}$, we get the desired unique extension – call it $\tilde{\partial}_{x_k} u \in L^2(M)'$ But by the Reisz Representation Theorem there exists a $v_k \in L^2(M)$ such that for $w \in L^2(M)$

$$\langle \tilde{\partial}_{x_k} u \mid w \rangle_{L^2(M),L^2(M)} = \int_M v_k(x) w(x)\,\mathrm{d}x. \tag{7.68}$$

Let $\phi \in D(M)$. We have

$$\langle \partial_{x_k} u \mid \phi \rangle_{D^*(M),D(M)} = \langle \tilde{\partial}_{x_k} u \mid \phi \rangle_{D^*(M),D(M)} = \langle \tilde{\partial}_{x_k} u \mid \phi \rangle_{L^2(M),L^2(M)} = \int_M v_k(x) w(x)\,\mathrm{d}x \tag{7.69}$$

and hence the distribution $\partial_{x_k} u$ can be identified with the $L^1_{\mathrm{loc}}(M)$ function v_k. But since $v_k \in L^2(M)$, $\partial_{x_k} u \in L^2(M)$. ■

Exercise 270. *For* $u : \mathbb{R}^d \to \mathbb{R}$, $h \in \mathbb{R}^d$, *define* $\mathcal{D}_h u$ *by* Finite difference trick

$$\mathcal{D}_h u(x) = u(x+h) - u(x). \tag{7.70}$$

Show $u \in H^1(\mathbb{R}^d)$ *iff* $u \in L^2(\mathbb{R}^d)$ *and there exists* $c > 0$ *such that for all* $h \in \mathbb{R}^d$, $\|\mathcal{D}_h u\|_{L^2(\mathbb{R}^d)} \leq c|h|$.

Proof. The previous exercise gives the reverse direction, so we only show the forward direction. Consider first $u \in D(\mathbb{R}^d)$. By Taylor's Theorem – in its most precise form – we have

$$u(x+h) - u(x) = \int_0^1 (\nabla u)(x+th) \cdot h \, \mathrm{d}t. \tag{7.71}$$

This statement can itself by proven by writing

$$v(t) = u(x+th) \tag{7.72}$$

for $t \in \mathbb{R}$, for which we have

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \sum_{k=1}^{d} \partial_{x_k} u(x+th) h_k = (\nabla u)(x+th) h \tag{7.73}$$

on which we can apply the Fundamental Theorem of Calculus to obtain

$$\int_0^1 \frac{\mathrm{d}v}{\mathrm{d}t} \, \mathrm{d}t = v(1) - v(0) = u(x+h) - u(x) \tag{7.74}$$

which gives Taylor's Theorem. We can then write,

$$\begin{aligned} |u(x+h) - u(x)| &= \left| \int_0^1 (\nabla u)(x+th) \cdot h \, \mathrm{d}t \right| && (7.75) \\ &\le \int_0^1 |(\nabla u)(x+th)| \, |h| \, \mathrm{d}t && (7.76) \\ &\le |h| \int_0^1 |(\nabla u)(x+th)| \, \mathrm{d}t && (7.77) \end{aligned}$$

using Cauchy-Schwarz on $\mathbb{R}^d$ and then

$$\begin{aligned} |u(x+h) - u(x)|^2 &\le |h|^2 \left[\int_0^1 |(\nabla u)(x+th)| \, \mathrm{d}t \right]^2 && (7.78) \\ &\le |h|^2 \int_0^1 |(\nabla u)(x+th)|^2 \, \mathrm{d}t && (7.79) \end{aligned}$$

using Cauchy-Schwarz on $L^2(\mathbb{R}^d)$. Taking integrals, we can write

$$
\begin{aligned}
\|\mathcal{D}_h u\|^2_{L^2(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} |u(x+h) - u(x)|^2 \,\mathrm{d}x && (7.80)\\
&\le |h|^2 \int_{\mathbb{R}^d} \int_0^1 |(\nabla u)(x+th)|^2 \,\mathrm{d}t\,\mathrm{d}x && (7.81)\\
&\le |h|^2 \int_0^1 \int_{\mathbb{R}^d} |(\nabla u)(x+th)|^2 \,\mathrm{d}x\,\mathrm{d}t && (7.82)\\
&\le |h|^2 \int_0^1 \|\nabla u\|^2_{L^2(\mathbb{R}^d)} \,\mathrm{d}t && (7.83)\\
&\le |h|^2 \|\nabla u\|^2_{L^2(\mathbb{R}^d)} \int_0^1 \mathrm{d}t && (7.84)\\
&\le |h|^2 \|\nabla u\|^2_{L^2(\mathbb{R}^d)} && (7.85)
\end{aligned}
$$

so taking square roots and setting $c = \|\nabla u\|_{L^2(\mathbb{R}^d)}$, by assumption finite since $u \in H^1(\mathbb{R}^d)$ gives the result for $D(\mathbb{R}^d)$. If we consider sequences in $D(\mathbb{R}^d)$, we can pass to the limit from $D(\mathbb{R}^d)$ to $H^1(\mathbb{R}^d)$, because $\|\mathcal{D}_h \cdot\|_{L^2(\mathbb{R}^d)}$ and $\|\nabla \cdot\|_{L^2(\mathbb{R}^d)}$ are continuous with respect to the $H^1(\mathbb{R}^d)$ topology. ■

Chapter 8

Elliptic partial differential equations

In this section we assume that M satisfies the following.

1. M is an open subset of $\mathbb{R}^d$.
2. M is bounded.
3. M is connected.
4. ∂M is smooth.
5. ∂M is regular.

Example 271. *Find $u : M \to \mathbb{R}$ such that* Dirichlet problem

$$-\Delta u + u = f \qquad x \in M \tag{8.1}$$
$$u = g \qquad x \in \partial M \tag{8.2}$$

with $f : M \to \mathbb{R}$ and $g : \partial M \to \mathbb{R}$ given.

Neumann problem

Example 272. *Find $u : M \to \mathbb{R}$ such that*

$$-\Delta u + u = f \qquad x \in M \tag{8.3}$$

$$\frac{\partial u}{\partial \nu} = \nabla u \cdot \nu = g \qquad x \in \partial M \tag{8.4}$$

where ν is the unit normal boundary, with $f : M \to \mathbb{R}$ and $g : \partial M \to \mathbb{R}$ given.

Note that in general both the equation and boundary conditions are needed to define the operator $(-\Delta(\cdot) + \mathcal{I}(\cdot))_{\mathrm{N}}$ in the previous example.

Other boundary conditions are also possible. We list another one below.

Example 273. *Find $u : M \to \mathbb{R}$ such that*

$$-\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}(a_{ij}(x)\frac{\partial u}{\partial x_j}) + a_0(x)u = f \qquad x \in M \tag{8.5}$$

$$u = g_0 \qquad x \in \Gamma_0 \tag{8.6}$$

$$\frac{\partial u}{\partial \nu_a} = \sum_{i,j=1}^{d} a_{ij}\frac{\partial u}{\partial x_j}\nu_i = g_1 \qquad x \in \Gamma_1 \tag{8.7}$$

where ν_i is the outwards unit normal and Γ_0, Γ_1 are measurable subsets of ∂M such that $\Gamma_0 \cap \Gamma_1 = \varnothing$ and $\Gamma_0 \cup \Gamma_1 = \partial M$. The term $\frac{\partial u}{\partial \nu_a}$ is called the co-normal derivative *of u with respect to a.*

8.1. Variational solutions of the Dirichlet and Neumann problems

Consider the problem

$$-\Delta u + u = f \qquad x \in M \tag{8.8}$$

$$\frac{\partial u}{\partial \nu} = \nabla u \cdot \nu = g \ \text{ or } \ u = g \qquad x \in \partial M \tag{8.9}$$

with $f \in L^2(M)$, $g \in \gamma(H^1(M)) = H^{1/2}(\partial M) \subseteq L^2(\partial M)$. Then there exists a $\tilde{g} \in H^1(M)$ – not necessarily unique – such that $\gamma\tilde{g} = \tilde{g}|_{\partial M} = g$.

Suppose $u \in H^1(M)$. Then $u = g$ on ∂M is equivalent to $u - \tilde{g} \in H_0^1(M)$ and $\gamma(u - \tilde{g}) = g - g = 0$.

Assume $u \in H^2(M)$ and $\tilde{g} \in H^2(M)$. Then the equation $-\Delta u + u = f$ is an equality among $L^2(M)$ functions – this will be called a *strong solution* if and only if the following hold.

1. $u - \tilde{g} \in H^2(M) \cap H_0^1(M)$.
2. $-\Delta u + u = f$ is true a.e. $x \in M$.

We note now that if u is a strong solution, then u is a solution to a variational formulation. Take an arbitrary test function $v \in H_0^1(M)$. Write

$$\int_M (-\Delta u + u - f) v \,\mathrm{d}x = \int_M \nabla u \cdot \nabla v \,\mathrm{d}x - \int_{\partial M} v (\nabla u \cdot \nu) \,\mathrm{d}S(x) \tag{8.10}$$

$$= \int_M \nabla u \cdot \nabla v \,\mathrm{d}x. \tag{8.11}$$

Remark 274. *Note that*

$$\int_M -\Delta u \; u \,\mathrm{d}x = \int_M |\nabla u|^2 \,\mathrm{d}x \geq 0 \tag{8.12}$$

which is why we write $-\Delta u$ to begin with.

We have thus shown that a strong solution u satisfies the following.

1. $u \cdot \tilde{g} \in H^2(M) \cap H_0^1(M)$.
2. $\int_M \nabla u \cdot \nabla v \,\mathrm{d}x + \int_M uv \,\mathrm{d}x = \int_M fv \,\mathrm{d}x$ for all $v \in H_0^1(M)$.

We can relax the $H^2(M)$ requirement to obtain the following problem.

1. $u \cdot \tilde{g} \in H_0^1(M)$.
2. $\int_M \nabla u \cdot \nabla v \,\mathrm{d}x + \int_M uv \,\mathrm{d}x = \int_M fv \,\mathrm{d}x$ for all $v \in H_0^1(M)$.

We call u satisfying this a *variational solution* of the Dirichlet problem.

Remark 275. *In the Dirichlet case, we impose boundary conditions by changing the function space over which the problem is defined. In the*

Neumann case, we impose them by changing the variational formulation to include an extra boundary term.

We can change variables for this problem to write

$$\tilde{u} = u - \tilde{g} \qquad \Longleftrightarrow \qquad u = \tilde{u} + \tilde{g} \tag{8.13}$$

for $\tilde{u} \in H_0^1(M)$ and hence the problem above is equivalent to the following problem.

1. $\tilde{u} \in H_0^1(M)$.
2. $\int_M \nabla\tilde{u} \cdot \nabla\tilde{v} \,\mathrm{d}x + \int_M \tilde{u}v \,\mathrm{d}x = \int_M fv \,\mathrm{d}x - \int_M \nabla\tilde{g}\nabla v \,\mathrm{d}x - \int_M \tilde{g}v \,\mathrm{d}x$ for all $v \in H_0^1(M)$.

Define

$$a(\tilde{u}, v) = \int_M \nabla\tilde{u} \cdot \nabla\tilde{v} \,\mathrm{d}x + \int_M \tilde{u}v \,\mathrm{d}x \quad b(v) = \int_M fv \,\mathrm{d}x - \int_M \nabla\tilde{g}\nabla v \,\mathrm{d}x - \int_M \tilde{g}v \,\mathrm{d}x. \tag{8.14}$$

Note that

$$a : H_0^1(M) \times H_0^1(M) \to \mathbb{R} \qquad b : H_0^1(M) \to \mathbb{R} \tag{8.15}$$

are bilinear and linear, res[ectively. And hence the variational problem above can be written as follows.

1. $\tilde{u} \in H$.
2. $a(\tilde{u}, v) = b(v)$ for all $v \in H$.

To get a unique solution, we need to assume the following.

1. b is continuous: there exists $c_b > 0$ such that $|b(v)| \leq c\|v\|_H$.
2. a is continuous: there exists $c_a > 0$ such that $|a(u, v)| \leq c\|u\|_H\|v\|_H$.
3. a is coercive: there exists $\alpha > 0$ such that $a(u, u) \geq \alpha\|u\|_H^2$.

Furthermore, the map

$$H' \to H \qquad b \mapsto u \tag{8.16}$$

is linear and continuous, so

$$\|u\|_H \leq \frac{1}{\alpha}\|b\|_{H'} = \frac{1}{\alpha} \sup_{\substack{v \in H \\ v \neq 0}} \frac{|b(v)|}{\|v\|_H} \tag{8.17}$$

by Lax-Milgram Theory.

The solution u does not depend on the choice of $\tilde{g}$ except through the corresponding choice of g: if we have $\tilde{g}_1, \tilde{g}_2$ such that $\gamma(\tilde{g}_1) = \gamma(\tilde{g}_2) = g$, and take $\tilde{u}_1, \tilde{u}_2$, then we can show $\tilde{u}_1 = \tilde{u}_2$. We can further show that the norm of u will only depend on the norm of $\tilde{g}$ through the norm of g.

If $a(u, v)$ is symmetric, then the solution u of the variational problem is also the unique solution of the following minimization problem:

$$u = \arg\min_{v \in H} J(v) = \frac{1}{2}a(v, v) - b(v). \tag{8.18}$$

Note that, in a sense that won't be made precise,

$$a(v + h, v + h) = a(v, v) + 2a(v, h) + O(\|h\|^2) \tag{8.19}$$

and so

$$J(v + h) - J(v) = a(v, h) - b(h) + O(\|h\|^2) \tag{8.20}$$

therefore a minimum should satisfy

$$a(u, h) - b(h) = 0 \tag{8.21}$$

for all appropriate h.

To apply Lax-Milgram to the previously-considered problem, we need to check the conditions. We have $H = H_0^1(M)$ and

$$a(\tilde{u}, v) = \int_M \nabla\tilde{u} \cdot \nabla\tilde{v} \,\mathrm{d}x + \int_M \tilde{u}v \,\mathrm{d}x \quad b(v) = \int_M fv \,\mathrm{d}x - \int_M \nabla\tilde{g}\nabla v \,\mathrm{d}x - \int_M \tilde{g}v \,\mathrm{d}x. \tag{8.22}$$

To check continuity of b write

$$|b(v)| \leq \|f\|_{L^2(M)}\|v\|_{L^2(M)} + \|\nabla\tilde{g}\|_{L^2(M)}\|\nabla v\|_{L^2(M)} + \|\tilde{g}\|_{L^2(M)}\|v\|_{L^2(M)} \tag{8.23}$$

$$\leq \left[\|f\|_{L^2(M)} + \|\tilde{g}\|_{H^1(M)}\right]\|v\|_{H_0^1(M)}. \tag{8.24}$$

To check continuity of a write

$$a(\tilde{u}, v) = \langle \tilde{u}, v \rangle_{H_0^1(M)} \leq \|\tilde{u}\|_{H_0^1(M)} \|v\|_{H_0^1(M)} \tag{8.25}$$

by Cauchy-Schwarz so a is continuous with $c_a = 1$. To check coercivity of a write

$$a(\tilde{u}, \tilde{u}) = \langle \tilde{u}, \tilde{u} \rangle_{H_0^1(M)} \leq \|\tilde{u}\|^2_{H_0^1(M)} \tag{8.26}$$

and so a is coercive with $\alpha = 1$. Thus the Lax-Milgram Theorem implies there exist a unique $\tilde{u} \in H_0^1(M)$ such that $a(\tilde{u}, v) = b(v)$ for all $v \in H_0^1(M)$. Furthermore, the map

$$L^2(M) \times H^1(M) \to H_0^1(M) \qquad (f, \tilde{g}) \mapsto u \tag{8.27}$$

is linear and continuous. Furthermore, since a is symmetric, it is the unique minimizer of

$$J(v) = \frac{1}{2}a(v,v) - b(v) = \frac{1}{2}\int_M \nabla\tilde{u}\cdot\nabla\tilde{v}\,\mathrm{d}x + \frac{1}{2}\int_M \tilde{u}v\,\mathrm{d}x - \int_M fv\,\mathrm{d}x + \int_M \nabla\tilde{g}\nabla v\,\mathrm{d}x + \int_M \tilde{g}v \tag{8.28}$$

which gives the energy principle for the given problem.

Note that under Dirichlet boundary conditions u is also the unique minimizer of

$$u = \operatorname*{arg\,min}_{\tilde{g}+H_0^1(M)} K(v) = \operatorname*{arg\,min}_{\tilde{g}+H_0^1(M)} \frac{1}{2}a(v,v) - \tilde{b}(v) = \operatorname*{arg\,min}_{\tilde{g}+H_0^1(M)} \frac{1}{2}a(v,v) - \int_M fv\,\mathrm{d}x. \tag{8.29}$$

Remark 276. *Here we again remark that the choice of function space on which the PDE is defined, along with choice of the bilinear form, determines the solution of our PDE. Here, we can see that the minimization problems for Neumann and Dirichlet boundary condition are different.*

Let's check that this extends the notion of a strong solution. Suppose now that u solves the variational problem. Take as a test function any $u \in D(M)$, which we can always take since $D(M) \subset H_0^1(M)$. We get

$$\langle \nabla u \mid \nabla v \rangle_{D^*(M),D(M)} + \langle u \mid v \rangle_{D^*(M),D(M)} = \langle f \mid v \rangle_{D^*(M),D(M)} \tag{8.30}$$

which by definition is the same as

$$-\langle \Delta u \mid v \rangle_{D^*(M),D(M)} + \langle u \mid v \rangle_{D^*(M),D(M)} = \langle f \mid v \rangle_{D^*(M),D(M)} \tag{8.31}$$

where Δ is taken in distributional sense. Hence, this is just

$$\langle -\Delta u + u - f \mid v \rangle_{D^*(M),D(M)} = 0 \qquad \forall v \in D(M). \tag{8.32}$$

with $-\Delta u, u, f \in L^2(M)$. We cannot conclude $u \in H^2(M)$ here, because $\sum_{i=1}^d \partial_{x_i} u \in L^2(M)$, but it's possible that $\partial_{x_i} u \notin L^2$ for some i. In fact, it's possible that $\partial_{x_i} u \notin L^1_{\text{loc}}$, though all singular parts must necessarily cancel when summed.

Remark 277. *In general the solution to a variational problem is not necessarily in $H^2(M)$, and hence is not a strong solution. We call it a* WEAK SOLUTION.

Hence for some problems, a strong solution may not exist. On the other hand, for other solution concepts such as an *ultra-weak solution*, we may not necessarily get uniqueness.

The variational formulation lends itself nicely to numerics: we can redefine the problem on a finite-dimensional subspace of $H^1(M)$, for which it becomes a linear system that can be solved numerically – this is the *finite element method*.

8.2. General elliptic problem

The general elliptic problem is handled analogously to the above.

8.3. Qualitative properties of variational solutions

Theorem 278. *Let M be a bounded domain such that the either ∂M is of class C^1 or M is convex. Let $f \in L^2(M)$ and u be the variational* Elliptic Regularity *solution to*

$$-\Delta u + u = f \qquad x \in M \tag{8.33}$$
$$u = 0 \qquad x \in \partial M \tag{8.34}$$

Then $u \in H^2(M)$ *and there exists* $c > 0$ *such that* $\|u\|_{H^2(M)} \leq c\|f\|_{L^2(M)}$.

If the assumptions fail, for instance if we are working with a heart-shaped domain, it's possible that some second derivative will diverge near a corner of the domain.

Stampacchia Maximum Principle

Theorem 279. *Let* M *be a bounded domain. Let* u *be a variational solution of the Dirichlet problem*

$$-\Delta u + u = f \qquad x \in M \tag{8.35}$$
$$u = g \qquad x \in \partial M \tag{8.36}$$

with $u \in \tilde{g} + H_0^1(M)$, $f \in L^2(M)$, $g \in \gamma(H^1(M))$, *and* $v \in H_0^1(M)$. *Then we have*

$$\min\left\{\inf_\Gamma g, \inf_D f\right\} \leq u \leq \max\left\{\sup_\Gamma g, \sup_D f\right\}. \tag{8.37}$$

Proof. Let $G \in C^1(\mathbb{R})$ such that

1. $|G'(s)| \leq M$.
2. G is increasing on $[0, \infty)$.
3. $G = 0$ on $(-\infty, 0]$.

Let $k = \max\{\sup_\Gamma g, \sup_M f\}$. Take $v = G(u - k)$, so $v \in H^1(M)$. Recall that for $J \in C^1(M)$ with $|J'(t)| \leq M$ and $J(0) = 0$, we have $f \in H^1(M)$ implies

$$J \circ f \in H^1(M) \qquad \nabla(J \circ f) = (J \circ f)\nabla f. \tag{8.38}$$

Apply this result with $J(u) = G(u - k) - G(-k)$. We now check that $v \in H_0^1(M)$. We have that on ∂M, $u = g$. We have $G(u - k) = G(g - k)$.

But $g \leq k$ by definition of k, so $g - k \leq 0$ and thus $G(g-k) = 0$ by construction. Hence $v \in H^1_0(M)$. We have $\nabla v = G'(u-k)\nabla u$. Thus

$$\int_M |\nabla u|^2 G'(u-k)\,\mathrm{d}x + \int_M uG(u-k)\,\mathrm{d}x = \int_M fG(u-k)\,\mathrm{d}x \tag{8.39}$$

hence subtracting $\int_M kG(u-k)\,\mathrm{d}x$ from both sides we get

$$\int_M |\nabla u|^2 G'(u-k)\,\mathrm{d}x + \int_M (u-k)G(u-k)\,\mathrm{d}x = \int_M (f-k)G(u-k)\,\mathrm{d}x. \tag{8.40}$$

The first term is non-negative. The third term on the right-hand side is negative. Thus

$$\int_M (u-k)G(u-k)\,\mathrm{d}x \leq 0 \tag{8.41}$$

Since $G(u-k) \geq 0$, we have $u - k \leq 0$, a.e. $x \in M$, and thus $u \leq k$. The other side of the inequality is proved similarly. ■

8.4. Spectral decomposition of elliptic operators

Some operators behave similarly to matrices – here, we consider compact operators. Let H be a separable Hilbert space. Let $T \in \mathcal{L}(H)$ be a continuous linear operator from H to itself. Recall that a linear operator is continuous iff it is bounded, and

$$\|T\|_{\mathcal{L}(H)} = \sup_{\substack{x\in H\\ x\neq 0}} \frac{\|Tx\|_H}{\|x\|_H} = \sup_{\substack{x\in H\\ \|x\|_H \leq 1}} \|Tx\|_H. \tag{8.42}$$

Recall also that T is compact iff $\overline{T(B(0,1))}$ is compact in H. Recall that in an infinite-dimensional space, $B(0,1)$ is not compact.

Definition 280. *Define* Space of compact operators

$$\mathcal{K}(H) = \{T \in \mathcal{L}(H) : T \text{ compact}\}. \tag{8.43}$$

We have that $\mathcal{K}(H)$ is a closed subspace of $\mathcal{L}(H)$ in the operator norm.

Hence if we have a sequence of operators $(T_n)_{n\geq 1}$ with $T_n \in \mathcal{K}(H)$ and there exists a $T \in \mathcal{L}(H)$ such that $\|T_n - T\|_{\mathcal{L}(H)} \xrightarrow{n\to\infty} 0$ then T is compact.

Finite rank operator

Definition 281. *We say that T is finite rank iff $\operatorname{Im} T = T(H)$ is finite-dimensional.*

Obviously, T finite rank is compact. We can thus use this to show T is compact by exhibiting a sequence of finite rank operators converging to T in operator norm. It turns out that, for a separable Hilbert space, the converse is also true.

Proposition 282. *T is compact if and only if there exists $(T_n)_{n\geq 1}$ such that $\|T_n - T\|_{\mathcal{L}(H)} \xrightarrow{n\to\infty} 0$, and T_n is finite rank for all n.*

For a general Banach space, the converse is false. This occurs because of a lack of a projection theorem, and was shown very recently.

Adjoint operator

Definition 283. *Take $T \in \mathcal{L}(H)$. Fix x. Consider the map*

$$H \to \mathbb{R} \qquad y \mapsto \langle x, Ty\rangle. \tag{8.44}$$

This is a continuous linear form because $|\langle x, Ty\rangle| \leq \|x\|_H \|Ty\|_H \leq \|T\|_{\mathcal{L}(H)}\|x\|_H\|y\|_H \leq c\|y\|_H$. By the Reisz Representation Theorem, there exists $\xi \in H$ such that

$$\langle x, Ty\rangle = \langle \xi, y\rangle. \tag{8.45}$$

The dependence of ξ on x is linear, so there exists a linear map

$$H \xrightarrow{T^*} H \qquad x \xmapsto{T^*} \xi. \tag{8.46}$$

Hence there exists a continuous linear operator T^* such that

$$\langle x, Ty\rangle = \langle T^* x, y\rangle. \tag{8.47}$$

Moreover

$$\|T^*x\|_H = \sup_{\substack{y\in H\\ y\neq 0}} \frac{|\langle T^*x, y\rangle|}{\|y\|_H} = \sup_{\substack{y\in H\\ y\neq 0}} \frac{|\langle x, Ty\rangle|}{\|y\|_H} \leq \sup_{\substack{y\in H\\ y\neq 0}} \frac{\|x\|_H\|T\|_{\mathcal{L}(H)}\|y\|_H}{\|y\|_H} = \|x\|_H\|T\|_{\mathcal{L}(H)} \tag{8.48}$$

hence $\|T^*\|_{\mathcal{L}} \leq \|T\|_{\mathcal{L}}$. *But,* $T^{**} = T$ *by definition, so* $\|T\|_{\mathcal{L}} \leq \|T^*\|_{\mathcal{L}}$ *and thus*

$$\|T^*\|_{\mathcal{L}} = \|T\|_{\mathcal{L}}. \tag{8.49}$$

We call T^* *the* ADJOINT OPERATOR *to* T.

PROPOSITION 284. T *is compact iff* T^* *is compact.*

DEFINITION 285. T *is* SELF-ADJOINT *if* $T^* = T$. Self-adjoint operator

A self-adjoint operator generalizes the concept of a symmetric matrix. Recall that a symmetric matrix is diagonalizable in an orthonormal basis, with real eigenvalues. The spectral theorem says that this is also true for compact self-adjoint operators.

REMARK 286. *Henceforth, let* T *be compact and self-adjoint. For non-self-adjoint operators, extensions are possible, but it's possible that the eigenvalues are complex, which we do not consider here.*

DEFINITION 287. *We say that* $\lambda \in \mathbb{R}$ *is an* EIGENVALUE *of* T *if there exists* $x \in H$, $x \neq 0$ *such that*

$$Tx = \lambda x \tag{8.50}$$

or equivalently $\ker(T - \lambda \operatorname{Id}) \neq \{0\}$. *We say that*

$$\operatorname{VP}(T) = \{\lambda \in \mathbb{R} : \lambda \textit{ is an eigenvalue of } T\}. \tag{8.51}$$

We say that $\lambda \in \sigma(T)$, *which we call the* SPECTRUM *of* T, *iff* $T - \lambda \operatorname{Id}$ *is not invertible. Note that*

$$\operatorname{VP}(T) \subseteq \sigma(T) \tag{8.52}$$

but it's possible, unlike in the finite-dimensional case, that $\operatorname{VP}(T) \neq \sigma(T)$. *We call* $\rho(T) = \mathbb{R} \setminus \sigma(T)$ *the* RESOLVENT *set. If* $\lambda \in \rho(T)$, *we can define the* RESOLVENT OPERATOR $R_\lambda(T) = (T - \lambda \operatorname{Id})^{-1}$, *which is continuous by the open mapping theorem.*

Recall that for a finite-dimensional space, the dimension of the kernel and the dimension of the image add up to the dimension of the full space.

Spectral decomposition of self-adjoint operators on a Hilbert space

Theorem 288. *Let H be a separable Hilbert space with $\dim H = \infty$. Let $T \in \mathcal{K}(H)$ be self-adjoint. Then*

1. $0 \in \sigma(T)$. *Thus T is not invertible.*
2. $\sigma(T) \setminus \{0\}$ *is a finite or countable set of eigenvalues of T. Thus* $\ker(T - \lambda \operatorname{Id}) \neq \{0\}$.
3. *Let $(\lambda_n)_{n\geq 1}$ be the sequence of eigenvalues. If $\sigma(T) \setminus \{0\}$ is countable, then $\lambda_n \xrightarrow{n\to\infty} 0$.*
4. *For all $n \geq 1$, $H_n = \ker(T - \lambda_n \operatorname{Id})$ is finite-dimensional. Moreover, $H_n \perp H_m$ for all $n \geq m$.*
5. *We have $H = \ker(T) \overset{\perp}{\oplus} \Big(\bigoplus_{n\geq 1}^{\perp} H_n\Big)$.*
6. *Take an orthonormal basis $\{e_{n,1}, .., e_{n,m_n}\}$ of H_n and a Hilbert basis $\{e_{0,1}, .., e_{0,j}, ..\}$ of $\ker(T)$. Then the union over all these bases is a Hilbert basis of of H, which can be ordered to give a sequence of eigenvectors of T.*

The reason that $0 \in \sigma(T)$ is because if we suppose the converse, then T is invertible, and since it is continuous and linear, we have that T^{-1} is invertible by the Open Mapping Theorem. If we take $\overline{B(0,1)} \in H$, then $\overline{T(\overline{B}(0,1))}$ is compact, so $T^{-1}(\overline{T(\overline{B}(0,1))})$ is compact. Obviously, $\overline{B}(0,1) \subseteq \overline{T(\overline{B}(0,1))}$ thus $\overline{B}(0,1)$ is compact, so H is finite-dimensional, which is a contradiction. Thus $0 \in \sigma(T)$ because T shrinks the unit ball into a compact set, so there is no way it could be invertible.

For $\lambda \neq 0$, $\ker(T - \lambda \operatorname{Id})$ is finite-dimensional for more-or-less the same reason. If we take $x \in \ker(T - \lambda \operatorname{Id})$, then $x = \frac{1}{\lambda} T x$. But if $\|x\| \leq 1$, then $x \in \frac{1}{\lambda} T(\overline{B}_{H_\lambda}(0,1)) \subseteq \frac{1}{\lambda} T(\overline{B}_H(0,1))$ is relatively compact, thus $\overline{B}_{H_\lambda}(0,1)$ is compact, and H_λ has to be finite-dimensional.

We now move to an example application.

Proposition 289. *Suppose M is bounded. There exists a Hilbert basis of $L^2(M)$, $(e_n)_{n\geq 1}$, and a sequence $(\lambda_n)_{n\geq 1}$ such that $\lambda_n \geq 0$, $\lambda_n \xrightarrow{n\to\infty} 0$, such that the following hold.*

1. *$e_n \in H_0^1(M)$.*
2. *e_n is a variational solution of $-\Delta e_n = \lambda_n e_n$ in M.*

Furthermore, if ∂M is C^1 on M or M is convex, then $e_n \in C^\infty(M) = \bigcap_{k\geq 0} H^k(M)$.

Proof. Consider the problem

$$\underbrace{\int_M \nabla u \cdot \nabla v \,\mathrm{d}x}_{a(u,v)} = \underbrace{\int_M fv \,\mathrm{d}x}_{b(v)} \qquad \forall v \in H_0^1(M) \qquad u \in H_0^1(M). \tag{8.53}$$

We check that a, b are continuous and coercive. Continuity is obvious. To check coercivity, write

$$a(u,u) = \int_M \|\nabla u\|^2_{L^2(M)} \geq \alpha \|u\|^2_{H^1(M)} = \alpha\Big[\|u\|^2_{L^2(M)} + \|\nabla u\|^2_{L^2(M)}\Big]. \tag{8.54}$$

For the ∇u term, using the Poincaré inequality with constant C, write

$$\|\nabla u\|^2_{L^2(M)} = \frac{1}{2}\|\nabla u\|^2_{L^2(M)} + \frac{1}{2}\|\nabla u\|^2_{L^2(M)} \geq \frac{1}{2}\|\nabla u\|^2_{L^2(M)} + \frac{C}{2}\|u\|^2_{L^2(M)} \geq \min\Big\{\frac{1}{2}, \frac{C}{2}\Big\}\|u\|^2_{H^1(M)} \tag{8.55}$$

and so a is coercive over $H_0^1(M)$. By the Lax-Milgram Theorem, there exists a unique solution $u \in H_0^1(M)$ and furthermore there exists $c > 0$ such that

$$\|u\|_{H^1(M)} \leq c\|f\|_{L^2(M)}. \tag{8.56}$$

Thus we can define a mapping

$$L^2(M) \xrightarrow{T} H^1(M) \qquad\qquad f \overset{T}{\mapsto} u. \tag{8.57}$$

Let $J : H^1(M) \to L^2(M)$ be a continuous embedding. Then J is a compact linear operator. Define

$$L^2(M) \xrightarrow{\tilde{T}} L^2(M) \qquad\qquad f \overset{\tilde{T}}{\mapsto} J(u). \tag{8.58}$$

The map $\tilde{T}$ is compact, so $J(T(\overline{B}_{L^2(M)}(0,1))) \subseteq J(T(\overline{B}_{H^1(M)}(0,C)))$. Thus we have a compact operator $\tilde{T} : L^2(M) \to L^2(M)$ which solves the variational formulation, which we henceforth identify with T. We now would like to show that this map is self-adjoint, so

$$\langle Tf, g\rangle_{L^2(M)} = \langle f, Tg\rangle_{L^2(M)} \tag{8.59}$$

where $Tf = u$, $Tg = w$, and

$$\int_M \nabla u \cdot \nabla v \,\mathrm{d}x = \int_M fv \,\mathrm{d}x \qquad \forall v \in H_0^1(M) \qquad u \in H_0^1(M) \tag{8.60}$$
$$\int_M \nabla w \cdot \nabla v \,\mathrm{d}x = \int_M gv \,\mathrm{d}x \qquad \forall v \in H_0^1(M) \qquad w \in H_0^1(M). \tag{8.61}$$

We can let $v = w$ on the top, and $v = u$ on the bottom. This gives

$$\int_M \nabla u \cdot \nabla w \,\mathrm{d}x = \int_M fw \,\mathrm{d}x \qquad \forall v \in H_0^1(M) \qquad u \in H_0^1(M) \tag{8.62}$$
$$\int_M \nabla w \cdot \nabla u \,\mathrm{d}x = \int_M gu \,\mathrm{d}x \qquad \forall v \in H_0^1(M) \qquad w \in H_0^1(M) \tag{8.63}$$

hence, since the bilinear a is symmetric,

$$\int_M fw \,\mathrm{d}x = \langle f, Tg\rangle = \langle Tf, g\rangle = \int_M gu \,\mathrm{d}x \tag{8.64}$$

which shows the map is self-adjoint. Finally, we can show that T is a non-negative operator in the sense $\langle Tf, f\rangle_{L^2(M)} \geq 0$. We can plug in u for v to obtain

$$0 \leq a(u,u) = b(v) = \langle u, f\rangle_{L^2(M)} = \langle Tf, f\rangle_{L^2(M)}. \tag{8.65}$$

If λ is an eigenvalue of T, then there exists a $u \neq 0$ such that $Tu = \lambda u$. Then $0 \leq \langle Tu, u\rangle_{L^2(M)} = \lambda\langle u, u\rangle_{L^2(M)}$ thus $\lambda \geq 0$. Hence $\ker(T) = \{0\}$. Suppose $f \in \ker(T)$. Then $u = Tf = 0$. Hence

$$0 = \int_M fv \,\mathrm{d}x \qquad\qquad \forall v \in H_0^1(M). \tag{8.66}$$

Taking $v \in D(M)$, we conclude that $f(x) = 0$ a.e. $x \in M$. By the Spectral Theorem, there exists a Hilbert basis $\{e_n\}_{n\geq 1}$ of $L^2(M)$ and a sequence $(\mu_n)_{n\geq 1}$ with $\mu_n \geq 0$ such that e_n is an eigenvector of T associated with eigenvalue $(\mu_n)_{n\geq 1}$. $(\mu_n)_{n\geq 1}$ can be ordered to be non-increasing with $\mu_n \xrightarrow{n\to\infty} 0$. The expression $Te_n = u = \lambda_n e_n$ so

$$\int_M \nabla e_n \cdot \nabla v \,\mathrm{d}x = \frac{1}{\mu_n} \int_M e_n v \,\mathrm{d}x \qquad \forall v \in H_0^1(M). \tag{8.67}$$

We conclude that there exists a Hilbert basis of $L^2(M)$ and a sequence $(\lambda_n)_{n\geq 1}$ with $\lambda_n > 0$, λ_n non-decreasing, $\lambda_n \xrightarrow{n\to\infty} \infty$, such that

$$\int_M \nabla e_n \cdot \nabla v \,\mathrm{d}x = \lambda_n \int_M e_n v \,\mathrm{d}x \qquad \forall v \in H_0^1(M). \tag{8.68}$$

By taking $v \in D(M)$ we get

$$-\langle \Delta e_n \mid v \rangle_{D(M)} = \lambda_n \langle e_n \mid v \rangle_{D(M)} \qquad \forall v \in H_0^1(M) \tag{8.69}$$

so $-\Delta e_n = \lambda_n e_n$ in $D(M)$. If ∂M is C^1 or M is convex, the Regularity Theorem tells us that $e_n \in H^2(M)$, and that $\Delta e_n \in H_0^1(M)$. Thus $-\Delta(\Delta e_n) = \lambda_n \Delta e_n$, so $e_n \in H^4(M)$. Thus $e_n \in \bigcap_{k\geq 1} H^{2k}(M) = C^\infty(M)$. ■

Chapter 9

Parabolic partial differential equations

Parabolic problems introduce a time variable, whose steady-state form is elliptic. We focus on bounded domains. We introduce a few examples.

Example 290. *Let $D \subset \mathbb{R}^d$ be a bounded domain, implicitly assumed connected, smooth, and regular so that Sobolev space theory holds. We seek to find a function $u : M \times [0, T] \to \mathbb{R}$ that satisfies the following.* Cauchy-Dirichlet problem

$$\frac{\partial u}{\partial t} - \Delta u = 0 \qquad u(x, 0) = u_0(x) \qquad \forall x \in M \tag{9.1}$$

$$(x, t) \in M \times [0, T] \qquad u(x, t) = 0 \qquad \forall (x, t) \in \partial M \times [0, T]. \tag{9.2}$$

For a given T, we call u a local solution*. If we have a unique solution for any T, then we can extend the solution to $[0, \infty)$, which gives a* global solution*. Call $Q = M \times [0, T]$ a* parabolic cylinder*, and define $\Sigma = \partial M \times [0, T]$*

Like for an elliptic problem, we need to introduce a variational formulation to be able to solve the problem. We introduce the setting now.

Definition 291. *Let $(X, \|\cdot\|)$ be a Banach space. Let $C^m([0,T], X)$ be the space of m times continuously differentiable functions $u : [0,T] \to X$. Such a function is differentiable at t_0 iff*

$$u_t = u(t_0) + u'(t_0)(t - t_0) + (t - t_0)\varepsilon(t) \tag{9.3}$$

with $\|\varepsilon(t)\|_X \xrightarrow{t \to t_0} 0$. Equivalently, for all sequences $t_n \xrightarrow{n\to\infty} t_0$, we have

$$\left\| \frac{u(t_n) - u(t_0)}{t_n - t_0} - u'(t_0) \right\|_X \xrightarrow{n\to\infty} 0. \tag{9.4}$$

We can define

$$\|u\|_{C^m([0,T],X)} = \max_{\substack{t\in[0,T]\\ 0\le l\le m}} \left\| \frac{\mathrm{d}^l u}{\mathrm{d}t^l}\,(t) \right\|_X. \tag{9.5}$$

We now extend Lebesgue spaces to the Bochner setting.

Definition 292. *Let X be a Banach space. Let $1 \le p \le \infty$. Let*

$$\|u\|_{L^p([0,T],X)} = \left(\int_0^T \|u(t)\|_X^p \,\mathrm{d}t \right)^{1/p} \qquad \|u\|_{L^\infty([0,T],X)} = \operatorname*{ess\,sup}_{t\in[0,T]} \|u(t)\|_X \tag{9.6}$$

and let

$$L^p([0,T],X) = \Big\{ u : [0,T] \to X \text{ measurable}, \|u\|_{L^p([0,T],X)} < \infty \Big\} \tag{9.7}$$

where integrability is taken in the Bochner sense, which is not defined here. This is a Banach space.

Definition 293. *Let X be a Hilbert space. Then $L^2([0,T],X)$ is a Hilbert space with inner product*

$$\langle u, v \rangle_{L^2([0,T],X)} = \int_0^T \langle u(t), v(t) \rangle_X \,\mathrm{d}t. \tag{9.8}$$

We also have $\|u\|_{L^2([0,T],X)} = \langle u, u \rangle_{L^2([0,T],X)}^{1/2}$.

In particular, we have that the map

$$[0,T] \to \mathbb{R} \qquad\qquad t \mapsto \langle u(t), v\rangle_X \tag{9.9}$$

is in $L^2([0,T])$ and

$$\|\langle u(\cdot), v\rangle_X\|^2_{L^2([0,T])} = \int_0^T |\langle u(t), v\rangle_X|^2 \,\mathrm{d}t \le \int_0^T \|u(t)\|_X^2 \|v\|_X^2 \,\mathrm{d}t \le \|v\|_X^2 \|u\|^2_{L^2([0,T],X)} \tag{9.10}$$

hence

$$\|\langle u(\cdot), v\rangle_X\|_{L^2([0,T])} \le \|v\|_X \|u\|_{L^2([0,T],X)}. \tag{9.11}$$

Definition 294. *Consider the Cauchy-Dirichlet problem. A function u is a* classical, smooth, *or* strong *solution iff*

$$u \in C^0([0,T], H^2(M) \cap H^1_0(M)) \cap C^1([0,T], L^2(M)). \tag{9.12}$$

The equation is valid almost everywhere in space and point-wise in time.

We can relax these regularity requirements. Take $v \in H^1_0(M)$ not dependent on t. Integrating, we claim that

$$\int_M \frac{\partial u}{\partial t}(t) v \,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t} \int_M u(t) v \,\mathrm{d}x. \tag{9.13}$$

We know $u \in C^1([0,T], L^2(M))$. This implies that $\int_M u(t)v\,\mathrm{d}x \in C^1([0,T])$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_M u(t) v \,\mathrm{d}x = \int_M u' v \,\mathrm{d}x \tag{9.14}$$

where u' is the derivative of $u : [0,T] \to L^2(M)$. To show this, we need to prove that u is continuous and differentiable. To show differentiability, we can take a sequence $t_n \xrightarrow{n\to\infty} t_0$, $t_n \neq t_0$, and need to show that

$$\left| \frac{\int_M u(t_n) v \,\mathrm{d}x - \int_M u(t_0) v \,\mathrm{d}x}{t_n - t_0} - \int_M u' v \,\mathrm{d}x \right| \xrightarrow{n\to\infty} 0. \tag{9.15}$$

We have that, since $u \in C^1([0,T], L^2(M))$ and $v \in L^2(M)$,

$$\left| \frac{\int_M u(t_n)v \,\mathrm{d}x - \int_M u(t_0)v \,\mathrm{d}x}{t_n - t_0} - \int_M u'v \,\mathrm{d}x \right| = \left| \int_M \frac{u(t_n)v - u(t_0)v}{t_n - t_0} - u'v \,\mathrm{d}x \right| \tag{9.16}$$

$$\leq \int_M \left| \frac{u(t_n)v - u(t_0)v}{t_n - t_0} - u'v \right| \mathrm{d}x \tag{9.17}$$

$$\leq \left\| \frac{u(t_n) - u(t_0)}{t_n - t_0} - u' \right\|_{L^2(M)} \|v\|_{L^2(M)} \tag{9.18}$$

and the claim follows from the definition of u' as the derivative of $u : [0,T] \to L^2(M)$ at t_0. The proof of continuity is similar.

If u is a strong solution then multiplying by v and integrating by parts gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_M uv \,\mathrm{d}x + \int_M \nabla u \cdot \nabla v \,\mathrm{d}x = 0 \qquad \forall v \in H_0^1(M) \tag{9.19}$$

which is the variational formulation.

As before, this formulation still makes sense under weaker regularity conditions.

Definition 295. *A function u is called a variational solution to the Cauchy-Dirichlet problem*

$$\frac{\partial u}{\partial t} - \Delta u = 0 \qquad u(x,0) = u_0(x) \qquad \forall x \in M \tag{9.20}$$

$$(x,t) \in M \times [0,T] \qquad u(x,t) = 0 \qquad \forall (x,t) \in \partial M \times [0,T]. \tag{9.21}$$

iff we have

$$u \in L^2([0,T], H_0^1(M)) \cap C^0([0,T], L^2(M)) \tag{9.22}$$

with

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_M uv \,\mathrm{d}x + \int_M \nabla u \cdot \nabla v \,\mathrm{d}x = 0 \quad \forall v \in H_0^1(M) \quad u(0) = u_0 \quad u_0 \in L^2(M). \tag{9.23}$$

Note that $\nabla u \in L^2([0,T], L^2(M))$, and $\nabla v \in L^2(M)$, hence $t \mapsto \int_M \nabla u(t) \cdot \nabla v \,\mathrm{d}x$ is in $L^2([0,T])$. Similarly, $u \in L^2([0,T], L^2(M))$, and $v \in L^2(M)$, hence $t \mapsto \int_M u(t) v \,\mathrm{d}x$ is in $L^2([0,T])$.

The equation is interpreted in the distributional sense in time, i.e. for all $\psi \in D([0,T])$ we have

$$\left\langle \frac{\mathrm{d}}{\mathrm{d}t} \int_M u(t) v \,\mathrm{d}x \;\middle|\; \psi(t) \right\rangle_{D^*([0,T]),D([0,T])} + \left\langle \int_M \nabla u \cdot \nabla v \,\mathrm{d}x \;\middle|\; \psi(t) \right\rangle_{D^*([0,T]),D([0,T])} = 0 \tag{9.24}$$

$$\implies \quad -\left\langle \int_M u(t) v \,\mathrm{d}x \;\middle|\; \frac{\mathrm{d}}{\mathrm{d}t}\, \psi(t) \right\rangle_{D^*([0,T]),D([0,T])} + \int_0^T \int_M \nabla u \cdot \nabla v \,\mathrm{d}x \; \psi(t) \,\mathrm{d}t = 0 \tag{9.25}$$

$$\implies \quad -\int_0^T \int_M u(t) v \,\frac{\mathrm{d}}{\mathrm{d}t}\, \psi(t) \,\mathrm{d}x \,\mathrm{d}t + \int_0^T \int_M \nabla u \cdot \nabla v \psi(t) \,\mathrm{d}x \,\mathrm{d}t = 0 \tag{9.26}$$

for all $\psi \in D([0,T]), v \in H_0^1(M)$. We can observe that $t \mapsto \int_M u(t) v \,\mathrm{d}x$ is in $H^1([0,T])$. Hence it's also possible to take $\psi \in H_0^1(M)$.

Similarly, suppose that u is a variational solution. Then we have

$$\left\langle \frac{\partial}{\partial t} u \;\middle|\; v(x)\psi(t) \right\rangle_{D^*(D\times[0,T]),D(D\times[0,T])} - \langle \Delta_x u \mid v(x)\psi(t) \rangle_{D^*(D\times[0,T]),D(D\times[0,T])} = 0 \tag{9.27}$$

since the space spanned by the functions of the form $v(x)\psi(t)$ with $v \in D(M)$ and $\psi \in D([0,T])$ is dense in $D(M \times [0,T])$. Hence the formulation is equivalent to

$$\left\langle \frac{\partial}{\partial t} u - \Delta_x u \;\middle|\; \phi \right\rangle_{D^*(D\times[0,T]),D(D\times[0,T])} = 0 \qquad \forall \phi \in D(D \times [0,T]). \tag{9.28}$$

The behavior of the Cauchy-Neumann problem is similar. A strong solution u is taken to be

$$u \in C^0([0,T], H^2(M)) \cap C^1([0,T], L^2(M)). \tag{9.29}$$

A variational solution, defined in the obvious way, is likewise taken to be

$$u \in L^2([0,T], H^1(M)) \cap C^0([0,T], L^2(M)). \tag{9.30}$$

9.1. Abstract variational framework

Let V, H be two separable Hilbert spaces. Typically, $V = H^1(M)$ or $V = H_0^1(M)$ depending on whether we are looking at the Dirichlet or Neumann problem, and $H = L^2(M)$. Assume $V \hookrightarrow H$ compactly, i.e. that the identity map between V and H is continuous and compact, and there exists a constant c such that $\|\cdot\|_H \leq c\|\cdot\|_V$. Assume V is dense in H. Define a continuous bilinear form

$$a : V \times V \to \mathbb{R}. \tag{9.31}$$

Consider the problem

1. Find $u \in L^2([0,T], V) \cap C^0([0,T], H)$ such that $\frac{\mathrm{d}}{\mathrm{d}t} \langle u(t), v\rangle_H + a(u(t), v) = 0$ for all $v \in V$ in $D^*([0,T])$ with $u(0) = u_0 \in H$ given.

For the Cauchy-Neumann, we take $H = L^2(M)$, $V = H^1(M)$, $a(u,v) = \int_M \nabla u \cdot \nabla v \,\mathrm{d}x$. For Cauchy-Dirichlet, we take the same but with $V = H_0^1(M)$. Assume the following.

1. There exists $\lambda \in \mathbb{R}$, $\alpha > 0$ such that $a(u,u) + \lambda\|u\|_H^2 \geq \alpha\|u\|_V^2$. for all $u \in V$.

2. The bilinear form a is symmetric.

For example, for the Cauchy-Neumann problem we get $\int_M |\nabla u|^2 + \int_M \|u\|_{L^2(M)}^2 \geq \|U\|_{H^1(M)}^2$ which gives the inequality with $\lambda = 1$. Define

$$a_\lambda(u,v) = a(u,v) + \lambda\langle u, v\rangle_H. \tag{9.32}$$

Then a_λ is a continuous coercive bilinear form on V. Take $f \in H$. Then

$$a_\lambda(u,v) = \langle f, v\rangle_H \qquad u \in V \qquad \forall v \in V \tag{9.33}$$

has a unique solution u with $\|u\|_V \leq c\|f\|_H$. The map $H \xrightarrow{T} V$, $f \overset{T}{\mapsto} u$ is continuous. On the other hand, $V \hookrightarrow H$ is compact. Thus, $T : H \to H$ is compact. Since a_λ is symmetric, T is self-adjoint. We also have $\langle Tu, u\rangle_H \geq 0$, so the eigenvalues of T are non-negative. We have $\ker T = \{0\}$. By the Spectral Theorem, there exists a Hilbert basis $\{e_n\}_{n\geq 1}$ of H and a sequence of eigenvalues of T with $\mu_n > 0$ and $\mu_n \searrow 0$. These satisfy

$$a_\lambda(\mu_n e_n, v) = \langle e_n, v\rangle_H \qquad \forall v \in V \tag{9.34}$$

thus $e_n \in V$ and

$$a(e_n, v) + \lambda \langle e_n, v \rangle_H = \frac{1}{\mu_n} \langle e_n, v \rangle_H \qquad \forall v \in V. \tag{9.35}$$

Thus $\{e_n\}_{n \geq 1}$ is a Hilbert basis of H, and there exists a sequence $(\lambda_n)_{n \geq 1}$ with $\lambda_n > -\lambda$ and $\lambda_n \nearrow \infty$ such that

$$a(e_n, v) = \lambda_n \langle e_n, v \rangle_H \qquad \forall v \in V \tag{9.36}$$

with $e_n \in V$. We aim to use this via the following lemma.

Lemma 296. *Let u be a solution of the variational problem introduced above. Then u is uniquely given by*

$$u(t) = \sum_{n=1}^{\infty} \langle u_0, e_n \rangle_H e^{-\lambda_n t} e_n. \tag{9.37}$$

Proof. Since $u \in C^0([0, T], H)$, we have $u(t) \in H$. Thus

$$u(t) = \sum_{n=1}^{\infty} u_n(t) e_n \tag{9.38}$$

but we know by definition that $u_n(t) = \langle u(t), e_n \rangle_H$. We can plug e_n in as a test function into the variational formulation

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle u(t), v \rangle_H + a(u(t), v) = 0 \qquad \forall v \in V \tag{9.39}$$

to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} u_n(t) + \lambda_n u_n(t) = 0 \tag{9.40}$$

in $D^*(M)$ which has the solution

$$u_n(t) = e^{-\lambda_n t} u_n(0) \tag{9.41}$$

which was shown in the earlier chapters. But $u_n(0) = \langle u_0, e_n \rangle_H$.

$\blacksquare$

Note that this does not imply that there exists a solution – only that if there is one, it must be expressible in the given way. We now seek to prove that there is a unique solution given by the formula.

We seek to reduce the problem to the case $\lambda = 0$. Introduce $w(t) = e^{-\lambda_t}(t)$ and substitute into the problem. We have $w(t) \in L^2([0,T]V) \cap C^0([0,T], H)$, so there is no change in regularity. This gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle w(t), v(t) \rangle_H + \Big[a(w(t), v) + \lambda \langle w(t), v \rangle_H \Big] = 0 \qquad \forall v \in V \tag{9.42}$$

where the terms inside the brackets are a continuous bilinear form on V. This bilinear form is coercive. We conclude that u is a solution to its variational problem if and only if w is a solution to its associated variational problem. We have thus reduced the problem to the case $\lambda = 0$, which is assumed henceforth.

We now show that the series in the lemma converges. Define the partial sums

$$S_m(t) = \sum_{n=1}^{m} u_{0,n} e^{-\lambda_n t} e_n. \tag{9.43}$$

We want to show that the S_n converge, and that their limit satisfies the variational formulation. Introduce the finite-dimensional space

$$P_m = \mathrm{span}\{e_1, .., e_m\}. \tag{9.44}$$

We claim S_m is a solution of the Galerkin variational formulation

$$S_m \in C^1([0,T], P_m) \quad \frac{\mathrm{d}}{\mathrm{d}t} \langle S_m(t), v \rangle_H + a(S_m(t), v) = 0 \quad \forall v \in P_m \text{ in } C^0([0,1]) \tag{9.45}$$

where since P_m is finite-dimensional we now consider the equation in the classical rather than distributional sense. To show this, take $v = e_\mu$ with $\mu \in \{1, .., m\}$. If the formula holds for all such μ, then we can conclude the claim. By orthogonality we have

$$\langle S_m(t), e_\mu \rangle_H = u_{0,\mu} e^{-\lambda_\mu t} \qquad \Longrightarrow \qquad \frac{\mathrm{d}}{\mathrm{d}t} \langle S_m(t), e_\mu \rangle_H = -\lambda_\mu u_{0,\mu} e^{-\lambda_\mu t} \tag{9.46}$$

$$a(e_n, v) = \lambda_n \langle e_n, v \rangle_H \qquad \Longrightarrow \qquad a(S_m(t), e_\mu) = \lambda_\mu u_{0\mu} e^{-\lambda_\mu t} \tag{9.47}$$

which cancel out. We seek to show that that $S_m(t)$ is a Cauchy sequence in $C^0([0,T],H)$, by showing that

$$\|S_{m+k} - S_m\|_{C^0([0,T],H)} \xrightarrow{m\to\infty} 0 \text{ uniformly in } k. \tag{9.48}$$

Write

$$\|S_{m+k} - S_m\|_{C^0([0,T],H)} = \left\|\sum_{p=1}^{k} \langle u_0, e_{m+p}\rangle e^{-\lambda_{m+p}t} e_{m+p}\right\|_{C^0([0,T],H)} \tag{9.49}$$

$$= \sup_{t\in[0,T]} \left\|\sum_{p=1}^{k} \langle u_0, e_{m+p}\rangle e^{-\lambda_{m+p}t} e_{m+p}\right\|_{H} \tag{9.50}$$

$$= \sup_{t\in[0,T]} \left[\sum_{p=1}^{k} |\langle u_0, e_{m+p}\rangle|^2 e^{-2\lambda_{m+p}t}\right]^{1/2} \tag{9.51}$$

$$\leq \sup_{t\in[0,T]} \left[\sum_{p=1}^{k} |\langle u_0, e_{m+p}\rangle|^2\right]^{1/2} \tag{9.52}$$

$$\leq \left[\sum_{p=1}^{k} |\langle u_0, e_{m+p}\rangle|^2\right]^{1/2} \tag{9.53}$$

$$\leq \left[\sum_{p=1}^{k} |u_{0,m+p}|^2\right]^{1/2} \tag{9.54}$$

by Pythagoras' Theorem and $\|e_{m+p}\|_H = 1$. We have used that $(\lambda_n)_{n\geq 1}$ is a positive increasing sequence, so the exponential term is less than 1. But

$$u_0 = \sum_{n=1}^{\infty} u_{0mn} e_n \quad \Longrightarrow \quad \sum_{n=1}^{\infty} |u_{0m}|^2 < \infty \quad \Longrightarrow \quad \sum_{p=1}^{\infty} |u_{0,m+p}|^2 \xrightarrow{m\to\infty} 0 \tag{9.55}$$

because the remainder sequence of a convergent sequence converges to zero. The value k is not present thus converges occurs uniformly in k, so S_m is Cauchy in $C^0([0,T],H)$. By replacing the supremum with an integral, we also immediately get convergence in $L^2([0,T],H)$. We'd like to obtain convergence in $L^2([0,T],V)$. Write

$$\|S_{m+k} - S_m\|^2_{L^2([0,T],V)} \le \int_0^T \|S_{m+k} - S_m\|^2_V \,\mathrm{d}t \tag{9.56}$$
$$\le \frac{1}{\alpha}\int_0^T a(S_{m+k} - S_m, S_{m+k} - S_m)\,\mathrm{d}t \tag{9.57}$$
$$\le \frac{1}{\alpha}\int_0^T \sum_{p=1}^{k}\sum_{q=1}^{k} a\Big(u_{0,m+p}e^{-\lambda_{m+p}t}e_{m+p}, u_{0,m+q}e^{-\lambda_{m+q}t}e_{m+q}\Big)\,\mathrm{d}t \tag{9.58}$$
$$\le \frac{1}{\alpha}\int_0^T \sum_{p=1}^{k}\sum_{q=1}^{k} u_{0,m+p}u_{0,m+q}e^{-\lambda_{m+p}t-\lambda_{m+q}t}a(e_{m+p}, e_{m+q})\,\mathrm{d}t \tag{9.59}$$
$$\le \frac{1}{\alpha}\int_0^T \sum_{p=1}^{k} |u_{0,m+p}|^2 e^{-2\lambda_{m+p}t}a(e_{m+p}, e_{m+p})\,\mathrm{d}t \tag{9.60}$$
$$\le \frac{1}{\alpha}\int_0^T \sum_{p=1}^{k} |u_{0,m+p}|^2 e^{-2\lambda_{m+p}t}\lambda_{m+p}\langle e_{m+p}, e_{m+p}\rangle_H\,\mathrm{d}t \tag{9.61}$$
$$\le \frac{1}{\alpha}\sum_{p=1}^{k} |u_{0,m+p}|^2 \int_0^T e^{-2\lambda_{m+p}t}\lambda_{m+p}\,\mathrm{d}t \tag{9.62}$$
$$\le \frac{1}{2\alpha}\sum_{p=1}^{k} |u_{0,m+p}|^2(1 - e^{-2\lambda_{m+p}T}) \tag{9.63}$$
$$\le \frac{1}{2\alpha}\sum_{p=1}^{k} |u_{0,m+p}|^2 \tag{9.64}$$

where we have used coercivity, and convergence follows using the steps of the previous argument. Thus S_m is Cauchy in $L^2([0,T],V)$.

We have shown that S_m is Cauchy in $C^0([0,T],H)$ and in $L^2([0,T],V)$. So there exist limits $u \in C^0([0,T],H)$ and $v \in L^2([0,T],V)$ of S_m. To conclude that u and v are the same function, we need to find a larger space within which both topologies embed continuously. Thus we'd like to show $C^0([0,T],H) \hookrightarrow L^2([0,T],H)$ and $L^2([0,T],V) \hookrightarrow L^2([0,T],H)$.

For $u \in C^0([0,T],H)$ we have

$$\|u\|_{L^2([0,T],H)} = \left[\int_0^T \|u(t)\|_H^2 \,\mathrm{d}t\right]^{1/2} \leq \sup_{t\in[0,T]} \|u(t)\|_H^2 \left[\int_0^T 1 \,\mathrm{d}t\right]^{1/2} = \sqrt{T}\|u\|_{C^0([0,T],H)} \tag{9.65}$$

which gives the continuous embedding.

For $u \in L^2([0,T],V)$ we have

$$\|u\|_{L^2([0,T],H)} \leq C\left[\int_0^T \|u\|_H^2 \,\mathrm{d}t\right]^{1/2} = C\|u\|_{L^2([0,T],V)} \tag{9.66}$$

by Sobolev embedding.

Since the $L^2([0,T],H)$ topology is Hausdorff and thus limits are unique, we conclude that

$$u = \sum_{n=1}^{\infty} u_{0,n} e^{-\lambda_n t} e_n \tag{9.67}$$

is a unique solution to the problem.

We now seek to show that for all $\psi \in D([0,T])$ and all $v \in V$ we have

$$-\int_0^T \langle u(t), v\rangle_H \,\frac{\mathrm{d}\psi}{\mathrm{d}t}\,(t)\,\mathrm{d}t + \int_0^T a(u(t),v)\psi(t)\,\mathrm{d}t = 0 \quad \forall \psi \in D([0,T]). \tag{9.68}$$

Take $\mu \in \mathbb{N}$, $v \in P_\mu = \operatorname{span}\{e_1, .., e_\mu\}$. Take $m \in \mathbb{N}$, $m \geq \mu$ so $P_\mu \subseteq P_m$ so $v \in P_m$. For all $m \geq \mu$ we have

$$-\int_0^T \langle S_m(t), v\rangle_H \,\frac{\mathrm{d}\psi}{\mathrm{d}t}\,(t)\,\mathrm{d}t + \int_0^T a(S_m(t),v)\psi(t)\,\mathrm{d}t = 0 \quad \forall \psi \in D([0,T]). \tag{9.69}$$

Taking $m \to \infty$, we have $S_m \xrightarrow{m\to\infty} u$ in $C^0([0,T],H)$. This implies that

$$\int_0^T \langle S_m(t), v\rangle_H \,\frac{\mathrm{d}\psi}{\mathrm{d}t}\,(t)\,\mathrm{d}t \xrightarrow{n\to\infty} \int_0^T \langle u(t), v\rangle_H \,\frac{\mathrm{d}\psi}{\mathrm{d}t}\,(t)\,\mathrm{d}t \tag{9.70}$$

through a straightforward application of inequalities. Showing convergence for the other integral proceeds similarly. We conclude

$$-\int_0^T \langle u(t), v\rangle_H \,\frac{\mathrm{d}\psi}{\mathrm{d}t}\,(t)\,\mathrm{d}t + \int_0^T a(u(t),v)\psi(t)\,\mathrm{d}t = 0 \quad \forall \psi \in D([0,T]) \quad \forall v \in P_\mu. \tag{9.71}$$

But μ is arbitrary, so

$$-\int_0^T \langle u(t), v\rangle_H \frac{\mathrm{d}\psi}{\mathrm{d}t}(t)\,\mathrm{d}t + \int_0^T a(u(t), v)\psi(t)\,\mathrm{d}t = 0 \quad \forall \psi \in D([0,T]) \quad \forall v \in \bigcup_{\mu=1}^{\infty} P_\mu. \tag{9.72}$$

We thus need to prove that $\bigcup_{\mu=1}^{\infty} P_\mu$ is dense in V. If this holds, then for any $v \in V$ we have a sequence $v_k \in \bigcup_{\mu=1}^{\infty} P_\mu$, so it suffices to prove that we can pass to the limit, which proceeds via similar argument as before – this gives the variational formulation we originally sought.

Note that $\bigcup_{\mu=1}^{\infty} P_\mu = \operatorname{span}\{e_1, ..\}$ where $\{e_n\}_{n\geq 1}$ is a Hilbert basis. Hence $\overline{\bigcup_{\mu=1}^{\infty} P_\mu} = H$. However, this doesn't immediately imply it is dense in V. By symmetry, the inner product $a(u,u)$ induces a norm $\|u\|_a$. We have

$$\alpha\|u\|_a \leq \|u\|_V \leq C\|u\|_a \tag{9.73}$$

by coercivity and continuity, so $\|u\|_a$ is equivalent to $\|u\|_V$. We claim that

$$g_n = \frac{e_n}{\sqrt{\lambda_n}} \tag{9.74}$$

is a Hilbert basis of $(V, a(\cdot,\cdot))$. Obviously $g_n \perp g_m$ for $n \neq m$. We proceed by contradiction to show that $\operatorname{span}\{e_n\}_{n\geq 1}$ is dense in V, so suppose otherwise, i.e. suppose that $\overline{\operatorname{span}\{e_n, ..\}} \not\subseteq V$. Thus there exists a $x \in \overline{\operatorname{span}\{e_n\}_{n\geq 1}}^V$ such that $x \in V$. Thus

$$y = x - \mathcal{P}_{\overline{\operatorname{span}\{e_n,..\}}} x \neq 0. \tag{9.75}$$

Hence $y \perp \overline{\operatorname{span}\{e_n\}_{n\geq 1}}^V$, so $y \perp \operatorname{span}\{e_n\}_{n\geq 1}$ and thus for all n we have

$$a(y, e_n) = 0 \quad \Longrightarrow \quad \lambda_n \langle e_n, y\rangle_H = 0 \quad \Longrightarrow \quad \langle e_n, y\rangle_H = 0 \quad \Longrightarrow \quad y = 0 \tag{9.76}$$

which is a contradiction. Hence $\bigcup_{\mu=1}^{\infty} P_\mu = \overline{\operatorname{span}\{e_n\}_{n\geq 1}}^V = V$. The result follows.