University of Cambridge Mathematical Tripos Part II

Analysis of Functions

Lecturer:
Péter Varjú
pv270@dpmms.cam.ac.uk

Notes by:
Robert Frederik Uy
rfu20@cam.ac.uk

Course schedule

Lebesgue integration theory

Review of integration: simple functions, monotone and dominated convergence; existence of Lebesgue measure; definition of L^p spaces and their completeness. The Lebesgue differentiation theorem. Egorov's theorem, Lusin's theorem. Mollification by convolution, continuity of translation and separability of L^p when $p \neq \infty$. [5]

Banach and Hilbert space analysis

Strong, weak and weak-* topologies; reflexive spaces. Review of the Riesz representation theorem for Hilbert spaces; the Radon–Nikodym theorem; the dual of L^p. Compactness: review of the Ascoli–Arzelà theorem; weak-* compactnesss of the unit ball for separable Banach spaces. The Riesz representation theorem for spaces of continuous functions. The Hahn–Banach theorem and its consequences: separation theorems; Mazur's theorem. [7]

Fourier analysis

Definition of Fourier transform in L^1; the Riemann–Lebesgue lemma. Fourier inversion theorem. Extension to L^2 by density and Plancherel's isometry. Duality between regularity in real variable and decay in Fourier variable. [3]

Generalized derivatives and function spaces

Definition of generalized derivatives and of the basic spaces in the theory of distributions: $\mathcal{D}/\mathcal{D}'$ and $\mathcal{S}/\mathcal{S}'$. The Fourier transform on $\mathcal{S}'$. Periodic distributions; Fourier series; the Poisson summation formula. Definition of the Sobolev spaces H^s in $\mathbb{R}^d$. Sobolev embedding. The Rellich–Kondrashov theorem. The trace theorem. [5]

Applications

Construction and regularity of solutions for elliptic PDEs with constant coefficients on $\mathbb{R}^n$. Construction and regularity of solutions for the Dirichlet problem of Laplace's equation. The spectral theorem for the Laplacian on a bounded domain. *The direct method of the Calculus of Variations.* [4]

Contents

1 Measure theory

1.1 Recap of elementary measure theory

We recall some basic notions from II Probability and Measure whilst establishing the notation for this course.

Let E be a set. A family $\mathcal{B} \subset \mathcal{P}(E)$ is a σ-algebra if $\varnothing \in \mathcal{B}$ and it is closed under countable unions and complements. A map $\mu\colon \mathcal{B} \to [0,\infty]$ is a measure if it is σ-additive, i.e., $\mu(\bigsqcup_n A_n) = \sum_n \mu(A_n)$ for a countable disjoint collection (A_n). A pair $(E,\mathcal{B})$ is called a measurable space, whereas a triple $(E,\mathcal{B},\mu)$ is called a measure space. If E is a topological space, we can (and, in this course, will) consider $\mathcal{B}$ to be the Borel σ-algebra.

A map $f\colon E \to \mathbb{C}$ is measurable if $f^{-1}(A) \in \mathcal{B}$ for all Borel $A \subset \mathbb{C}$. If $f\colon E \to [0,\infty]$ is measurable, then $\int f\,d\mu$ is well-defined (in $[0,\infty]$). We say that $f\colon E \to \mathbb{C}$ is integrable if it is measurable and $\int |f|\,d\mu < \infty$. We then denote by $\mathcal{L}^1(E,\mathcal{B},\mu)$ the set of all integrable functions $E \to \mathbb{C}$. Writing $f \sim g \iff f = g$ a.e., we then define $L^1(E,\mathcal{B},\mu) := \mathcal{L}^1(E,\mathcal{B},\mu)/\sim$. However, it is of course standard to refer to an equivalence class $[f] \in L^1(E)$ using a concrete function f from it.

Theorem 1.1 Dominated convergence theorem
Let $(E,\mathcal{B},\mu)$ be a measure space. Let $g, f, f_1, f_2 \in L^1(E,\mathcal{B},\mu)$. Suppose $f_n(x) \to f(x)$ and $|f_n(x)| \le g(x)$ for a.e. $x \in E$. Then

$$\int_E f_n\,d\mu \to \int_E f\,d\mu$$

Recall that a measure space $(E,\mathcal{B},\mu)$ is σ-finite if there exists $(A_n)_{n\in\mathbb{N}}$ such that $E = \bigcup_{n\in\mathbb{N}} A_n$ and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.

Theorem 1.2 Fubini's theorem
Let $(E,\mathcal{A},\mu)$ and $(F,\mathcal{B},\mu)$ be two σ-finite measure spaces. Let $f\colon E\times F \to \mathbb{C}$ be $\mathcal{A}\otimes\mathcal{B}$-measurable.
(i) *If the map $x \mapsto \int_F f(x,y)\,d\nu$ in $L^1(E,\mathcal{A},\mu)$, then $f \in L^1(E\times F)$.*
(ii) *If $f \in L^1(E\times F)$, then*

$$\int_{E\times F} f\,d\mu\otimes\nu = \int_E\int_F f(x,y)\,d\nu(y)\,d\mu(x) = \int_F\int_E f(x,y)\,d\mu(x)\,d\nu(y)$$

1.2 Signed and complex measures

Definition Complex and signed measures
Let $(E,\mathcal{B})$ be a measurable space. A set function $\mu\colon \mathcal{B} \to \mathbb{C}$ is a complex measure if it is σ-additive. We then say that $(E,\mathcal{B},\mu)$ is a complex measure space. If $\mu(\mathcal{B}) \subset \mathbb{R}$, we call μ a signed measure and $(E,\mathcal{B},\mu)$ a signed measure space.

Remark. Previously, we really have been considering *positive* measures, though we will always just refer to such maps as *measures*. Note that not every (positive) measure is a complex measure.

Given a complex measure μ, define the *real* and *imaginary parts* $\operatorname{Re}\mu$ and $\operatorname{Im}\mu$ as

$$\forall\, A \in \mathcal{B} \qquad \operatorname{Re}\mu(A) = \operatorname{Re}(\mu(A)), \qquad \operatorname{Im}\mu(A) = \operatorname{Im}(\mu(A))$$

It is easy to verify that these are signed measures.

Definition Positive and negative sets
Let $(E, \mathcal{B}, \mu)$ be a signed measure space. We say that A is a positive (resp. negative) set if $\mu(B) \geq 0$ (resp. $\mu(B) \leq 0$) for all $B \subset A$.

Theorem 1.3 Hahn decomposition
Let $(E, \mathcal{B}, \mu)$ be a signed measure space. Then there exists a positive set $P \in \mathcal{B}$ and a negative set $N \in \mathcal{B}$ such that $E = P \sqcup N$.

In proving the Hahn decomposition theorem, we will ultimately be defining P be a positive set of the largest possible measure. However, we first show the existence of non-trivial positive sets via the following lemma.

Lemma 1.4
For all $A \subset \mathcal{B}$, there is a positive set $D \subset A$ such that $\mu(D) \geq \mu(A)$.

Proof. If A is positive, then we can simply take $D = A$ and we are done. It remains to carefully consider the case A is not positive. Pick negative set $B_1 \subset A$ such that $\mu(B_1)$ is "as negative as possible" in the approximate sense that there is no $B \subset A$ and $k \in \mathbb{N}$ for which $\mu(B_1) > -1/k \geq \mu(B)$. Let $A_1 = A \backslash B_1$. Continue inductively to define $B_2, A_2, B_3, A_3, \cdots$ in such a way that $B_{j+1} \subset A_j$ and $\mu(B_{j+1}) > -1/k$ for some $k \in \mathbb{N}$ only if this is so for all subsets of A_i; we then take $A_{j+1} = A_j \backslash B_{j+1}$.

Now, take $D = \bigcap A_j$. Then $A = D \sqcup B_1 \sqcup B_2 \sqcup \cdots$, so $\mu(A) = \mu(D) + \sum_j \mu(B_j)$. Since $\mu(B_j) \geq 0$ for each j, we have $\mu(D) \geq \mu(A)$. It then remains to show that D is positive. Note that $\mu(B_j) \to 0$ by convergence of $\sum_j \mu(B_j)$. Fix $k \in \mathbb{N}$. Then $\mu(B_i) > -1/k$ for some i. Then $\mu(B) > -1/k$ for all $B \subset A_{i-1}$ and thus for all $B \subset D$. Hence, D must be positive. ■

Proof of Theorem 1.3. Note that a union of positive sets $A = A_1 \cup A_2 \cup \cdots$ is positive; indeed, for any $B \subset A$, we have $B = (B \cap A_1) \sqcup (B \cap A_2 \backslash A_1) \sqcup (B \cap A_3 \backslash (A_1 \cup A_2)) \sqcup \cdots$ which implies that $\mu(B) \geq 0$ by σ-additivity.

Let $s = \sup\{\mu(A) \colon A \subset E \text{ positive}\}$ and let (P_i) be a sequence of positive sets in $\mathcal{B}$ such that $\lim \mu(P_i) = s$. Then by above, we know that $P = \bigcup P_i$ is positive. Since $\mu(P) \geq \mu(P_i)$, we also have $\mu(P) = s$.

Now, suppose, on the contrary, that $N := E \backslash P$ is not a negative set. Then there exists $B \subset N$ such that $\mu(B) > 0$. By Lemma 1.4, there exists $D \subset B$ positive such that $\mu(D) \geq \mu(B)$. It then follows that $\mu(P \cup D) > \mu(P)$ with $P \cup D$ a positive set — a contradiction! ■

Remarks

- A key corollary of the Hahn decomposition theorem is the Jordan decomposition of signed measures.
- The decomposition $E = P \sqcup N$ is manifestly non-unique as we can, for instance, 'move' a negligible subset from P to N. However, this makes no difference in the Jordan decomposition of μ.

Corollary 1.5 Jordan decomposition of signed measures
Let $(E, \mathcal{B}, \mu)$ be a signed measure space, with $E = P \sqcup N$ for positive P and negative N. Then $\mu^+ := \mu|_P$ and $\mu^- := -\mu|_N$ are positive measures satisfying $\mu = \mu^+ - \mu^-$.

Proof. Trivial. ■

Now that we've proven the Hahn and Jordan decompositions, we can now extend Lebesgue integration to signed and complex measures.

Definition Integral with respect to signed measures
Let $(E, \mathcal{B}, \mu)$ be a signed measure space. A measurable function $f\colon E \to \mathbb{C}$ is integrable if it is integrable with respect to μ^+ and μ^-. In that case, we then define its integral to be

$$\int_E f\, d\mu := \int_E f\, d\mu^+ - \int_E f\, d\mu^-$$

Definition Integral with respect to complex measures
Let $(E, \mathcal{B}, \mu)$ be a complex measure space. A measurable function $f\colon E \to \mathbb{C}$ is integrable with respect to $\operatorname{Re}\mu$ *and* $\operatorname{Im}\mu$. *In that case, we then define its integral to be*

$$\int_E f\, d\mu := \int_E f\, d\operatorname{Re}\mu + i\int_E f\, d\operatorname{Im}\mu$$

Finally, we end this subsection with a brief discussion on the Banach space structure of the space of complex measures. Given a signed measure μ, we can define its *total variation measure* to be

$$|\mu| := \mu^+ + \mu^-$$

and its *total variation norm* to be

$$||\mu|| := |\mu|(E) = \mu^+(E) + \mu^-(E)$$

On the Example Sheet, you will extend these notions to complex measures. It can then be shown that the space of complex measures on a measurable space $(E, \mathcal{B})$ forms a Banach space with respect to the total variation norm.

1.3 Radon-Nikodym theorem

Motivation: In II Probability and Measure, you met some examples of continuous and discrete random variables. Are there probability distributions that do not belong to either category? How do we know if a random variable has a density?

Let us start off with some definitions. Let μ, ν be (positive) measures on $(E, \mathcal{B})$.

Definition Absolutely continuous measure
We say that ν is absolutely continuous wrt μ if, for all $A \in \mathcal{B}$, $\mu(A) = 0 \Longrightarrow \nu(A) = 0$. We then write $\nu \ll \mu$.

Definition Singular measure
We say that ν is singular wrt μ if there exists a decomposition $E = A \sqcup B$ such that $\nu(A) = 0$ and $\mu(B) = 0$. We then write $\nu \perp \mu$.

Remark. It is easy to see that $\nu \perp \mu \iff \mu \perp \nu$.

Definition Concentrated on a set
We say that μ is concentrated on a set $A \in \mathcal{B}$ if $\mu(E \backslash A) = 0$.

Remark. $\mu \perp \nu$ iff there exist disjoint sets $A, B \in \mathcal{B}$ such that μ is concentrated on A and ν is concentrated on B.

We now state the two main theorems of this subsection.

Theorem 1.6 Radon-Nikodym theorem
Let μ, ν be finite measures on $(E, \mathcal{B})$. Suppose that $\nu \ll \mu$. Then there exists $f \in L^1(E, \mathcal{B}, \mu)$ such that

$$\nu(A) = \int_A f \, d\mu$$

for all $A \in \mathcal{B}$. If g is another function that satisfies the conclusion of this theorem, then $f = g$ μ-a.e.

The function f in the theorem above is called the *Radon-Nikodym (RN) derivative*, often denoted by

$$\frac{d\nu}{d\mu}$$

Remark. We could extend this theorem by replace finite with σ-finite, but this comes at the expense of having $f \in L^1$.

Theorem 1.7 Lebesgue decomposition
Let μ, ν be finite measures. Then there exist unique measures ν_a, ν_s such that $\nu = \nu_a + \nu_s$ with $\nu_a \ll \mu$ and $\nu_s \perp \mu$.

The idea behind the proof of these theorems is as follows. For each $t \in \mathbb{R}_{\geq 0}$, let $P_t \sqcup N_t$ be a Hahn-decomposition of $\nu - t\mu$. Then, for all measurable $A \subset P_t$, we have $(\nu - t\mu)(A) \geq 0$ and so $\nu(A) \geq t\mu(A)$. Similarly, for all measurable $A \subset N_t$, we have $\nu(A) \leq t\mu(A)$. Let $t_1 < t_2$. if $A \in P_{t_1} \cap N_{t_2}$, then

$$t_1\mu(A) \leq \nu(A) \leq t_2\mu(A)$$

so it is reasonable to expect that the Radon-Nikodym derivative will be between t_1 and t_2 on $P_{t_1} \cap N_{t_2}$.

The above intuition motivates the following definition: For each $n \in \mathbb{Z}_{\geq 0}$, define

$$f_n(x) := \sup\{t \in 2^{-n}\mathbb{Z}_{\geq 0} \colon x \in P_t\}$$

where $2^{-n}\mathbb{Z}_{\geq 0}$ is the set of all numbers of the form $2^{-n}a$ with $a \in \mathbb{Z}_{\geq 0}$. Set $f(x) := \lim_{n\to\infty} f_n(x)$ for each x.

Lemma 1.8
There exists $A \in \mathcal{B}$ such that $\mu(E \backslash A) = 0$ and

$$\nu(B) = \int_B f \, d\mu$$

for all $B \subset A$.

Proof. For each $n \in \mathbb{Z}_{\geq 0}$, let $A_n := \{x \colon f_n(x) < \infty\}$. We first show that $\mu(E \backslash A_n) = 0$. To do this, fix $t \in 2^{-n}\mathbb{Z}_{\geq 0}$. Then $E \backslash A_n \subset \bigcup_{s \geq t} P_s$, with $\bigcup_{s \geq t} P_s$ positive for $\nu - t\mu$. We thus deduce that $\nu(E) \geq \nu(E \backslash A_n) \geq t\mu(E \backslash A_n)$. Taking $t \to \infty$, we conclude that $\mu(E \backslash A_n) = 0$.

Now, fix $B \subset A_n$. For $s \in 2^{-n}\mathbb{Z}_{\geq 0}$, let

$$B_t = \{x \in B \colon f_n(x) = t\}$$

Note that $B = \bigsqcup_t B_t$ and that, if $x \in B_t$, then $x \in P_t$ but $x \notin P_{t+2^{-n}}$. Thus, $B_t \subset P_t \cap N_{t+2^{-n}}$ and so we have

$$t\mu(B_t) \leq \nu(B_t) \leq (t + 2^{-n})\mu(B_t)$$

This implies that

$$\left|\nu(B_t) - \int_{B_t} f_n \, d\mu\right| \leq 2^{-n}\mu(B_t)$$

from which it follows that

$$\left|\nu(B) - \int_{B} f_n \, d\mu\right| \leq 2^{-n}\mu(B)$$

Finally, set $A := \bigcap_n A_n$. Picking $B \subset A$, the above holds for all n, so by the monotone convergence theorem, we conclude that

$$\int_B f \, d\mu = \lim_{n\to\infty} \int_B f_n \, d\mu = \nu(B)$$

Note also that $\mu(E\backslash A) = \mu(\bigcup(E\backslash A_n)) = 0$. ∎

Proof of Theorem 1.6. Since $\nu \ll \mu$, we have $\nu(E\backslash A) = 0$. Then for each $B \in \mathcal{B}$,

$$\nu(B) = \nu(B \cap A) + \nu(B\backslash A) = \int_{B\cap A} f \, d\mu = \int_B f \, d\mu$$

Thus, f is a Radon-Nikodym derivative.

It remains to show uniqueness. Let f, g be two RN derivatives. Suppose that, on the contrary, $f \neq g$ μ-a.e. Then there exists $\varepsilon > 0$ such that either $B = \{x \colon f(x) - g(x) > \varepsilon\}$ or $\{x \colon f(x) - g(x) < \varepsilon\}$ has positive measure. WLOG, let it be the first one. Then we have

$$0 = \nu(B) - \nu(B) = \int_B (f - g) \, d\mu \geq \int_B \varepsilon \, d\mu = \varepsilon\mu(B)$$

as required. ∎

Proof of Theorem 1.7. We first prove existence. Pick $A \in \mathcal{B}$ as in the previous lemma. Set $\nu_a = \nu|_A$ and $\nu_s = \nu|_{E\backslash A}$. Note that ν_s is concentrated on $E\backslash A$ whereas μ is concentrated on A, so $\nu \perp \mu$. Moreover, $\nu_a \ll \mu$ since f is a RN derivative for ν_a.

For uniqueness, suppose $\nu = \nu_a + \nu_s = \tilde{\nu}_a + \tilde{\nu}_s$. Then we can pick μ-null sets $D, \tilde{D}$ on which $\nu_s, \tilde{\nu}_s$ concentrate. Then $F := D \cup \tilde{D}$ is a μ-null set on which both ν_s and $\tilde{\nu}_s$ concentrate. Now, fix $G \in \mathcal{B}$. Since $\nu_a, \tilde{\nu}_a \ll \mu$, we have

$$\nu_a(G \cap F) = \tilde{\nu}_a(G \cap F) = 0$$

On the other hand, since $\nu_s, \tilde{\nu}_s$ both concentrate on F, we have

$$\nu_s(G\backslash F) = \tilde{\nu}_s(G\backslash F) = 0$$

It then follows that

$$\nu_a(G\backslash F) = \tilde{\nu}_a(G\backslash F) + \tilde{nu}_s(G\backslash F) - \nu_s(G\backslash F) = \tilde{\nu}_a(G\backslash F)$$

Putting everything together, we obtain $\nu_a(B) = \tilde{\nu}_a(B)$ and hence $\nu_a = \tilde{\nu}_a$. ∎

Example *Bernoulli convolution*
Fix $\lambda \in (0, 1)$. Let $(X_n)_{n\in\mathbb{N}}$ be an iid sequence of random variables that take values in $\{0, 1\}$ with equal probabilities. Let ν_λ be the distribution of the random variable

$$Y = \sum_{n=1}^{\infty} \lambda^n X_n$$

The first m values $X_1, \cdots, X_m$ determine the value of Y up to an error of at most $\sum_{n=m+1}^{\infty} 3^{-n} = 2/3^{m+1}$. If K is the set of values of Y, then it can be covered by 2^m intervals of length $2/3^{m+1}$. Then we have the estimate

$$|K| \leq 2^m \cdot \frac{2}{3^{m+1}} \to 0$$

On the other hand, $\nu_{1/3}(\mathbb{R}\backslash K) = 0$, showing that $\nu_{1/3} \perp dx$.

Remark. Solomyak has proven that $\nu_\lambda \ll dx$ for almost all $\lambda \in [1/2, 1)$. However, in a later example sheet, you will see that $\nu_\varphi \perp dx$ for $\varphi = (\sqrt{5} - 1)/2$. It remains a major open problem whether $v_{2/3} \ll dx$ or $\nu_{2/3} \perp dx$.

1.4 Lebesgue differentiation theorem

Motivation: From the previous subsection we know that an absolutely continuous measure $\mu \ll dx$ has a Radon-Nikodym derivative. But how do we find it? One might try to consider

$$\frac{\mu(B(x,r))}{|B(x,r)|}$$

and try to show that it converges to the Radon-Nikodym derivative as $r \to 0$. Indeed, this can be done via the Lebesgue differentiation theorem.

Another reason for studying this theorem is to obtain a generalisation of the fundamental theorem of calculus for the theory of Lebesgue integration.

Definition Lebesgue point
Let $f \in L^1(\mathbb{R}^d)$. A point $x \in \mathbb{R}^d$ is a Lebesgue point of f if

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0$$

Remark. If x is a Lebesgue point of f, then

$$\left| \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy - f(x) \right| \leq \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy \to 0$$

which gives us a generalisation of $F'(x) = f(x)$ that also makes sense in higher dimensions.

Theorem 1.9 Lebesgue differentiation theorem
Let $f \in L^1(\mathbb{R}^d)$. Then Lebesgue-a.e. $x \in \mathbb{R}^d$ is a Lebesgue point of f.

Definition Hardy-Littlewood maximal function
The Hardy-Littlewood maximal function Mf of $f \in L^1(\mathbb{R}^d)$ is

$$Mf(x) = \sup_{r\in\mathbb{R}_{>0}} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(t)|\, dt$$

Proposition 1.10 Maximal inequality
Let $f \in L^1(\mathbb{R}^d)$ and $t \in \mathbb{R}_{>0}$. Then

$$|\{x \in \mathbb{R}^d \colon Mf(x) > t\}| \leq \frac{5^d}{t} ||f||_1$$

Remark. The constant 5^d can be substantially improved; the best constant is known to grow at most linearly in d. It remains an open problem whether or not the constant can be made independent of d.

Now, let us prove the Lebesgue differentiation theorem assuming the above proposition. To do this, we will make use of the following density result which we will prove in the subsequent subsection.

Lemma 1.11
For all $f \in L^1(\mathbb{R}^d)$ and $\varepsilon > 0$, there exists $g \in C_c(\mathbb{R}^d)$ such that $||f - g||_1 < \varepsilon$.

Proof of Theorem 1.9. Fix $f \in L^1(\mathbb{R}^d)$. For each $t \in \mathbb{R}_{>0}$, define

$$A_t = \left\{ x \in \mathbb{R}^d \colon \limsup_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy > 2t \right\}$$

Observe that it suffices to show that $|A_t| = 0$ for all $t > 0$.

Fix $\varepsilon > 0$. By density of $C_c(\mathbb{R}^d)$ in $L^1(\mathbb{R}^d)$, we can pick $g \in C_c(\mathbb{R}^d)$ such that $||f - g||_1 < \varepsilon$. By the triangle inequality, we obtain the estimate

$$\begin{aligned}\frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy \leq{} & \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - g(y)|\, dy \\ & + \frac{1}{|B(x,r)|} \int_{B(x,r)} |g(y) - g(x)|\, dy \\ & + |f(x) - g(x)|\end{aligned}$$

Since g is continuous, we have

$$\limsup_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |g(y) - g(x)|\, dy = 0$$

so the estimate becomes

$$\frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy \leq M[f-g](x) + |f(x) - g(x)|$$

Now, if $x \in A_t$, then either $|M[f-g](x)| > t/2$ or $|f(x) - g(x)| > t/2$. By the maximal inequality (Proposition 1.10), we can estimate

$$|\{x \in \mathbb{R}^d \colon M[f-g](x) > t\}| \leq \frac{5^d}{t} ||f - g||_1 < \frac{5^d \varepsilon}{t}$$

On the other hand, we can use Markov's inequality to estimate

$$|\{x \in \mathbb{R}^d \colon |f(x) - g(x)| > t\}| \leq \frac{||f - g||_1}{t} < \frac{\varepsilon}{t}$$

Putting everything together, we have, for each $t > 0$,

$$|A_t| < \frac{5^d + 1}{t} \varepsilon$$

Taking $\varepsilon \to 0$ yields the required result. ■

Now, we need to prove the maximal inequality. Our proof will make use of the following lemma:

Lemma 1.12 Vitali covering lemma
Let $\mathcal{U}$ be a collection of balls in $\mathbb{R}^d$ with bounded diameter. Then there is a disjoint (finite or countably infinite) collection $\{V_1, V_2, \cdots\} \subset \mathcal{U}$ such that

$$\left| \bigcup_j V_j \right| = \infty \qquad \text{or} \qquad \bigcup_{U \in \mathcal{U}} U \subset \bigcup_j 5 \cdot V_j$$

where $5 \cdot V_j$ denotes the dilation of the ball V_j about its centre by a factor of 5.

Proof. We will construct the collection inductively. Pick $V_1 \in \mathcal{U}$ such that $\mathrm{diam}(V_1) \geq \mathrm{diam}(U)/2$ for all $U \in \mathcal{U}$. Suppose we have chosen $V_1, \cdots, V_n$ for some $n \geq 1$. Then pick $V_{n+1} \in \mathcal{U}$ such that it is disjoint from $V_1 \cup \cdots \cup V_n$ and $\mathrm{diam}(V_{n+1}) \geq \mathrm{diam}(U)/2$ for all $U \in \mathcal{U}$ for which U is disjoint from $V_1 \cup \cdots \cup V_n$.

If $|\bigcup_j V_j| = \sum_j |V_j| = \infty$, then we are done. Otherwise, we must have $\{V_j\}_j$ is finite or $\mathrm{diam}(V_n) \to 0$. Fix $U \in \mathcal{U}$. In either case, U will not be disjoint from $V_1 \cup \cdots \cup V_n$ for some n. Let N be the smallest such n. This implies that $U \cap V_N \neq \varnothing$. By construction, we have $\mathrm{diam}(U) \leq 2\mathrm{diam}(V_N)$. As an exercise, you can check (using the triangle inequality) that $U \subset 5 \cdot V_N$, completing the proof. ∎

Proof of Proposition 1.10. Note that, for each $x \in \mathbb{R}^d$ such that $Mf(x) > t$, we can pick $r(x) > 0$ such that

$$|B(x, r(x))| \leq t^{-1} \int_{B(x,r(x))} |f(t)|\, dt \leq t^{-1}||f||_1 \tag{†}$$

Denote by $U(x) = B(x, r(x))$. Observe that

$$\{x \in \mathbb{R}^d \colon Mf(x) > t\} \subset \bigcup_{x\colon Mf(x)>t} U(x)$$

We want

$$\left|\bigcup U(x)\right| \leq t^{-1} \int_{\bigcup U(x)} |f(t)|\, dt$$

Of course, if the balls are disjoint, the above does hold by σ-additivity. While this is not necessarily the case, the Vitali covering lemma helps! By (†), the set $\{\mathrm{diam}(U(x))\colon x \in \mathbb{R}^d, Mf(x) > t\}$ is bounded, so we can apply the lemma to $\mathcal{U} = \{U(x)\colon Mf(x) > t\}$. Each V_j given by the lemma satisfies (†), so we have

$$\left|\bigcup_{x\colon Mf(x)>t} U(x)\right| \leq \left|\bigcup_j 5 \cdot V_j\right| \leq 5^d \sum_j |V_j| \leq 5^d t^{-1}||f||_1$$

as required. ∎

1.5 Littlewood's principles

In his *Lectures on the Theory of Functions*, Littlewood stated three principles: "Every (measurable) set is nearly a finite sum of intervals; every function (of class L^p) is nearly continuous; every convergent sequence of functions is nearly uniformly convergent."

The first principle can be stated precisely as follows:

Proposition 1.13
Let $A \subset \mathbb{R}^d$ with $|A| < \infty$. Then for every $\varepsilon > 0$, there exists $B \subset \mathbb{R}^d$ which is a finite union of rectangles, such that $|A \triangle B| < \varepsilon$.

Exercise. Prove the above proposition.

Next, a more precise formulation of the second principle is given by Lusin's theorem. Its proof makes use of the regularity of finite Borel measures on metric spaces, which we now recall from II Probability and Measure.

Lemma 1.14 Regularity of finite Borel measures
Let (E, dist) be a metric space, $\mathcal{B}$ be the Borel σ-algebra of E, and μ a finite Borel measure on $(E, \mathcal{B})$. Then for all $A \in \mathcal{B}$ and $\varepsilon > 0$, there exists a closed set $K \subset E$ and an open set $U \subset E$ such that $K \subset A \subset U$ and $\mu(U \backslash K) < \varepsilon$.

Theorem 1.15 (Weak) Lusin's theorem
Let (E, dist) be a metric space and μ a finite Borel measure. If $f: E \to \mathbb{C}$ is measurable, then for every $\varepsilon > 0$, there exists a closed set $K \subset E$ such that $\mu(E \backslash K) < \varepsilon$ and $f|_K$ is continuous.

Proof. We will prove this result in steps: we first consider the special case of indicator functions and gradually extend the result to measurable functions.

STEP 1: Indicator functions on Borel sets
Let $A \in \mathcal{B}$. Fix $\varepsilon > 0$. By regularity of μ, we can pick closed $F \subset E$ and open $U \subset E$ such that $F \subset A \subset U$ and $\mu(U \backslash F) < \varepsilon$. Recall from II Linear Analysis that metric spaces are normal, so by Urysohn's lemma, there exists a continuous function $g: E \to [0, 1]$ such that $g|_{E \backslash U} = 0$ and $g|_F = 1$. In this case, we can set $K = (E \backslash U) \cup F$. Note that $\mathbb{1}_A|_K = g|_K$ is continuous, and $\mu(E \backslash K) = \mu(U \backslash F) < \varepsilon$, as required.

STEP 2: Simple functions
Fix $\varepsilon > 0$. Suppose f is a simple function, given by $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ with each $A_i \in \mathcal{B}$. Apply Step 1 for each $\mathbb{1}_{A_i}$ to obtain closed sets K_i satisfying $\mu(E \backslash K_i) < \varepsilon/n$ and $\mathbb{1}_{A_i}|_{K_i}$ continuous. Set $K = \bigcap_{i=1}^n K_i$. Note that K is closed, $\mu(E \backslash K) \le \sum_{i=1}^n \mu(E \backslash K_i) < \varepsilon$ and $\mathbb{1}_{A_i}|_K$ continuous.

STEP 3: Bounded measurable functions
Fix $\varepsilon > 0$. Suppose f is a bounded measurable function. Recall from II Probability and Measure that simple functions are dense in $L^\infty(E)$, so we can pick a sequence $(f_n)_{n \in \mathbb{N}}$ of simple functions that uniformly converge to f. For each f_n, apply Step 2 to get closed K_n such that $\mu(E \backslash K_n) < \varepsilon 2^{-n}$ and $f_n|_{K_n}$ is continuous. Set $K = \bigcap_{n \in \mathbb{N}} K_n$. Note that K is closed, $\mu(E \backslash K) \le \sum_{n=1}^\infty \mu(E \backslash K_n) < \varepsilon$ and $f|_K$ is continuous (by the uniform limit theorem).

STEP 4: Measurable functions
Fix $\varepsilon > 0$. Suppose f is a measurable function. For each $n \in \mathbb{N}$, define $E_n = \{x \in E: |f(x)| \le n\}$. Note that $(\mu(E_n))_{n \in \mathbb{N}}$ is a monotonically increasing sequence, bounded above by $\mu(E) < \infty$. Thus, $\mu(E_n) \to \mu(E)$. Pick $N \in \mathbb{N}$ such that $\mu(E \backslash E_N) < \varepsilon/2$. Apply Step 3 to the bounded measurable function $f \mathbb{1}_{E_N}$, yielding a closed $\tilde{K} \subset E$ such that $\mu(E \backslash \tilde{K}) < \varepsilon/2$ and $f \mathbb{1}_{E_N}|_{\tilde{K}}$ is continuous. Now, consider the Borel set $\tilde{K} \cap E_N$. Note that $\mu(E \backslash (\tilde{K} \cap E_N)) \le \mu(E \backslash E_N) + \mu(E_N \backslash K) < \varepsilon$. By inner regularity of μ, we can pick closed $K \subset E$ such that $K \subset \tilde{K} \cap E_N$ and $\mu((\tilde{K} \cap E_N) \backslash K) < \varepsilon - \mu(E \backslash (\tilde{K} \cap E_N))$. Then $\mu(E \backslash K) < \varepsilon$ and $f|_K$ is continuous. ■

Theorem 1.16 (Strong) Lusin's theorem
Let (E, dist) be a metric space and μ a finite Borel measure. If $f: E \to \mathbb{C}$ is measurable, then, for every $\varepsilon > 0$, there exists a closed set $K \subset E$ and $g \in C(E)$ such that $f = g$ on K, and $\mu(E \backslash K) < \varepsilon$. If, in addition, f is also bounded, then the conclusion holds with $||g||_\infty = ||f||_\infty$.

Proof. We simply apply Tietze's extension theorem (cf. II Linear Analysis, including variants in Example Sheet 3 Q1 and Q2). ■

As an application of Lusin's theorem, let's prove density of $C_c(\mathbb{R}^d)$ in $L^1(\mathbb{R}^d)$:

Proof of Lemma 1.11. Fix $f \in L^1(\mathbb{R}^d)$ and $\varepsilon > 0$. For $r > 0$, set

$$E_r := \{x \in \mathbb{R}^d \colon x \in [-r,r]^d, |f(x)| \leq r\}$$

Note that $|f(x)\mathbb{1}_{E_r}(x)| \leq |f(x)|$ and $f(x)\mathbb{1}_{E_r}(x) \to f(x)$ for a.e. x. Thus, by the dominated convergence theorem, $||f\mathbb{1}_{E_r} - f||_1 \to 0$ as $r \to \infty$. Pick $R > 0$ sufficiently large such that $||f\mathbb{1}_{E_R} - f||_1 < \varepsilon/3$. Note that $|[-R,R]^d| < \infty$, so we can apply (weak) Lusin's theorem (Theorem 1.15) to the measurable function $f_R := (f\mathbb{1}_{E_R})|_{[-R,R]^d}$ defined on $[-R,R]^d$. We thus obtain a closed subset $K \subset [-R,R]^d$ such that $|[-R,R]^d \backslash K| < \varepsilon/6R$ and $f_R|_K$ is continuous. Note that K is also closed in $\mathbb{R}^d$. Thus, by Tietze's extension theorem, there exists $g \in C(\mathbb{R}^d)$ such that $g|_K = f_R|_K$ and $||g||_\infty = ||f_R|_K||_\infty = R$. Now, set $B := \mathbb{R}^d \backslash (-R-\delta, R+\delta)^d$, where $\delta > 0$ is chosen sufficiently small such that $|(-R-\delta,R+\delta)^d \backslash [-R,R]^d| < \varepsilon/3R$. Note that B and K are both closed, so by Urysohn's lemma, there exists $\phi \colon \mathbb{R}^d \to [0,1]$ such that $\phi|_B = 0$ and $\phi|_K = 1$. Set $h = g\phi$. Note that $||h||_\infty \leq ||g||_\infty$ by construction. Finally, observe that $h \in C_c(\mathbb{R}^d)$ and

$$\begin{aligned} ||h-f||_1 &\leq ||f\mathbb{1}_{E_R} - f||_1 + \int_{[-R,R]^d \backslash K} |h(x) - f(x)\mathbb{1}_{E_R}(x)|\,dx \\ &\quad + \int_{(-R-\delta,R+\delta)^d \backslash [-R,R]^d} |h(x) - f(x)\mathbb{1}_{E_R}(x)|\,dx \\ &< \frac{\varepsilon}{3} + \left|[-R,R]^d \backslash K\right| \sup_{[-R,R]^d \backslash K} |h - f\mathbb{1}_{E_R}| \\ &\quad + \left|(-R-\delta,R+\delta)^d \backslash [-R,R]^d\right| \sup_{(-R-\delta,R+\delta)^d \backslash [-R,R]^d} |h| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{6R} \cdot 2R + \frac{\varepsilon}{3R} \cdot R = \varepsilon \end{aligned}$$

as required. ■

Finally, the third of Littlewood's principles is made rigorous by Egorov's theorem.

Theorem 1.17 Egorov's theorem
Let $(E, \mathcal{B}, \mu)$ be a finite measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions such that $f_n(x) \to f(x)$ for μ-a.e. $x \in E$. Then for every $\varepsilon > 0$, there exists a set $A \in \mathcal{B}$ such that $(f_n|_A)_{n\in\mathbb{N}}$ converges uniformly and $\mu(E\backslash A) < \varepsilon$.

Proof. For each $k, N \in \mathbb{N}$, define

$$A_{k,N} = \left\{x \in E \colon |f(x) - f_n(x)| < \frac{1}{k} \ \forall\, n \geq N\right\}$$

Note that $(A_{k,N})_{N\in\mathbb{N}}$ is an increasing sequence, with

$$\bigcup_{N\in\mathbb{N}} A_{k,N} = \{x \in E \colon f_n(x) \to f(x)\}$$

Thus, we have $\mu\left(E\backslash \bigcup_{N\in\mathbb{N}} A_{k,N}\right) = 0$. Since $(E\backslash \bigcup_{N=1}^n A_{k,N})_{n\in\mathbb{N}}$ is a decreasing sequence with $\mu(E) < \infty$, we have $\mu\left(E\backslash \bigcup_{N=1}^n A_{k,N}\right) \downarrow 0$ and thus, for each $k \in \mathbb{N}$, we can pick $n(k)$ such that

$$\mu\left(E\backslash \bigcup_{N=1}^{n} A_{k,N}\right) < \frac{\varepsilon}{2^k}$$

Now, define $A = \bigcap_k A_{k,n(k)}$. Observe that $\mu(E\backslash A) < \sum_{k=1}^\infty \frac{\varepsilon}{2^k} = \varepsilon$ and that $|f_N(x) - f(x)| < 1/k$ for all $N \geq n(k)$ and $x \in A$, so $(f_n|_A)_{n\in\mathbb{N}}$ converges uniformly. ■

1.6 Riesz representation theorem

Let E be a compact metric space and $\mathcal{B}$ the Borel σ-algebra. We denote by $C(E)$ the space of continuous functions on E. Recall from II Linear Analysis that this is a Banach space when endowed with the norm $||\cdot||_\infty$.

A *linear functional* is a map $L\colon C(E) \to \mathbb{C}$ such that $L(a_1 f_1 + a_2 f_2) = a_1 L(f_1) + a_2 L(f_2)$ for all $a_1, a_2 \in \mathbb{C}$ and $f_1, f_2 \in C(E)$. We say that a linear functional L is *bounded* if there exists $C > 0$ such that $|L(f)| \leq C||f||_\infty$ for all $f \in C(E)$. We say that a linear functional L is *positive* if $L(f) \geq 0$ for all $f \in C(E)$ with $f(E) \subset \mathbb{R}_{\geq 0}$.

Example
Let E be a compact metric space. For any finite Borel measure μ,

$$L(f) = \int_E f \, d\mu$$

defines a bounded positive linear functional on $C(E)$.

Theorem 1.18 Riesz representation theorem
Let E be a compact metric space, and let L be a positive bounded linear functional on $C(E)$. Then there is a unique finite Borel measure μ such that

$$L(f) = \int f \, d\mu$$

for all $f \in C(E)$. Moreover, $||L|| = \mu(E)$.

> There is a version of the theorem for bounded linear functionals that may not be positive.
>
> **Theorem 1.19** Riesz representation theorem
> *For every bounded linear function L on $C(E)$, there exists a unique complex Borel measure μ such that*
>
> $$L(f) = \int f \, d\mu$$
>
> *for all $f \in C(E)$. Moreover, $||L|| = ||\mu||$, where $||\mu||$ is the total variation norm of the μ.*

Proof of uniqueness. Let μ_1, μ_2 be finite measures such that $\int f \, d\mu_1 = \int f \, d\mu_2$ for all $f \in C(E)$. Fix $A \in \mathcal{B}$. Pick K_1, K_2 compact and U_1, U_2 open such that $K_j \subset A \subset U_j$ and $\mu_j(U_j \backslash K_j) < \varepsilon$ for $j \in \{1, 2\}$. Set $U = U_1 \cap U_2$ and $K = K_1 \cup K_2$. Then $K \subset A \subset U$ and $\mu_j(U \backslash K) < \varepsilon$ for $j \in \{1, 2\}$. By Urysohn, we can pick $g \in C(E)$ such that $g|_K = 1$ and $g|_{E\backslash U} = 0$. Then

$$\left| \int \mathbb{1}_A \, d\mu_j - \int g \, d\mu_j \right| < \varepsilon$$

Since $\int g \, d\mu_1 = \int g \, d\mu_2$, it follows that

$$|\mu_1(A) - \mu_2(A)| < \varepsilon.$$

Taking $\varepsilon \to 0$, we deduce that $\mu_1(A) = \mu_2(A)$. Hence, $\mu_1 = \mu_2$. ■

We will not be proving existence in this course.

1.7 Lebesgue spaces

Let $(E, \mathcal{B}, \mu)$ be a measure space. For a measurable function $f\colon E \to \mathbb{C}$ and $p \in [1, \infty)$, we write

$$||f||_p := \left(\int_E |f|^p \, d\mu \right)^{1/p}$$

and

$$||f||_\infty := \inf\{t \geq 0 \colon |f(x)| \leq t \text{ for } \mu\text{-a.e. } x \in E\}$$

where, by convention, the infimum of an empty set is taken to be ∞. For $p \in [1, \infty]$, we denote by $\mathcal{L}^p(E, \mathcal{B}, \mu)$ the set of all measurable functions $E \to \mathbb{C}$ for which $|||f||_p < \infty$. Writing $f \sim g \iff f = g$ a.e., we then define the Lebesgue space $L^p(E, \mathcal{B}, \mu) := \mathcal{L}^p(E, \mathcal{B}, \mu)/\sim$. Moreover, in the case that E is a topological space, there is also a notion of *local* Lebesgue spaces:

Definition Local Lebesgue space L^p_{loc}
Let E be a topological space and μ a Borel measure. For $p \in [1, \infty]$, we define

$$L^p_{\text{loc}}(E) := \{f \colon E \to \mathbb{C} \colon f\mathbb{1}_K \in L^p(E) \ \forall \text{ compact } K \subset E$$

Recall the Minkowski inequality: if $f, g \in L^p$, then $f + g \in L^p$ with

$$||f + g||_p \leq ||f||_p + ||g||_p$$

It is then easy to check that L^p is a Banach space for $p \in [1, \infty]$.

Finally, we also recall Hölder's inequality: if $f \in L^p(E)$ and $g \in L^q(E)$ for $p, q \in [1, \infty]$ that satisfy $1/p + 1/q = 1$, then $fg \in L^1(E)$ with

$$||fg||_1 \leq ||f||_p ||g||_q$$

Now, let us begin with a generalisation of Minkowski's inequality:

Theorem 1.20 Minkowski's integral inequality
Let $(E, \mathcal{A}, \mu)$ and $(F, \mathcal{B}, \nu)$ be σ-finite measure spaces and $p \in [1, \infty]$. Suppose $G\colon E \times F \to \mathbb{R}$ are $\mathcal{A} \otimes \mathcal{B}$-measurable. If $\int_F ||G(\cdot, y)||_p \, d\nu(y) < \infty$, then

$$g(x) = \int G(x, y) \, d\nu(y)$$

exists for μ-a.e. x and

$$||g||_p \leq \int ||G(\cdot, y)||_p \, d\nu(y)$$

Remark. Note that, if $F = \{1, 2\}$ with counting measure ν, then the above reduces to the usual Minkowki inequality.

Proof of Theorem 1.20. The result is vacuously true when g vanishes μ-a.e., so we may assume WLOG that this is not the case.

For $p = \infty$, note that $|G(x, y)| \leq ||G(\cdot, y)||_\infty$ for $\mu \otimes \nu$-a.e. (x, y), so

$$|g(x)| \leq \int |G(x, y)| \, d\nu(y) \leq \int ||G(\cdot, y)||_\infty \, d\nu(y)$$

for μ-a.e. $x \in E$.

Now, suppose $p < \infty$. If $p = 1$, the result is immediate from Fubini's theorem, so it remains to carefully consider the case $p \in (1, \infty)$. Observe that

$$\begin{aligned}
\int_E \left| \int_F G(x,y)\, d\nu(y) \right|^p d\mu(x) &\le \int_E \left(\int_F |G(x,y)|\, d\nu(y) \right) |g(x)|^{p-1}\, d\mu(x) \\
&= \int_F ||G(\cdot, y) g^{p-1}||_1 \, d\nu(y) \\
&\le \int_F ||G(\cdot, y)||_p ||g^{p-1}||_{p/(p-1)} \, d\nu(y) \\
&= ||g||_p^{p-1} \int_F ||G(\cdot, y)||_p \, d\nu(y)
\end{aligned}$$

Dividing both sides by $||g||_p^{p-1}$ yields the result. ■

For the rest of this subsection, we specialise to $E = \mathbf{R}^d$ equipped with the Lebesgue measure.

Proposition 1.21
For every $p \in [1, \infty)$, $C_c(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.

Proof. The proof is identical to the $p = 1$ case. ■

Proposition 1.22
The space of $L^p(\mathbb{R}^d)$ is separable for all $p \in [1, \infty)$.

Proof. Let $p \in [1, \infty)$. Fix $f \in L^p(\mathbb{R}^d)$ and $\varepsilon > 0$. Let

$$\mathcal{R}_{\mathbb{Q}} := \left\{ \prod_{i=1}^d (a_i, b_i] \colon a_i, b_i \in \mathbb{Q} \right\}$$

and

$$S_{\mathbb{Q}} := \left\{ s = \sum_{k=1}^N (\alpha_k + i\beta_k) \mathbb{1}_{R_k} \colon N \in \mathbb{N}, \alpha_k, \beta_k \in \mathbb{Q}, R_k \in \mathcal{R}_k \right\}$$

Note that $S_{\mathbb{Q}}$ is countable. We would then be done if we can show that $S_{\mathbb{Q}}$ is dense in $L^p(\mathbb{R}^d)$.

Fix $f \in L^p(\mathbb{R}^d)$ and $\varepsilon > 0$. By density of $C_c(\mathbb{R}^d)$ in $L^p(\mathbb{R}^d)$, we can pick $h \in C_c(\mathbb{R}^d)$ such that $||h - f||_p < \varepsilon/2$. We can then pick $K > 0$ such that $\operatorname{supp} h \subset [-K/2, K/2]^d$. Choose $0 < \eta < \varepsilon/2K^d$. Since h is continuous, we can pick $\delta > 0$ such that $|x - y| < \delta \implies |h(x) - h(y)| < \eta$. Pick $N \in \mathbb{N}$ sufficiently large such that $N^d \ge 1/\delta$. Partition $[-K/2, K/2]$ into a finite union of cubes $\{C_k\}_{k=1}^N$ from $\mathcal{R}_{\mathbb{Q}}$ of side length less than $1/N$. For each cube C_k in the partition, pick $\alpha_k, \beta_k \in \mathbb{Q}$ such that $|\alpha_k + i\beta_k - h(x)| < \eta$ for all $x \in C_k$. Now, observe that

$$\left|\left| \sum_{k=1}^N (\alpha_k + i\beta_k) \mathbb{1}_{C_k} - h \right|\right|_p \le \sum_{k=1}^N ||(\alpha_k + i\beta_k - h)\mathbb{1}_{C_k}||_p < K^d \eta < \frac{\varepsilon}{2}$$

Hence,

$$\left|\left| \sum_{k=1}^N (\alpha_k + i\beta_k) \mathbb{1}_{C_k} - f \right|\right|_p < \varepsilon$$

as required. ■

For $a \in \mathbb{R}^d$, denote by τ_a the translation-by-a operator, $\tau_a f(x) = f(x - a)$.

Proposition 1.23
Let $p \in [1, \infty)$ and $f \in L^p(\mathbb{R}^d)$. Then the map

$$\begin{aligned} \mathbb{R} &\to L^p(\mathbb{R}^d) \\ a &\mapsto \tau_a f \end{aligned}$$

is continuous.

Proof. Fix $\varepsilon > 0$. Pick $h \in C_c(\mathbb{R}^d)$ such that $||h - f||_1 < \varepsilon/3$. Note that h is uniformly continuous, so for every $\eta > 0$, there exists $\delta > 0$ such that $|x - y| < \delta \implies |h(x) - h(y)| < \eta$. Then, for any $a, b \in \mathbb{R}^d$ with $|a - b| < \delta$, we have

$$|\tau_a h(x) - \tau_b h(x)| = |h(x - a) - h(x - b)| < \eta$$

Note that $|\tau_a h(x) - \tau_b h(x)| > 0$ only if $x \in \operatorname{supp} \tau_a h \cup \operatorname{supp} \tau_b h$. The measure of this set is bounded above by $C = C(h)$. Thus, we can estimate

$$||\tau_a h - \tau_b h||_p \leq \eta C^{1/p}$$

from which it follows that

$$\begin{aligned} ||\tau_a f - \tau_b f||_p &\leq ||\tau_a f - \tau_a h||_p + ||\tau_a h - \tau_b h||_p + ||\tau_b h - \tau_b f||_p \\ &= 2||f - h||_p + ||\tau_a h - \tau_b h||_p \\ &< \frac{2\varepsilon}{3} + \eta C^{1/p} \end{aligned}$$

Picking η sufficiently small, we then get $||\tau_a f - \tau_b f||_p < \varepsilon$, as required. ■

Theorem 1.24
Let $(E, \mathcal{B}, \mu)$ be a σ-finite measure space. Let $p \in [1, \infty)$ and let q be defined by $p^{-1} + q^{-1} = 1$. Then for all $g \in L^q(E)$, the map

$$L_g(f) = \int fg \, d\mu$$

is a bounded linear functional on $L^p(E)$ with $||L_g|| = ||g||_q$. Conversely, for every bounded linear functional L on $L^p(E)$, there is a unique $g \in L^q(E)$ such that $L = L_g$.

Remark. Taking $E = \mathbb{N}$ equipped with counting measure, we recover the familiar result $\ell_q^* \equiv \ell_p$ from II Linear Analysis.

Proof. The fact that $L_g \in L^p(E)^*$ with $||L_g|| \leq ||g||_q$ is immediate from linearity of integration and Hölder's inequality. Now, let us show that $||L_g|| = ||g||_q$. If $p > 1$, set

$$f(x) = \begin{cases} g(x)^{-1}|g(x)|^q, & g(x) \neq 0 \\ 0, & g(x) = 0 \end{cases}$$

and observe that $||f||_p^p = ||g^{p(q-1)}||_1 = ||g||_q^q$ and $L_g(f) = ||g||_q^q = ||f||_p ||g||_q$. In the case $p = 1, q = \infty$, we want to show that the set $A := \{x \in E \colon |g(x)| > ||L||\}$ has measure zero. Suppose, on the contrary, that $\mu(A) > 0$. Set

$$f(x) = \begin{cases} g(x)^{-1}, & x \in A \\ 0, & x \notin A \end{cases}$$

and note that $||f||_1 < ||L||^{-1}\mu(A)$ and $L(f) = \int fg \, d\mu = \mu(A)$. Thus, $\mu(A) \leq ||L_g|| ||f||_1 < \mu(A)$ — a contradiction!

Now, we prove the converse. Uniqueness follows easily from the fact that $||L_{g_1} - L_{g_2}|| = ||L_{g_1-g_2}|| = ||g_1 - g_2||_q$. For existence, fix a bounded linear function L on $L^p(E)$. We will first prove the result for finite measure space and extend it to the σ-finite case.

CASE 1: $\mu(E) < \infty$

In this case, $L^\infty(E) \subset L^p(E)$, and we may define the function $\nu(A) = L(\mathbb{1}_A)$ for $A \in \mathcal{B}$. This is finitely additive because L is linear. To show that ν is σ-additivity, it suffices to show that $\lim_{n\to\infty} \nu(A_n) = 0$ for every decreasing sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{B}$ with $\bigcap_n A_n = \varnothing$. Indeed,

$$|\nu(A_n)| = |L(\mathbb{1}_A)| \leq ||L||||\mathbb{1}_A||_p = ||L||\,\mu(A)^{1/p} \to 0$$

Thus, we have now shown that ν is a complex measure.

Now, note that $\nu(A) = 0 \implies \mathbb{1}_A = 0$ a.e. $\implies L(1_{\mathbb{A}}) = 0 \implies \nu(A) = 0$, so $\nu \ll \mu$. By the Radon-Nikodym theorem (Theorem 1.6), we can pick $g \in L^1(E)$ such that

$$L(\mathbb{1}_A) = \nu(A) = \int \mathbb{1}_A g\, d\mu$$

By linearity of L and integration, we have that $L(f) = \int fg\, d\mu$ for all simple functions f. By density of simple functions in $L^\infty(E)$, the same holds for all $f \in L^\infty(E)$. Finally, density of $L^\infty(E)$ in $L^p(E)$ gives us $L(f) = \int fg\, d\mu$ for all $f \in L^p(E)$.

To complete the proof in this case, it suffices to show that $g \in L^q(E)$. If $p > 1$, consider

$$g_R(x) = \begin{cases} g(x), & |g(x)| \leq R \\ R\operatorname{sign}(g(x)), & \text{otherwise} \end{cases}$$

and let

$$f_R(x) = \begin{cases} g_R(x)^{-1} g_R(q)^q, & g_R(x) \neq 0 \\ 0, & g_R(x) = 0 \end{cases}$$

Note that $||f_R||_p = ||g_R||_q^{q-1}$ and that

$$L(f_R) = \int f_R g\, d\mu \geq \int f_R g_R\, d\mu \geq ||g_R||_q^q$$

Thus, $||g_R||_q^q \leq ||g_R||_q^{q-1}||L||$ and so $||g_R||_q \leq ||L||$. Taking $R \to \infty$, we get that $g \in L^q$ with $||g||_q \leq ||L||$.

CASE 2: μ is σ-finite

In this case, we can pick an increasing sequence $(E_n)_{n\in\mathbb{N}}$ such that $\bigcup_n E_n = E$ and $\mu(E_n) < \infty$ for each $n \in \mathbb{N}$. Note that there are natural inclusions $L^p(E_1) \subset L^p(E_2) \subset \cdots \subset L^p(E)$ by identifying a function on a smaller subset with a function on the larger subset but which vanishes outside the smaller subset. Applying the result of Case 1, we can pick $g_n \in L^p(E_n)$, for each n, such that $L|_{L^p(E_n)} = L_{g_n}$. By injectivity of $g \mapsto L_g$, we have that $g_m|_{E_n} = g_n$ whenever $m > n$. It follows that we can define $g\colon E \to \mathbb{C}$ by setting $g(x) = g_n(x)$ whenever $x \in E_n$. It is measurable since each g_n is measurable and $E = \bigcup_n E_n$. Since $||g_n||_q \leq ||L||$, we also have that $||g||_q \leq ||L||$ by the monotone convergence theorem. By construction, $L(f) = L_g(f)$ for all $f \in \bigcup_n L^p(E_n)$. Note that this space is dense in $L^p(E)$ (since each $f \in L^p(E)$ can be approximated by $f\mathbb{1}_{E_n} \in L^p(E_n) \subset L^p(E)$). Hence, we in fact have $L(f) = L_g(f)$ for all $f \in L^p(E)$. ■

Remark. Note the technique used in the proof to prove σ-additivity, namely that it suffices to show finite additivity and that $\lim_{n\to\infty} \nu(A_n) = 0$ for every decreasing sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{B}$ with $\bigcap_n A_n = \varnothing$. Indeed, let $(E_n)_{n\in\mathbb{N}}$ be a sequence of pairwise disjoint sets in $\mathcal{B}$. Set $S_N = \bigcup_{n=1}^N E_n$. By finite additivity, $\mu(S_N) = \sum_{n=1}^N \mu(E_n)$. Note also that $(S_\infty \backslash S_N)_{N\in\mathbb{N}}$ is a decreasing sequence with $\bigcap_N (S_\infty \backslash S_N) = \varnothing$. Thus, by the second condition, we have $\mu(S_\infty) = \lim_{N\to\infty} \mu(S_N) = \sum_{n=1}^\infty \mu(E_n)$, as required.

1.8 Convolutions and mollifications

Definition Convolution
Let f, g be measurable functions $\mathbb{R}^d \to \mathbb{C}$. We define their convolution to be

$$(f \star g)(x) = \int f(t)g(x-t)\,dt$$

if the above integral exists for a.e. $x \in \mathbb{R}^d$.

Lemma 1.25
Let f, g be measurable functions $\mathbb{R}^d \to \mathbb{C}$.
(i) $(f \star g)(x) = (g * f)(x)$ *if the integral exists at x*
(ii) *For any $a \in \mathbb{R}^d$, $\tau_a(f \star g) = \tau_a(f) \star g = f \star \tau_a(g)$ provided the convolutions exist.*

Now, when do convolutions exist and, if they do, how regular are they?

Theorem 1.26 Young's inequality
Let $p, q, r \in [1, \infty]$ satisfy $p^{-1} + q^{-1} = r^{-1} + 1$. If $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$, then the integral $f \star g$ exists for a.e. x. Moreover, $f \star g \in L^r(\mathbb{R}^n)$ with the estimate

$$||f \star g||_r \leq ||f||_p ||g||_q$$

In lectures, we will only prove two special cases.

Proof for $q = 1, p = r$. Consider the map

$$G\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$$
$$(x, t) \mapsto g(t)f(x-t)$$

Note that this is measurable and that

$$\int ||G(\cdot, t)||_p\,dt = \int |g(t)|||\tau_t f||_p\,dt = ||f||_p||g||_1 < \infty$$

Thus, by Minkowski's integral inequality (Theorem 1.20), we have that $g \star f$ exists for a.e. $x \in X$ (and thus $f \star g$ as well). Moreover, we have the estimate

$$||f \star g||_r \leq \int ||G(\cdot, t)||_p\,dt = ||f||_p||q||_1$$

as required. ■

Proof for $r = \infty$. Since $p^{-1} + q^{-1} = 1$, we have

$$|(f \star g)(x)| \leq \int |f(t)g(x-t)|\,dt \leq ||f||_p||q||_q$$

by Hölder's inequality. ■

Remark. Note that, in the $r = \infty$ case, the proof actually gives us a stronger result! In this case, the integral exists and the inequality holds for all $x \in \mathbb{R}^d$, not just a.e.

Lemma 1.27
Let $f, g \in L^1(\mathbb{R}^d)$ and $h \in L^\infty(\mathbb{R}^d)$. Then

$$(f \star g) \star h = f \star (g \star h)$$

Proof. Existence of the convolutions follows immediately from Young's inequality (Theorem 1.26). Observe that

$$[(f \star g) \star h](x) = \int (f \star g)(t) h(x-t)\, dt = \int \left[\int f(s) g(t-s)\, ds \right] h(x-t)\, dt$$

Now, note that

$$\int \int |f(s) g(t-s) h(x-t)|\, dt\, ds \leq ||f||_1 ||g||_1 ||h||_\infty < \infty$$

We can thus use Fubini to deduce that

$$\begin{aligned}
[(f \star g) \star h](x) &= \iint f(s) g(t-s) h(x-t)\, dt\, ds \\
&= \iint f(s) g(u) h(x-s-u)\, du\, ds \\
&= \int f(s) (g \star h)(x-s)\, ds \\
&= [f \star (g \star h)](x)
\end{aligned}$$

as required. ■

> **Remark.** The result still holds under the conditions $f \in L^p(\mathbb{R}^d), g \in L^q(\mathbb{R}^d), h \in L^r(\mathbb{R}^d)$ if $p^{-1} + q^{-1} + r^{-1} \geq 2$.

Lemma 1.28
Let $p, q \in [1, \infty]$ satisfy $p^{-1} + q^{-1} = 1$. If $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$, then $f \star g \in C(\mathbb{R}^d)$.

Proof. By symmetry in p and q, we may assume WLOG that $q < \infty$. By Young's inequality (Theorem 1.26) and the remark after the proof for $r = \infty$, we see that, for every $x, y \in \mathbb{R}^d$,

$$\begin{aligned}
|(f \star g)(x) - (f \star g)(y)| &= |(f \star g)(x) - [\tau_{y-x}(f \star g)](x)| \\
&= |[f \star (g - \tau_{y-x} g)](x)| \\
&\leq ||f||_p ||g - \tau_{y-x} g||_q
\end{aligned}$$

By continuity of translations in L^q (Proposition 1.23), we deduce that $|(f \star g)(x) - (f \star g)(y)| \to 0$ as $y \to x$, as required. ■

Notation *Multi-index notation*
Given $\alpha \in \mathbb{Z}^d_{\geq 0}$ and a function $f : \mathbb{R}^d \to \mathbb{C}$, we write

$$|\alpha| = \sum_{i=1}^d |\alpha_i|$$

and

$$D^\alpha = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$$

whenever the partial derivative exists.

Definition Schwartz space
The Schwartz space on $\mathbb{R}^d$ is defined by

$$\mathcal{S}(\mathbb{R}^d) := \left\{ f \in C^\infty(\mathbb{R}^d) \colon \| |x|^n D^\alpha f \|_\infty < \infty \right\}$$

for all $n \in \mathbb{Z}_{\geq 0}$ and $\alpha \in \mathbb{Z}^d_{\geq 0}$.

Proposition 1.29
Let $f \in L^p(\mathbb{R}^d)$ for some $p \in [1, \infty]$. If $g \in \mathcal{S}(\mathbb{R}^d)$, then $f \star g \in C^\infty(\mathbb{R}^d)$ and

$$D^\alpha (f \star g) = f \star (D^\alpha g)$$

for every multi-index $\alpha \in \mathbb{Z}^d_{\geq 0}$.

Proof. By iteration, it suffices to show the result for $|\alpha| = 1$. Fix a multi-index α with $|\alpha| = 1$. Then we have

$$\begin{aligned} \frac{(f \star g)(x + t\alpha) - (f \star g)(x)}{t} - (f \star D^\alpha g)(x) &= \frac{[\tau_{-t\alpha}(f \star g)](x) - (f \star g)(x)}{t} - (f \star D^\alpha g)(x) \\ &= f \star \left(\frac{\tau_{-t\alpha} g - g}{t} - D^\alpha g \right)(x) \end{aligned}$$

For each $x \in \mathbb{R}^d$, the mean value theorem gives us $\tilde{s}(x)$ between 0 and t for which

$$\frac{(\tau_{-t\alpha})g(x) - g(x)}{t} = D^\alpha g(x + \tilde{s}(x)\alpha)$$

Apply the mean value theorem once more, we obtain $s(x)$ between 0 and $\tilde{s}(x)$ such that

$$\frac{(\tau_{-t\alpha})g(x) - g(x)}{t} - D^\alpha g(x) = \tilde{s}(x) D^{2\alpha} g(x + s(x)\alpha)$$

and thus

$$\left| \frac{(\tau_{-t\alpha})g(x) - g(x)}{t} - D^\alpha g(x) \right| \leq |t| \left| D^{2\alpha} g(x + s(x)\alpha) \right|$$

Next, we note that $|D^{2\alpha} g(x)| \leq C(|x| + 10)^{-d-1}$ holds for $C = \|(|x| + 10)^{d+1} D^{2\alpha} g(x)\|_\infty$ (which is finite since $g \in \mathcal{S}(\mathbb{R}^d)$. It then follows that

$$\left\| \frac{(\tau_{-t\alpha})g(x) - g(x)}{t} - D^\alpha g(x) \right\|_q \leq C|t| \, \|(|x + s(x)\alpha| + 10)^{-d-1}\|_q$$

For every $t \in (-1, 1)$, we then obtain the estimate

$$\left\| \frac{(\tau_{-t\alpha})g(x) - g(x)}{t} - D^\alpha g(x) \right\|_q \leq C|t| \, \|(|x| + 9)^{-d-1}\|_q$$

Noting that $\|(|x| + 9)^{-d-1}\|_q^q \leq \int_{\mathbb{R}^d} |x|^{-2q} \, dx \lesssim \int_0^\infty r^{-2q} \, dr < \infty$, we deduce that

$$\left\| \frac{(\tau_{-t\alpha})g(x) - g(x)}{t} - D^\alpha g(x) \right\|_q \to 0$$

and the result follows from Young's inequality (Theorem 1.26). ∎

Definition Approximation of the identity
We say that a sequence of functions $(f_n)_{n \in \mathbb{N}}$ in $L^1(\mathbb{R}^d)$ is an approximation of the identity if

(i) $f_n(x) \geq 0$ *for all $x \in \mathbb{R}^d$ and all $n \in \mathbb{N}$*
(ii) $\displaystyle\int f_n \, dx = 1$ *for all $n \in \mathbb{N}$*
(iii) $\displaystyle\int_{|x| > \varepsilon} f_n(x) \, dx \to 0$ *for all $\varepsilon > 0$*

Lemma 1.30
Let $f \in L^1(\mathbb{R}^d)$ be nonnegative with $\int f\,dx = 1$. The sequence (f_n) defined by

$$f_n(x) = n^d f(nx)$$

is an approximation of the identity.

Proof. It remains to show (iii). By change of variables and dominated convergence, we have

$$\int_{|x|>\varepsilon} f_n(x)\,dx = \int_{|x|>n\varepsilon} f(x)\,dx \to 0$$

as required. ■

Theorem 1.31
Let (f_n) be an approximation of the identity. For every $g \in L^p(\mathbb{R}^d)$, we have $||f_n \star g - g||_p \to 0$ as $n \to \infty$.

Proof. WLOG, we may assume that $f_n(x) = 0$ on $|x| > 1$ for all $n \in \mathbb{N}$. To see why, consider

$$\tilde{f}_n(x) = \begin{cases} f_n(x)\left(\int_{|x|\le 1} f_n\,dx\right)^{-1}, & |x| \le 1 \\ 0, & |x| > 1 \end{cases}$$

It is easy to check that $||f_n - \tilde{f}_n||_1 \to 0$. By Young's inequality (Theorem 1.26), we then have $||f_n \star g - \tilde{f}_n \star g||_p \le ||f_n - \tilde{f}_n||_1 ||g||_p \to 0$. Thus, $||f_n \star g - g||_p \to 0$ iff $||\tilde{f}_n \star g - g||_p \to 0$.

Fix $\varepsilon > 0$. Pick $h \in C_c(\mathbb{R}^d)$ such that $||g - h|| < \varepsilon$. Since h is uniformly continuous, we can pick $\delta > 0$ such that $|x - y| < \varepsilon \implies |h(x) - h(y)| < \varepsilon$. Pick $N \in \mathbb{N}$ sufficiently large such that $\int_{|x|>\delta} f_N(x)\,dx < \varepsilon$. We then have

$$\begin{aligned} |(f_n \star h)(x) - h(x)| &\le \int f_n(t)|h(x-t) - h(x)|\,dt \\ &= \int_{|t|<\delta} f_n(t)|h(x-t) - h(x)|\,dt + \int_{|t|\ge\delta} f_n(t)|h(x-t) - h(x)|\,dt \\ &\le \varepsilon \int f_n(t)\,dt + 2||h||_\infty \int f_n(t)\,dt \\ &\le \varepsilon(1 + 2||h||_\infty) \end{aligned}$$

Pick $R > 0$ such that $\operatorname{supp} h \subset [-R, R]^d$. Then we have

$$\operatorname{supp}(h \star f_n) \subset \operatorname{supp} h + \operatorname{supp} f_n \subset [-R-1, R+1]^d$$

We then have that

$$||f_n \star h - h||_p \le [2(R+1)]^{d/p}(1 + 2||h||_\infty)\varepsilon$$

Hence, by the triangle inequality, we conclude that

$$\begin{aligned} ||f_n \star g - g||_p &\le ||f_n \star (g - h)||_p + ||f_n \star h - h||_p + ||h - g||_p \\ &\le 2||g - h||_p + |f_n \star h - h||_p \\ &= \left\{2 + [2(R+1)]^{d/p}(1 + 2||h||_\infty)\right\}\varepsilon \end{aligned}$$

as required. ■

Approximations of the identity can be particularly useful when they are C_c^∞. A standard example would be to consider the function

$$\eta(x) = \begin{cases} C \exp\left(-\frac{1}{1-|x|^2}\right), & |x| < 1 \\ 0, & |x| \geq 1 \end{cases}$$

where C is chosen so that $\int \eta = 1$. By Lemma 1.30, we then get an approximation of the identity (η_n). This is sometimes called the *standard mollifier.*

Corollary 1.32
For every $p \in [1, \infty)$, $C_c^\infty(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.

Proof. Fix $f \in L^p$ and $\varepsilon > 0$. Pick $R > 0$ sufficiently large such that $||f - f\mathbb{1}_{[-R,R]^d}|| < \varepsilon/2$. Set $g = f\mathbb{1}_{[-R,R]^d}$. By Theorem 1.31, we can pick $n \in \mathbb{N}$ such that $||g - g \star \eta_n|| < \varepsilon/2$. By Proposition 1.29, $g \star \eta_n \in C^\infty$. Since η_n is compactly supported, we have $g \star \eta_n \in C_c^\infty$. Finally, note that $||f - g \star \eta_n|| < \varepsilon$. ∎

2 Hilbert and Banach space analysis

Recall from II Linear Analysis that a *normed space* is a vector space X over $\mathbb{C}$ equipped with a map $||\cdot||\colon X \to \mathbb{R}_{\geq 0}$ such that

(i) $||x|| \geq 0$ for all $x \in X$, with $||x|| = 0 \iff x = 0$

(ii) $||\lambda x|| = |\lambda| ||x||$ for all $x \in X, \lambda \in \mathbb{C}$

(iii) $||x+y|| \leq ||x|| + ||y||$ for all $x, y \in X$

In this course, the field will always be $\mathbb{C}$ unless otherwise stated.

A normed space X that is complete is called a *Banach space*. If X is a Banach space equipped with a sesquilinear form $\langle \cdot, \cdot \rangle \colon X \times X \to \mathbb{C}$ such that $||x||^2 = \langle x, x \rangle$, then we say it is a *Hilbert space*.

A *linear functional* on a normed space X is a map $f\colon X \to \mathbb{C}$ such that $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$ for all $\lambda \in \mathbb{C}$ and $x, y \in X$. We say that a linear functional $f\colon X \to \mathbb{C}$ is *bounded* if

$$||f|| := \sup_{x \in B_X} |f(x)| < \infty$$

We denote by X^* the space of bounded linear functions $X \to \mathbb{C}$. It is a Banach space with respect to $||\cdot||$ as defined above.

Notation. We write $f(x) = \langle x, f \rangle$ for $f \in X^*$ and $x \in X$.

2.1 Hahn-Banach theorem and its consequences

Theorem 2.1 Hahn-Banach theorem
Let X be a normed space and $Y \subset X$ a subspace of X. For every $g \in Y^*$, there exists $f \in X^*$ such that $f|_Y = g$ and $||f|| = ||g||$.

An important consequence of the Hahn-Banach theorem is that

$$||x|| = \max_{f \in B_{X^*}} |\langle x, f \rangle|$$