ANALYSIS OF FUNCTIONS
(D COURSE - PART II MATHEMATICAL TRIPOS)

CLÉMENT MOUHOT - C.MOUHOT@DPMMS.CAM.AC.UK

Contents

Preamble. These notes are inspired by the books of Lieb & Loss [4], Brézis [1], Kolmogorov-Fomin [2, 3] and Rudin [5]. There is no claim of originality in the material presented here which is standard, apart from the arrangement and presentation. There will be three example sheets and a mock exam.

1. Integration of functions

Structure of the Tripos imposes constraints. We take for granted the construction of the Lebesgue measure as well as familiarity with abstract measure theory. This is covered in the course *Probability & Measure* or the books already mentioned. We focus on the applications to the analysis of functions.

1.1. **Recall: Lebesgue measure and integration.** Measure theory invention (Borel, Lebesgue, Radon, Fréchet...) was motivated by defining length and volume for more and more complicated sets; Lebesgue integration theory relies on measure theory and its invention was motivated by solving some theoretical and practical defiencies in Riemann integration theory.

Exercise 1. *Show that the pointwise limit of a sequence of Riemann-integrable functions is not necessarily Riemann-integrable.*

1.1.1. *Recalls on measure theory.* Consider a set X, and $\mathcal{P}(X)$ denotes the set of its subsets.

Definition 1.1 (Algebra). *$\mathcal{A} \subset \mathcal{P}(X)$ is an algebra is if it (i) is stable by **finite** union, (ii) is stable by absolute difference ($A \in \mathcal{A} \Rightarrow X \setminus A \in \mathcal{A}$), (iii) contains the whole set X.*

Definition 1.2 (σ-Algebra). *$\mathcal{A} \subset \mathcal{P}(X)$ is a σ-algebra is if it (i) is stable by **countable** union, (ii) is stable by absolute difference, (iii) contains the whole set X. $(X, \mathcal{A})$ is then a **measurable space**.*

Remark 1.3. *Compare with a* topology $\mathcal{T} \subset \mathcal{P}(X)$ *on X: contains X and $\emptyset$, stable by **any** union, **finite** intersections.*

The property of begin a σ-algebra is stable by intersection, therefore there is a notion of smallest σ-algebra containing a given set $\mathcal{M} \subset \mathcal{P}(X)$. When $\mathcal{M} = \mathcal{T}$ is the topology given on X, this is the *Borel sets*. In $\mathbb{R}^n$ it is the smaller σ-algebra containing all open balls, denoted $\mathcal{B}(\mathbb{R}^n)$.

Date: Lent 2017.

Definition 1.4 (Measure)**.** *A measure on* $(X,\mathcal{A})$ *is an application* $\mu : \mathcal{A} \to [0,+\infty]$ *s.t.* $\mu(\emptyset)=0$ *and* μ *is* σ*-additive (i.e. countably additive).* $(X,\mathcal{A},\mu)$ *is then a* ***measure space****. The measure space is said to* ***complete*** *if* $A \in \mathcal{A}$ *with* $\mu(A)=0$ *and* $B \subset A$ *implies* $B \in \mathcal{A}$ *(and of course* $\mu(B)=0$*).*

Exercise 2. *Construct an additive function that is* not σ*-additive on the Borel sets.*

The completion of the Borel sets $\mathcal{B}(\mathbb{R}^n)$ is denoted $\mathcal{L}(\mathbb{R}^n)$.

Exercise 3. *What are the cardinals of* $\mathcal{B}(\mathbb{R}^n)$ *and* $\mathcal{L}(\mathbb{R}^n)$*?*

> **Theorem** (Existence of the Lebesgue mesure) *There exists a unique measure on* $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ *s.t.* $\mu(\prod_{i=1}^n [a_i,b_i]) = \prod_{i=1}^n (b_i - a_i)$ *(measure of hypercubes).*

Remark 1.5. *This measure is* ***σ-finite****: there is a countable increasing covering sequence of sets with finite measures (e.g.* $[-N,N]^n$, $N \in \mathbb{N}^*$*).*

Definition 1.6 (Measurable functions)**.** *Consider two measurable spaces* $(X,\mathcal{A})$, $(Y,\mathcal{B})$*, then a function* $f : X \to Y$ *is measurable if for any* $B \in \mathcal{B}$ *then* $f^{-1}(B) \in \mathcal{A}$*.*

Remark 1.7. *When* $\mathcal{B}$ *is the Borel sets, we call* f ***Borel function*** *and it is enough to check the definition for any* $B \in \mathcal{T}$*. Standard operations preserve measurability: composition, coordinate concatenation, coordinate restriction...*

Proposition 1.8 (Sequence of measurable functions)**.** *Consider two measurable spaces* $(X,\mathcal{A})$, $(Y,\mathcal{B})$ *with* Y *metric space and* $\mathcal{B}$ *the Borel sets of* Y*. Consider a sequence of measurable functions* $f_k : X \to Y$ *s.t.* f_k *converges pointwise to* f*. Then* f *is measurable.*

Proof. Since $\mathcal{B}$ can be formed from open sets through the operations of countable union, countable intersection, and complement, it is enough to check that $f^{-1}(U) \in \mathcal{A}$ for any $U \in \mathcal{T}_Y$. Using the metric d_Y we define

$$U_n := \{x \in Y \ : \ d_Y(x, Y \setminus U) > 1/n\}, \quad F_n := \{x \in Y \ : \ d_Y(x, Y \setminus U) \geq 1/n\}$$

and characterise U and $f^{-1}(U)$ as

$$U = \bigcup_{n\geq 1} U_n = \bigcup_{n \geq 1} F_n \quad \text{with} \quad U_n \subset F_n \subset U_{n+1} \subset F_{n+1} \subset \cdots \subset U.$$

$$f^{-1}(U) = f^{-1}\left(\bigcup_{n\geq 1} U_n\right) = \bigcup_{n\geq 1} f^{-1}(U_n) \subset \bigcup_{n\geq 1}\bigcup_{\ell\geq 1}\bigcap_{k\geq \ell} f_k^{-1}(U_n).$$

Observe that the last set inclusion is not necessarily an equality as U_n is open. However for our particular sets we have $\cup_{\ell\geq 1}\cap_{k\geq\ell} f_k^{-1}(U_n) \subset \cup_{\ell\geq 1}\cap_{k\geq\ell} f_k^{-1}(F_n)$ and by closure of F_n we have $\cup_{\ell\geq 1}\cap_{k\geq\ell} f_k^{-1}(F_n) \subset f^{-1}(F_n)$, therefore

$$f^{-1}(U) = \bigcup_{n\geq 1} f^{-1}(U_n) \subset \bigcup_{n\geq 1}\bigcup_{\ell\geq 1}\bigcap_{k\geq \ell} f_k^{-1}(U_n) \subset \bigcup_{n\geq 1}\bigcup_{\ell\geq 1}\bigcap_{k\geq \ell} f_k^{-1}(F_n) \subset \bigcup_{n\geq 1} f^{-1}(F_n) = f^{-1}(U)$$

which implies set equality and finally (countable union and intersection) that $f^{-1}(U) \in \mathcal{A}$. □

Exercise 4. *(1) When* $(Y,\mathcal{B}) = (\mathbb{R},\mathcal{B}(\mathbb{R}))$*, adapt the proof and generalise the limit to limsup and liminf. (2) Is the statement still true when* Y *is merely a topological space? (3) Is the statement still true when the pointwise convergence is replaced by the convergence almost everywhere? (*Hint: it requires the space to be complete: build a counter-example otherwise, and prove it with this property.*)*

1.2. **Integrability and convergence theorems.** Riemann integration theory is based on subdividing the *input space* with the total order of the real line (whereas the output space can be a general Banach space). By contrast Lebesgue integration theory is based on subdivising the *output space* with the order of the the real line whereas the input space can be general measure space. The respective approximation tools to go from discrete to continuous analysis are (1) *Riemann sums* on the one hand and (2) *simple functions* on the other hand. We consider from now *real or complex-valued functions* over a measure space $(X, \mathcal{A}, \mu)$. The real line is always endowed with Borel sets.

Definition 1.9 (Simple functions)**.** *A function $f : X \to [0, +\infty)$ is **simple** if it is measurable and takes a finite number of values.*

Remark 1.10. *The* characteristic function *of a set $A \subset X$ is denoted χ_A (returns one on A, zero elsewhere). Check that χ_A measurable iff $A \in \mathcal{A}$.*

Proposition 1.11. *Let $f : X \to [0, +\infty]$ measurable, then there is an increasing sequence s_k of simple functions converging pointwise to f (including "convergence to $+\infty$" in $[0, +\infty]$).*

Proof. Define for any $n \geq 1$: $B_n := \{x \mid f(x) > n\}$ and $A^i_n := \{x \mid f(x) \in [(i-1)/2^n, i/2^n]\}$ for $i = 1, \dots, n2^n$. These sets are all in $\mathcal{A}$ and partition X. Define s_n so that $s_n \equiv (i-1)/2^n$ on A^i_n and $s_n \equiv n$ on B_n. The proof of the convergence and monotonicity are straightforward. □

Definition 1.12 (Integral of simple functions)**.** *The integral of a simple function $s := \sum_{i=1}^n \alpha_i \chi_{A_i}$, $\alpha_i \in [0, +\infty)$, $A_i \in \mathcal{A}$ on the set $E \in \mathcal{A}$ is defined as*

$$\int_E s \, \mathrm{d}\mu := \sum_{i=1}^n \alpha_i \mu(A_i \cap E).$$

Remark 1.13. *Observe then that the map $A \in \mathcal{A} \mapsto \int_A s \, \mathrm{d}\mu$ is a measure (σ-additivity is inherited from that of μ by finite sum).*

Definition 1.14 (Integral of positive measurable functions)**.** *Let $f : X \to [0, +\infty]$ measurable. Its integral on $E \in \mathcal{A}$ is defined as*

$$\int_E f \, \mathrm{d}\mu := \sup \left\{ \int_E s \, \mathrm{d}\mu \mid s \text{ simple function and } s \leq f \right\} \in [0, +\infty].$$

Remark 1.15. *The set $[0, +\infty] = [0, +\infty) \cup \{+\infty\}$ is metrised with $d(x, y) = |\arctan x - \arctan y|$ and its Borel sets can be formed from the basis $(a, +\infty]$, $a \geq 0$. Observe that the integral* always *makes sense in $[0, +\infty]$ for positive measurable functions. Standard properties of the integral are satisfied.*

Exercise 5. *(1) Prove Chebyshev's inequality from this definition: $\mu(\{x \mid f(x) \geq \alpha\}) \leq \alpha^{-1} \int_X f \, \mathrm{d}\mu$. (2) Prove that if $f : X \to [0, +\infty]$ is measurable with finite integral then $\mu(\{x \mid f(x) = +\infty\}) = 0$.*

Theorem 1.16 (Beppo Levi - Lebesgue's monotone convergence theorem)**.** *Consider an increasing sequence of measurable functions $f_k : X \to [0, +\infty]$ ($f_k \leq f_{k+1}$) that converges pointwise to f, then*

$$\forall\, E \in \mathcal{A}, \quad \lim_{k \to +\infty} \int_E f_k \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu.$$

Proof. By considering $f_k \chi_E$ it is enough to consider $E = X$. The sequence $\int_X f_k \, \mathrm{d}\mu$ of $[0, \infty]$ is increasing therefore converges to some $\alpha \in [0, +\infty]$. Since $\int_X f_k \, \mathrm{d}\mu \leq \int_X f \, \mathrm{d}\mu$ by monotonicity, this limit $\alpha \leq \int_X f \, \mathrm{d}\mu$. Let s a simple function below f and $c \in (0, 1)$. Define $E_k := \{x \in X \mid f_k(x) \geq cs(x)\}$. Check that $E_k \in \mathcal{A}$ measurable, $E_k \subset E_{k+1}$, and $\cup_{k \geq 1} E_k = X$. By continuity from below of the measure $A \in \mathcal{A} \to \int_A s \, \mathrm{d}\mu$, we have $\int_X s \, \mathrm{d}\mu = \lim_{k \to \infty} \int_{E_k} s \, \mathrm{d}\mu$. Take $k \to +\infty$ and $c \to 1^-$ in

$$\int_X f_k \, \mathrm{d}\mu \geq \int_{E_k} f_k \, \mathrm{d}\mu \geq \int_{E_k} cs \, \mathrm{d}\mu = c \int_{E_k} s \, \mathrm{d}\mu.$$

to get $\alpha \geq \int_X s \, \mathrm{d}\mu$. Since it is true for all simple function s below f it implies $\alpha \geq \int_X f \, \mathrm{d}\mu$, which concludes the proof. □

Exercise 6. *(1) For any sequence $f_k : X \to [0,+\infty]$ prove that $\int_X (\sum_{k\ge 1} f_k)\,\mathrm{d}\mu = \sum_{k\ge 1}\int_X f_k\,\mathrm{d}\mu$. (2) Prove that for $f : X \to [0,+\infty]$ measurable, $\nu : A \in \mapsto \int_A f\,\mathrm{d}\mu$ is a measure, and that for $g : X \to [0,+\infty]$ measurable, $\int_X g\,\mathrm{d}\nu = \int_X fg\,\mathrm{d}\mu$. (3) (Fatou's lemma) For any sequence $f_k : X \to [0,+\infty]$ prove that $\int_X (\liminf f_k)\,\mathrm{d}\mu \le \liminf(\int_X f_k\,\mathrm{d}\mu)$.*

Definition 1.17 (Integral of real or complex valued functions)**.** *A measurable function $f : X \to \mathbb{C}$ is **integrable** if $|f| : X \to [0,+\infty)$ (whose integral is defined above) satisfies $\int_X |f|\,\mathrm{d}\mu < +\infty$. Its integral is then computed by splitting real/imaginary and positive/negative parts.*

Theorem 1.18 (Lebesgue's dominated convergence theorem)**.** *Let $f_k : X \to \mathbb{C}$ be a sequence of measurable functions that converges pointwise to f. We assume (*domination*) that there is $g : X \to [0,+\infty]$ measurable and with finite integral so that $|f_k| \le g$ for all k. Then f_k and f are integrable and $\int_X |f_k - f|\,\mathrm{d}\mu \to 0$.*

Proof. The domination immediately implies the integrability of f_k and f. Consider $h_k := 2g - |f - f_k| \ge 0$ valued in $[0,+\infty]$ which converges pointwise to $2g$. Use Fatou's lemma:

$$\int_X 2g\,\mathrm{d}\mu = \int_X \liminf h_k\,\mathrm{d}\mu \le \liminf \int_X h_k = \int_X 2g\,\mathrm{d}\mu - \limsup \int_X |f_k - f|\,\mathrm{d}\mu$$

which implies $\limsup \int_X |f_k - f|\,\mathrm{d}\mu = 0$ and concludes the proof. □

Remark 1.19. *The previous results extend when the pointwise convergence is replaced by the pointwise convergence almost everywhere on some $A \in \mathcal{A}$ with $\mu(X \setminus A) = 0$: replace f by $f\chi_A$.*

Exercise 7. *Formulate and prove continuity/differentiability of $t \mapsto \int_X F(t,\cdot)\,\mathrm{d}\mu$ when F satisfies proper continuity/differentiability and domination assumption.*

Remark 1.20. *In summary the construction of the theory of integration of complex-valued function has followed the "standard process": (1) characteristic functions using the base measure, (2) simple functions by linear combination, (3) positive functions by Beppo-Levi, (4) real or complex-valued functions through the modulus.*

Exercise 8. *(1) Prove that for a bounded function on $[a,b]$, Riemann-integrability implies Lebesgue-measurability (for the Lesbesgue σ-algebra) but not necessarily Borel-measurability, and implies Lebesgue-integrability with both integrals agreeing. (2) Prove that the Riemann-integrability is equivalent to the set of discontinuity having zero Lebesgue measure.*

1.3. **Lebesgue spaces: completeness, separability.**

Definition 1.21. *We denote, for $p \in [1,+\infty]$, $L^p(X)$ the set of equivalence classes of measurable functions $f : X \to \mathbb{R}$ (or $\to \mathbb{C}$) such that $|f|^p$ is integrable (resp. f essentially bounded when $p = +\infty$, i.e. bounded outside a nullset set), for the relation of almost everywhere equality.*

Theorem 1.22 (Riesz-Fischer)**.** *Endowed with $\|f\|_{L^p(X)} := (\int_X |f|^p\,\mathrm{d}\mu)^{1/p}$ (resp. the essential supremum $\|f\|_{L^\infty(X)} := \inf\{M \ge 0\ s.t.\ \mu(\{x \in X \mid |f(x)| \ge M\}) = 0\}$ when $p = +\infty$), the space $L^p(X)$ is a Banach space (i.e. a complete normed vector space).*

Exercise 9. *Suppose $f \in L^\infty(X)$ is supported on a set of finite measure, then prove that $f \in L^p(X)$ for all $p \in [1,+\infty)$, and $\|f\|_{L^p} \to \|f\|_{L^\infty}$ as $p \to +\infty$.*

Proof. Standard properties (Minkowski and Hölder inequalities) are left to the reader. We prove the completeness. Assume first $p \in [1,+\infty)$: We prove the following auxiliary result: *consider a sequence g_k of $L^p(X)$ functions so that $\sum \|g_k\|_{L^p(X)} < +\infty$, then there exists $G \in L^p(X)$ so that $\sum_{k=1}^n g_k \to G$ pointwise almost everywhere and in $L^p(X)$.*

Proof of the auxiliary result: Define $h_n := \sum_{k=1}^n |g_k|$ and $h := \sum_{k=1}^\infty |g_k|$ two functions from X to $[0,+\infty]$. Since h_n^p increases and converges pointwise to h^p in $[0,+\infty]$, the Beppo-Levi theorem implies $\int_X h_n^p\,\mathrm{d}\mu \to \int_X h^p\,\mathrm{d}\mu$. Moreover by triangular inequality $\|h_n\|_{L^p(X)} \le \sum_{k=1}^n \|g_k\|_{L^p(X)} \le M$, and thus $h \in L^p(X)$ and is finite almost everywhere. Hence, the series $\sum g_k$ converges absolutely almost everywhere, hence converges almost everywhere, and we call this pointwise limit G. This limit satisfies $|G(x)| \leq |\sum_{k=1}^{\infty} g_k| \leq \sum_{k=1}^{\infty} |g_k(x)| = h(x)$ and is therefore L^p. Finally the integral

$$\int_X \left| G(x) - \sum_{k=1}^{n} g_k(x) \right| \, \mathrm{d}\mu(x) \to 0$$

by applying the dominated convergence theorem: the integrand converges pointwise to zero, and the following domination holds: $|G(x) - \sum_{k=1}^{n} g_k(x)|^p \leq 2^p h^p$, where h^p integrable. This concludes the proof of the auxiliary result.

Back to the main proof: consider f_n Cauchy sequence, we choose a subsequence so that the series $g_k := f_{\varphi(k+1)} - f_{\varphi(k)}$ satisfies $\|g_k\|_p \leq \frac{1}{2^k}$, which implies that $\sum_{k=1}^{\infty} \|g_k\|_{L^p(X)} < +\infty$. The auxiliary result shows that there is a $G \in L^p(X)$ so that $\sum_{k=1}^{n} g_k = f_{\varphi(n+1)} - f_{\varphi(1)} \to G$ pointwise almost everywhere and in $L^p(X)$, and we define $f := G + f_{\varphi(1)}$. Then (Cauchy property) $\int_X |f_m - f_n|^p \,\mathrm{d}\mu \to 0$ as $m, n \to \infty$ and taking $m = \varphi(k) \to \infty$ we get $f_n \to f$ in L^p.

When $p = +\infty$: by removing the nullset set $A := \cup_{m,n} A_{m,n}$ with $A_{m,n} := \{x \mid |f_m(x) - f_n(x)| > \|f_n - f_m\|_\infty$, we are left with a sequence uniformly Cauchy for $\|\cdot\|_\infty$ on $X \setminus A$ and the rest of the proof is standard. ☐

Remark 1.23. *Observe that we have proved the following auxiliary result interesting* per se*: the convergence in L^p implies the pointwise convergence almost everywhere of' a subsequence.*

Theorem 1.24 (Abstract density result)**.** *Suppose that $(X, \mathcal{A}, \mu)$ is a measure space and $p \in [1, +\infty]$. Then the simple functions that belong to $L^p(X)$ are dense in $L^p(X)$.*

Remark 1.25. *Note that a simple function $s = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$ belongs to L^p ($p \in [1, +\infty)$) iff $\mu(A_i)$ finite as soon as $\alpha_i \neq 0$, and all simple functions belong to L^∞.*

Proof. For f real or complex, split real/imagnary positive/negative parts and approximate each: it is enough to deal with non-negative functions. For $f \geq 0$, use again s_n so that $s_n \equiv (i-1)/2^n$ on A_n^i and $s_n \equiv n$ on B_n with $B_n := \{x \mid f(x) > n\}$ and $A_n^i := \{x \mid f(x) \in [(i-1)/2^n, i/2^n]\}$ for $i = 1, \ldots, n2^n$. This produces a sequence that converges pointwise and from below towards $f \geq 0$. The Beppo-Levi theorem then implies the convergence in L^p norm. ☐

Theorem 1.26 (Density-separability result in $\mathbb{R}^n$)**.** *Consider O open set of $\mathbb{R}^n$ and $p \in [1, +\infty)$, then $L^p(O)$ is separable, and moreover smooth functions compactly supported in O are dense in it.*

We will need and admit a key property of the Lebesgue measure:

> **Theorem** (Regularity of the Lebesgue mesure) *A **regular measure** on a topological space (with its Borel sets) is a measure for which every measurable set can be approximated from above by open sets (outer regularity) and from below by compact sets (inner regularity). The Lebesgue measure on $\mathbb{R}^n$ is regular.*

Exercise 10. *For a regular measure, completing the σ-algebra with the nullset sets does not bring any complications in most proofs: prove indeed that any measurable set with finite measure is squeezed between two Borel sets with same measure.*

Proof. Let us prove separability: Consider the countable base $\mathcal{C}$ of open sets of $\mathbb{R}^n$ made up of hypercubes with rational coordinates.

Then we first show that: *any open set O can be covered with a countable union of rational cubes with disjoint interiors.* The proof is done by the following inductive proceedure: on the grid $\mathbb{Z}^n$ retains the cubes fully inside O, discard those fully outside, for those remaining bisect them into 2^n smaller half-length cubes and iterate the process.

Then we approximate any simple functions by simple functions using $A_i \in \mathcal{C}$ and rational values $\alpha_i \in \mathbb{Q}$ as follows: the outer regularity of the measure allows approximation of A_i by an open set O_i (with small error on the measure), then we cover O_i with a countable collection of closed cubes $C_{i,k}$

as above, with $\mu(O_i) = \sum_{k=1}^{\infty} \mu(C_{i,k})$, and since the series converges the partial sum $\sum_{k=1}^{N} \mu(C_{i,k})$ can approximate the target measure with small error, with a finite number of those cubes; we finally approximate the coefficient with a rational number.

Then for the second point of the statement we construct smooth compactly supported approximation of the characteristic function on each cube. The approximation of the characteristic function by a continuous function is straightforward with a piecewise affine function. The approximation by a smooth compactly function requires the use the function e^{-1/x^2} as seen in Analysis II. ☐

Remark 1.27. *Approximation relies most of the time on the convolution by an approximation of the unit. But this **still** requires the construction of a smooth compactly supported function, using* e^{-1/x^2} *or a variant. And it is important to observe that the convergence of* $\varphi_k * f \to f$ *in* L^p *for* φ_k *approximation of the unit relies on the fact that the translation operator is continuous in* L^p, $p \in [1, +\infty)$ *(prove it).*

Exercise 11. *Try to generalise as far as possible the previous statement to an abstract topological space* X *instead of* $\mathbb{R}^n$*, replacing smooth by continuous (hint: second-countable,* T^4 */ Urysohn's lemma. . .).*

Exercise 12. *Prove that* $L^\infty(\mathbb{R}^n)$ *is **not** separable (hint: consider the family* $\chi_{B_r(0)}$ *for* $r > 0$*).*

1.4. How regular are measurable and integrable functions?

Definition 1.28 (Lesbegue points)**.** *For a function* $f : \mathbb{R}^n \to \mathbb{C}$ *we say that* x *is a **Lebesgue point** if* $\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \,\mathrm{d}\mu(y) \to 0$ *(averaging limit).*

Exercise 13. *Prove that all points of continuity are Lebesgue.*

Theorem 1.29 (Lebesgue's differentation and density Theorems)**.** *For an integrable function* $f \in L^1(\mathbb{R}^n)$*, almost every points are Lebesgue (differentiation). This implies on the Borel sets (density): let* $E \in \mathcal{B}(\mathbb{R}^n)$ *then for almost every* $x \in \mathbb{R}^n$ *the density ratio* $\frac{|E \cap B_r(x)|}{|B_r(x)|} \to \chi_E(x)$ *as* $r \to 0$.

Proof of the density theorem assuming the differentiation theorem. For $|x| \le M$, $r < 1$, the ratio rewrites $\frac{|E\cap B_r(x)|}{|B_r(x)|} = \frac{|E\cap B_{M+1}(x)\cap B_r(x)|}{|B_r(x)|}$, then apply the differentiation theorem to the integrable function $f = \chi_{E\cap B_{M+1}(x)}$. ☐

Proof of the Lebesgue's differentiation Theorem. General strategy: The desired limit is satisfied for continuous functions and L^1 functions can be approximated by continous functions in L^1, hence it is enough to prove that the "amount" of non-Lebesgue points is controlled by the L^1 norm of the function. This "amount" will be measured by the following operator.

*Step 1: Define the **Hardy-Littlewood maximal operator*** $M_f(x) := \sup_{r>0} m_f(x, r)$ *with* $m_f(x, r) := \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f| \,\mathrm{d}\mu$*. Then for* $a > 0$, $E_a := \{x \mid M_f(x) > a\}$ *is an open set.*

If $x \in E_a$, then $M_f(x) > a$ and there is $r > 0$ s.t. $m_f(x, r) > a$. Let $\epsilon > 0$ s.t. $\left(\frac{r}{r+\epsilon}\right)^n > \frac{a}{m_f(x,r)}$. Then in a neighborhood of x, i.e. for $y \in B(x, \epsilon)$, we have

$$m_f(y, r+\epsilon) = \frac{1}{\mu(B(y, r+\epsilon))} \int_{B(y,r+\epsilon)} |f| \,\mathrm{d}\mu \ge \left(\frac{r}{r+\epsilon}\right)^n \frac{1}{\mu(B(x, r))} \int_{B(x,r)} |f| \,\mathrm{d}\mu = \left(\frac{r}{r+\epsilon}\right)^n m_f(x, r) > a$$

where we have used $B(x, r) \subset B(y, r + \epsilon)$. This proves the step. Remark it implies that $M_f : \mathbb{R}^n \to [0, +\infty]$ is measurable, since the sets $(a, +\infty)$ generate the Borel sets.

Step 2: Vitali's covering lemma: Consider a set $X \subset \mathbb{R}^n$ *that is included in a finite number of open balls* $X \subset \cup_{i=1}^{N} B(x_i, r_i)$*. There a subset of indeces* $J \subset \{1, \dots, N\}$ *s.t. the collection of balls* $(B(x_j, r_j))_{j\in J}$ *are pairwise disjoint and* $X \subset \cup_{j\in J} B(x_j, 3r_j)$*.*

Assume wlog the radiuses are ranked $r_1 \ge r_2 \ge \cdots \ge r_N$. Then consider $B(x_1, r_1)$: all balls that intersect it are included in $B(x_1, 3r_1)$, remove them, then consider the ball with the second largest radius in this new collection and argue similarly, and continue inductively. By finite induction it builds the desired collection.

Step 3: For all $a > 0$ one has $\mu(E_a) \leq \frac{3^n}{a} \|f\|_{L^1(\mathbb{R}^n)}$[1].

Consider any compact set $K \subset E_a$: for any $x \in K \subset E_a$ there is r_x so that $m_f(x, r_x) > a$. Then $K \subset \cup_{x \in K} B(x, r_x) \subset \cup_{i=1}^N B(x_i, r_{x_i})$ (finite covering by compactness). The Vitali covering Lemma gives then indeces J s.t. $K \subset \cup_J B(x_j, 3r_{x_j})$ with disjoint balls $B(x_j, r_{x_j})$, $j \in J$. We deduce

$$\mu(K) \leq \sum_{j \in J} \mu(B(x_j, 3r_{x_j})) = 3^n \sum_{j \in J} \mu(B(x_j, r_{x_j})) \leq \frac{3^n}{a} \sum_{j \in J} \int_{B(x_j, r_{x_j})} |f| \, \mathrm{d}\mu \leq \frac{3^n}{a} \int_{\mathbb{R}^n} |f| \, \mathrm{d}\mu.$$

Since (inner regularity) $\mu(E_a) = \sup\{\mu(K) \mid K \subset E_a \text{ compact}\}$, it concludes this step.

Step 4: Conclusion.

Define the operators $t_f(x, r) = \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(x) - f(y)| \, \mathrm{d}\mu(y)$ and $T_f(x) = \sup_{r>0} t_f(x, r)$. Suppose $f = g + h$ with $g \in L^1(\mathbb{R}^n)$ continuous and $h \in L^1(\mathbb{R}^n)$. Then

$$t_f(x, r) \leq t_g(x, r) + \frac{1}{\mu(B(x, r))} \int_{B(x,r)} |h(y)| \, \mathrm{d}\mu(y) + |h(x)|$$

and taking $r \to 0$ and using that all points are Lebesgue for a continuous function: $T_f(x) \leq M_h(x) + |h(x)|$. Therefore

$$\mu\left(\left\{x \mid T_f(x) > \frac{1}{k}\right\}\right) \leq \mu\left(\left\{x \mid M_h(x) > \frac{1}{2k}\right\}\right) + \mu\left(\left\{x \mid |h(x)| > \frac{1}{2k}\right\}\right) \leq 2k(3^n + 1)\|h\|_{L^1(\mathbb{R}^n)}$$

where we have used the weak integrability of M_h and Chebychef's inequality. The density of continuous functions in L^1 then mean the decomposition $f = g + h$ exists with $\|h\|_{L^1}$ as small as wanted, which implies $\mu(\{x \mid T_f(x) > \frac{1}{k}\}) = 0$ for all k non-zero integer, and so $\mu(\{x \mid T_f(x) \neq 0\}) = 0$. □

Let us explore further the links between integrability and differentiability, as emphasized by the name of the previous theorem.

Theorem 1.30. *Consider $f \in L^1(\mathbb{R})$ and define $F(x) = \int_{-\infty}^x f(y) \, \mathrm{d}\mu$, then F is differentiable almost everywhere with $F' = f$ (understood as an equality between elements of $L^1(\mathbb{R})$, i.e. almost everywhere).*

Proof. Observe that $\frac{F(x+\delta) - F(x)}{\delta} = \frac{1}{\mu([x, x+\delta]} \int_{[x,x+\delta]} f(y) \, \mathrm{d}\mu$. Then

$$\frac{1}{\mu([x, x+\delta])} \int_{[x,x+\delta]} |f(y) - f(x)| \, \mathrm{d}\mu \leq \frac{2}{\mu(B(x, \delta))} \int_{B(x,\delta)} |f(y) - f(x)| \, \mathrm{d}\mu$$

which goes to zero for almost every $x \in \mathbb{R}$ by the previous theorem. Hence for almost every $x \in \mathbb{R}$, one has $\frac{F(x+\delta) - F(x)}{\delta} \to f(x)$, which concludes the proof. □

Exercise 14. *Is the converse of this result true? More precisely if f differentiable almost everywhere with $f' \in L^1(\mathbb{R})$, do we always have $f(y) - f(x) = \int_x^y f'(z) \, \mathrm{d}\mu$?*

Remark 1.31. *One can prove that "being the integral of an L^1 function" (which* includes *assuming $f(y) - f(x) = \int_x^y f'(z) \, \mathrm{d}z$ as emphasized by the previous exercise) is equivalent to the* absolute continuity*: for any $\varepsilon > 0$ there is $\delta > 0$ s.t. for any finite collection of pairwise disjoint intervals (a_k, b_k), $k = 1, \dots, n$ with $\sum_{k=1}^n (b_k - a_k) \leq \delta$, then $\sum_{k=1}^n |f(b_k) - f(a_k)| \leq \varepsilon$.*

We finally turn to the link between pointwise and uniform convergence, and between measurability and continuity.

Theorem 1.32 (Egorov)**.** *Let $f_n : \mathbb{R}^d \to \mathbb{C}$ be a sequence of measurable functions, a Borel set A with finite measure and assume that f_k converges pointwise to a function f on A. Then for any $\varepsilon > 0$ there is a Borel set $A_\varepsilon \subset A$ with $\mu(A \setminus A_\varepsilon) \leq \varepsilon$ s.t. f_n converges uniformly to f on A.*

---

[1] For those interested in digging more on the web, this means that M_f is "weak L^1" with "weak L^1 norm" bounded in terms of the L^1 norm of f. This is often called Hardy-Littlewood maximal inequality.

Proof. Define for $k, n \geq 1$ the sets $E_n^{(k)} := \cap_{p \geq n}\{x \in A \mid |f_p(x) - f(x)| \leq \frac{1}{k}\}$, prove for any $k, n \geq 1$: $E_n^{(k)} \subset E_{n+1}^{(k)}$, $E_n^{(k+1)} \subset E_n^{(k)}$ and prove for any $k \geq 1$: $A = \cup_n E_n^{(k)}$. At k fixed, $A = \cup_n E_n^{(k)}$ with increasing union so (continuity from above of measure) there is n_k so that $\Delta_k := A \setminus E_{n_k}^{(k)}$ has measure $\mu(\Delta_k) \leq \varepsilon/2^k$. Define $\Delta = \cup_{k \geq 1}\Delta_k$ which has measure less than ε, and $A_\varepsilon := A \setminus \Delta$. On A_ε the convergence is uniform: for any k, $A_\varepsilon \subset E_{n_k}^{(k)}$ and thus $\sup_{x \in A_\varepsilon} |f_p(x) - f(x)| \leq 1/k$ for all $p \geq n_k$. □

Exercise 15. *Prove that the assumption $\mu(A) < +\infty$ is necessary for the result to hold.*

Theorem 1.33 (Lusin's Theorem - take 1). *If $f : \mathbb{R} \to \mathbb{C}$ is measurable and $\varepsilon > 0$, there is some Lebesgue measurable set $E \subset \mathbb{R}$ with $\mu(E) < \varepsilon$ so that $f_{|\mathbb{R}\setminus E}$ is continuous.*

Proof. Let us first present a proof which shows an interesting approximation by step functions.

Step 1. Given a measurable set $F \subset \mathbb{R}$ with $\mu(F) < \infty$ and $\varepsilon > 0$, there exists a finite union K of intervals with $\mu(F \Delta K) < \varepsilon$.

We have already proved it when studying the separability of $L^1(\mathbb{R})$.

Step 2. Given a simple function s on a measurable set F with finite measure and $\varepsilon > 0$, there is a step function S so that $\mu(\{x \in \mathbb{R} \mid S(x) \neq s(x)\}) < \varepsilon$.

Let $s = \sum_{k=1}^n \alpha_k \chi_{A_k}$ with $A_k \subset F$, let K_k be a finite union of intervals with $\mu(A_k \Delta K_k) < \varepsilon/n$, and let $S = \sum_{k=1}^n \alpha_k \chi_{K_k}$. Then $S(x) = s(x)$ on $F \setminus (\sum_{k=1}^n (A_k \Delta K_k))$, and $\mu(\sum_{k=1}^n (A_k \Delta K_k)) < \varepsilon$.

Step 3. If $f : F \to \mathbb{C}$ measurable with F measurable set with finite measure, there is a sequence of step functions converging to f almost everywhere.

Take a sequence of simple functions s_n that converges pointwise to f and for each n take S_n step function so that $S_n = s_n$ except on a set D_n with measure $\mu(D_n) < 2^{-n}$. Then $D := \cap_{N \geq 1} \cup_{n \geq N} D_n$ has zero measure and $S_n \to f$ on $F \setminus D$.

Step 4. Conclusion. Apply the previous step to each $F_\ell = [\ell, \ell + 1)$ to get sequences of step functions S_n^ℓ converging to f almost everywhere F_ℓ. Egorov's Theorem shows that there is $A_\ell \subset F_\ell$ with $\mu(A_\ell) \leq \frac{\varepsilon}{3 \cdot 2^{|\ell|}}$ s.t. the convergence $S_n^\ell \to f$ is uniform in $F_\ell \setminus A_\ell$. Let $A := \cup_{\ell \in \mathbb{Z}} A_\ell$, which has measure $\mu(A) \leq \varepsilon$. The set D of points of discontinuity of all step functions on all intervals is countable, hence $E := \mathbb{Z} \cup D \cup A$ has measure $\mu(E) \leq \varepsilon$. The function S_n concatenating all S_n^ℓ is continuous on $\mathbb{R} \setminus E$ and converges uniformly to f on $\mathbb{R} \setminus E$, which concludes the proof.

An alternative short proof. It is enough to consider $f : F \to \mathbb{R}$ with $\mu(F) < +\infty$, by applying on each F_ℓ and to the real or imaginary parts of f. Let $(V_n)_n$ be an enumeration of the open intervals with rational endpoints in $\mathbb{R}$. Fix compact sets $K_n \subset f^{-1}(V_n)$ and $K'_n \subset F \setminus f^{-1}(V_n)$ for each n so that $\mu(F \setminus (K_n \cup K'_n)) < \varepsilon/2^n$ (inner regularity). Fix open sets U_n s.t. $K_n \subset U_n$ and $U_n \cap K'_n = \emptyset$. Now, for $K := \cap_n (K_n \cup K'_n)$, $\mu(F \setminus K) < \varepsilon$. Given $x \in K$ and an n with $f(x) \in V_n$, $x \in K_n \subset U_n$ and $f(U_n \cap K) \subset V_n$. Since the V_n are a base of neighbourhoods, this shows that $f_{|K}$ is continuous. □

Remark 1.34. *This theorem does not claim that f is continuous at every $x \in \mathbb{R} \setminus E$. It is the restriction of f that is continuous. To illustrate the difference, consider $f = \chi_{\mathbb{Q}}$, which is nowhere continuous. However, its restriction to $\mathbb{R} \setminus \mathbb{Q}$ is continuous (constantly zero).*

Theorem 1.35 (Lusin's Theorem - take 2). *If $f : \mathbb{R} \to \mathbb{C}$ is measurable and $\varepsilon > 0$, there is some measurable set $G \subset \mathbb{R}$ with $\mu(G) < \varepsilon$ and a continuous function $g : \mathbb{R} \to \mathbb{C}$ so that $f = g$ on $\mathbb{R} \setminus G$.*

Proof. Apply the previous theorem: Take E with $\mu(E) < \varepsilon/2$ s.t. $f_{\mathbb{R}\setminus E}$ is continuous. Take (outer regularity) $G \subset \mathbb{R}$ open set that includes E and s.t. $\mu(G) \leq \varepsilon$. This set G is a pairwise disjoint countable union of open intervals $G = \cup_{k \geq 1}(a_k, b_k)$. Finally define g by $g := f$ on $\mathbb{R} \setminus G$, and

$$g(x) := f(a_k) + \frac{x - a_k}{b_k - a_k}\big(f(b_k) - f(a_k)\big) \quad \text{on} \quad (a_k, b_k).$$

□

Exercise 16. *The full power of Lusin's theorem is* not *required to prove that continuous functions are dense in $L^1(\mathbb{R})$. Give however a new proof of it by using Lusin's theorem.*