Analysis of Functions

Dr. Claude Warnick

February 23, 2021

Schedule

ANALYSIS OF FUNCTIONS (D) **24 lectures, Lent Term**

Part II Linear Analysis and Part II Probability and Measure are essential.

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]

Appropriate books

H. Brézis *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Universitext, Springer 2011

A.N. Kolmogorov, S.V. Fomin *Elements of the Theory of Functions and Functional Analysis.* Dover Books on Mathematics 1999

E.H. Lieb and M. Loss *Analysis.* Second edition, AMS 2001

Contents

Comments and Acknowledgements

These notes contain considerably more material than will be lectured or is contained in the schedules. The Appendices contain background material which may have been covered in the pre-requisite courses Linear Analysis and Measure Theory, but do not directly form part of the course. Some of the material on Distributions and the Fourier Transform formed part of a longer course I lectured on just these topics, and so goes into more detail than we will cover in this course. However, I have not drastically trimmed the notes, as some of the material not covered in lectures may nevertheless be interesting to students. There are many bonus exercises in these notes, marked by $(*)$, the example sheet exercises have their own numbers.

The textbooks I have found useful in preparing these notes include

- "Real Analysis", Stein and Shakarchi;
- "Measure Theory and Integration", Taylor;
- "Distributions, Theory and Applications", Duistermaat and Kolk;
- "Introduction to the theory of distributions", Friedlander and Joshi;
- "Fourier Analysis", Körner;
- "Functional Analysis", Rudin;
- "Topology", Munkres.

Please pass any corrections to `c.m.warnick@maths.cam.ac.uk`.

Chapter 1

Lebesgue Integration Theory

1.1 Introduction

This course can be thought of as "putting the *function* into functional analysis"[1]. The course *Linear Analysis* builds on earlier material in the Tripos to describe how a vector space structure (the linear part) interacts with a topological structure (the analysis part). This leads to a very beautiful abstract theory of Banach and Hilbert spaces, as well as other more general topological vector spaces. In this course, we shall see how many of these abstract results relate to more concrete spaces, in particular spaces of functions.

You are (hopefully) familiar from Part IB with the space $C^0([a,b])$ of continuous functions $f : [a,b] \to \mathbb{C}$, equipped with the norm:

$$\|f\|_{C^0} = \sup_{a \leqslant x \leqslant b} |f(x)| .$$

The completeness of this space follows from standard results concerning the uniform convergence of sequences of uniformly continuous functions, hence this is a Banach space. This space and its generalisations are important in many applications (for example in the proof of the Picard-Lindelöf Theorem, and the Schauder Theory for elliptic PDE).

Other spaces of functions naturally arise in many settings. For example, when studying Fourier series defined on $[a,b]$, it is natural to consider the space of continuous functions $f : [a,b] \to \mathbb{C}$ equipped with the norm:

$$\|f\|_{L^2} = \left(\int_a^b |f(x)|^2 \, dx \right)^{\frac{1}{2}} .$$

This norm comes from an inner product in a natural way. This space is not complete: we can construct a sequence of continuous functions f_i such that $(f_i)_{i\in\mathbb{N}}$ is Cauchy with respect to the L^2 norm, but for which there is no continuous function f such that $f_i \to f$ in L^2. One might hope to fix this by considering the space $\mathscr{R}([a,b])$ of Riemann integrable functions, however, we encounter two issues. Firstly, there are non-zero $f \in \mathscr{R}([a,b])$ such that $\|f\|_{L^2} = 0$, so $\|\cdot\|_{L^2}$ ceases to be a norm. In order to avoid this we can work instead

[1] More ambitiously, one could attempt to "put the *fun* into putting the *function* into functional analysis", but we do not aim so high.

with the space $R([a,b]) = \mathscr{R}([a,b])/\sim$ where we quotient by the equivalence relation $f \sim g$ if $\|f-g\|_{L^2} = 0$. The second issue is more serious: even working with the space $R([a,b])$, we do not find completeness with respect to the L^2 norm.

Exercise($*$). a) Find a sequence of continuous functions $f_i : [a,b] \to \mathbb{C}$ such that $(f_i)_{i\in\mathbb{N}}$ is Cauchy with respect to the L^2 norm, but for which there is no continuous function $f : [a,b] \to \mathbb{C}$ such that $f_i \to f$ in L^2.

b) Find a sequence $f_i \in R([a,b])$ such that $(f_i)_{i\in\mathbb{N}}$ is Cauchy with respect to the L^2 norm, but for which there is no $f \in R([a,b])$ such that $f_i \to f$ in L^2.

The solution to this problem is hopefully familiar to you. We should abandon the Riemann integral and work instead with the Lebesgue integral. This brings in our second pre-requisite course, *Probability and Measure*. The construction of the theory of measures and the Lebesgue integral is considerably more involved than that of the Riemann integral, however the pay-off is that the resulting theory of integration is much more powerful. In this course, we shall briefly review the theory of Lebesgue integration that you should have learned last term, before moving on to make use of measure theory, in combination with functional analysis, to understand various function spaces with importance in many branches of analysis.

1.2 Spaces of differentiable functions

Before reviewing integration, we briefly state some facts about the spaces of smooth functions. Let $\Omega \subset \mathbb{R}^n$ be an open set. We denote by $C^k(\Omega)$ the space of all k-times continuously differentiable complex valued functions on Ω, and by

$$C^\infty(\Omega) = \bigcap_{k=0}^{\infty} C^k(\Omega),$$

the set of smooth functions on Ω.

When dealing with partial derivatives of high orders, the notation can get rather messy. To mitigate this, it's convenient to introduce *multi-indices*. We define a multi-index α to be an element of $(\mathbb{Z}_{\geqslant 0})^n$, i.e. a $n-$vector of non-negative integers $\alpha = (\alpha_1, \ldots, \alpha_n)$. We define $|\alpha| = \alpha_1 + \ldots + \alpha_n$ and

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f,$$

in other words, we differentiate α_1 times with respect to x_1, α_2 times with respect to x_2 and so on. When it's unambiguous on which variables the derivative acts, we will also use the more compact notation:

$$D_i := \frac{\partial}{\partial x_i},$$

and

$$D^\alpha := \frac{\partial^{|\alpha|}}{\partial x^\alpha}.$$

For a vector $x \in \mathbb{R}^n$, we will also use the notation:

$$x^\alpha := (x_1)^{\alpha_1}(x_2)^{\alpha_2}\cdots(x_n)^{\alpha_n},$$

finally, we define

$$\alpha! = \alpha_1!\alpha_2!\cdots\alpha_n!.$$

The spaces $C^k(\Omega)$ and $C^\infty(\Omega)$ are vector spaces over $\mathbb{C}$, where addition and scalar multiplication are defined pointwise. If $\phi_1, \phi_2 \in C^k(\Omega)$ and $\lambda \in \mathbb{C}$, we define the maps $\phi_1 + \phi_2$, $\lambda\phi_1$ by

$$\begin{aligned} \phi_1 + \phi_2 \quad : \quad & \Omega \to \mathbb{C}, \\ & x \mapsto \phi_1(x) + \phi_2(x), \end{aligned} \qquad \begin{aligned} \lambda\phi_1 \quad : \quad & \Omega \to \mathbb{C}, \\ & x \mapsto \lambda\phi_1(x). \end{aligned} \tag{1.1}$$

> **Exercise(∗).** Show that with the definitions (1.1) the space $C^k(\Omega)$ is a vector space over $\mathbb{C}$, and that $C^l(\Omega)$ is a vector subspace of $C^k(\Omega)$ provided $k \leqslant l \leqslant \infty$.

Definition 1.1. *If $\phi \in C^0(\Omega)$, the* support *of ϕ is the set:*

$$\operatorname{supp}\phi = \overline{\{x \in \Omega : \phi(x) \neq 0\}},$$

where the closure is understood to be relative[2] *to Ω. That is $\operatorname{supp}\phi$ is the closure of the set on which ϕ is not zero. We say that ϕ has compact support if $\operatorname{supp}\phi$ is compact.*

For $0 \leqslant k \leqslant \infty$, we define $C^k_c(\Omega)$ to be the subset of $C^k(\Omega)$ consisting of functions with compact support. $C^k_c(\Omega)$ is a vector subspace of $C^k(\Omega)$.

Theorem 1.1. *There exists a function $\psi \in C^\infty_c(\mathbb{R}^n)$ such that*

i) $\psi \geqslant 0$

ii) $\psi(0) \neq 0$

iii) $\operatorname{supp}\psi \subset B_1(0) := \{x \in \mathbb{R}^n : |x| < 1\}$

iv) We have:

$$\int_{\mathbb{R}^n} \psi(x)dx = 1.$$

Proof. First, we note that the function:

$$\chi(t) = \begin{cases} 0 & t \leqslant 0 \\ e^{-\frac{1}{t}} & t > 0 \end{cases}$$

is smooth, i.e. $\chi \in C^\infty(\mathbb{R})$. Moreover, $\chi \geqslant 0$ and $\chi(1) \neq 0$. We define $\psi_0(x) = \chi\left(1 - 2|x|^2\right)$. Since the map $x \mapsto |x|^2$ is smooth, $\psi_0 \in C^\infty(\mathbb{R}^n)$. We set:

$$\psi(x) = \frac{\psi_0(x)}{\int_{\mathbb{R}^n}\psi_0(x)dx}.$$

It is easy to verify that ψ satisfies conditions $i) - iv)$. □

[2] If $\Omega \subset \mathbb{R}^n$ is open, and $A \subset \Omega$, then the closure of A relative to Ω is the intersection of Ω with the closure of A as a subset of $\mathbb{R}^n$. Note that the closure of A relative to Ω may not be closed as a subset of $\mathbb{R}^n$.

Corollary 1.2. *For $\Omega \subset \mathbb{R}^n$ open, $C_c^\infty(\Omega)$ is not the trivial subspace $\{0\} \subset C^\infty(\Omega)$.*

Proof. Since Ω is open, there exists ϵ, x such that the ball $B_\epsilon(x) = \{y \in \mathbb{R}^n : |y - x| < \epsilon\}$ is contained in Ω. The function $y \mapsto \psi\left[\epsilon^{-1}(y-x)\right]$ is easily seen to belong to $C_c^\infty(\Omega)$. □

Exercise(∗). Construct explicitly a function $\psi \in C_c^\infty(\mathbb{R}^n)$ such that

i) $0 \leqslant \psi \leqslant 1$

ii) $\operatorname{supp} \psi \subset B_2(0)$

iii) $\psi(x) = 1$ for $|x| \leqslant 1$.

Suppose $\Omega \subset \Omega'$, where both are open subsets of $\mathbb{R}^n$. If $\phi \in C_c^k(\Omega)$, then we can extend ϕ to a function on Ω' by setting $\phi = 0$ on $\Omega' \setminus \Omega$. This extended function will be smooth in Ω' and we do not alter the support, so in this way we see that $C_c^k(\Omega)$ is a vector subspace of $C_c^k(\Omega')$.

The following result is useful:

Lemma 1.3. *Suppose $\Omega \subset \mathbb{R}^n$ is open and $K \subset \Omega$ is compact. Then $d(K, \partial\Omega) > 0$, where:*

$$d(K, \partial\Omega) := \inf_{x \in K, y \in \partial\Omega} |x - y| .$$

Proof. K is compact, so $K \subset B_R(0)$ for some $R > 0$. Let $\Omega_R = \Omega \cap B_R(0)$. Ω_R is open and bounded, with $K \subset \Omega_R$. It suffices to show that $d(K, \partial\Omega_R) > 0$. Since Ω_R is bounded, $\partial\Omega_R$ is compact. Therefore the map:

$$\begin{aligned} f \quad : \quad & K \times \partial\Omega_R \to \mathbb{R}_{\geq 0}, \\ & (x, y) \mapsto |x - y| , \end{aligned}$$

is a continuous map on a compact set, hence it achieves its minimum d at (x_0, y_0). Suppose that $d = 0$, then $x_0 = y_0$, but $x_0 \in K \subset \Omega_R$ and $y_0 \in \partial\Omega_R \subset \Omega_R^c$ a contradiction. Thus $d > 0$ and we're done. □

Corollary 1.4. *If $\phi \in C_c^k(\Omega)$, extend ϕ to $\mathbb{R}^n$ by $\phi = 0$ on Ω^c. Define $\tau_x\phi$ by:*

$$\begin{aligned} \tau_x\phi \quad : \quad & \Omega \to \mathbb{C}, \\ & y \mapsto \phi(y - x). \end{aligned} \tag{1.2}$$

Then there exists $\epsilon > 0$ such that $\tau_x\phi \in C_c^k(\Omega)$ for all $x \in B_\epsilon(0)$.

Proof. We have

$$\operatorname{supp} \tau_x\phi = \operatorname{supp} \phi + x$$

Since $\operatorname{supp} \phi$ is compact, $\operatorname{supp} \tau_x\phi$ is just a translate of a compact set, so is compact as a subset of $\mathbb{R}^n$. We need to check that $\operatorname{supp} \tau_x\phi \subset \Omega$. We have $d(\operatorname{supp} \phi, \partial\Omega) = \delta > 0$. Set $\epsilon = \delta/2$. Then we have, by Lemma 1.3

$$\operatorname{supp} \phi + B_\epsilon(0) \subset \Omega$$

but if $x \in B_\epsilon(0)$, then $\operatorname{supp} \tau_x\phi \subset \operatorname{supp} \phi + B_\epsilon(0)$ and we're done. □

1.3 Review of integration

In this section we will briefly recall the main definitions and theorems of the theory of measure and Lebesgue integration. We shall focus on the Lebesgue measure on $\mathbb{R}^n$. Appendix B gives a much fuller account of the theory, and in particular contains proofs of the results claimed below.

We start with the basic definition of a measure space.

Definition 1.2. *Given a set E, a collection of subsets $\mathcal{E}$ of E is called a σ-algebra if:*

i) $\emptyset \in \mathcal{E}$

ii) $A \in \mathcal{E}$ *implies* $A^c = \{x \in E : x \notin A\} \in \mathcal{E}$

iii) $A_n \in \mathcal{E}$ *for* $n \in \mathbb{N}$ *implies* $\cup_n A_n \in \mathcal{E}$.

The pair $(E, \mathcal{E})$ *is called a* measurable space *and elements of* $\mathcal{E}$ *are called* measurable sets. *A* measure *on* $(E, \mathcal{E})$ *is a function* $\mu : \mathcal{E} \to [0, \infty]$ *such that:*

i) $\mu(\emptyset) = 0$

ii) If $A_n \in \mathcal{E}$ *for* $n \in \mathbb{N}$ *are disjoint, then* $\mu(\cup_n A_n) = \sum_n \mu(A_n)$.

A triple $(E, \mathcal{E}, \mu)$ *is called a measure space.*

A simple example of a measure space is given by taking $\mathcal{E} = 2^E$ and $\mu(A) = \#A$. This is the counting measure. Given any collection $\mathcal{A}$ of subsets of E, we can define $\sigma(\mathcal{A})$, the σ-algebra generated by $\mathcal{A}$ to be the intersection of all σ-algebras containing $\mathcal{A}$. If E is a topological space, the *Borel* algebra is the σ-algebra generated by the open sets of E, written $\mathcal{B}(E)$.

A particular case of interest is $E = \mathbb{R}^n$, on which we can define a σ-algebra $\mathcal{M}$, and measure λ with the following properties:

i) $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}$

ii) If A is a rectangle, i.e. $A = (a_1, b_1] \times \cdots \times (a_n, b_n]$, then $\lambda(A) = (b_1 - a_1) \cdots (b_n - a_n)$.

iii) $A \in \mathcal{M}$ if and only for any $\epsilon > 0$ there exists an open set O and a closed set C such that $C \subset A \subset O$ and
$$\lambda(O \setminus C) < \epsilon.$$

Since any open set in $\mathbb{R}^n$ is the countable union of disjoint rectangles these conditions determine $\mathcal{M}$, λ uniquely. We note that if $\lambda(A) < \infty$, then the set C in *iii*) above may be assumed to be compact. Property *iii*) is sometimes referred to as Borel regularity. We call $\mathcal{M}$ the σ-algebra of Lebesgue measurable sets, and λ is the Lebesgue measure. For the Lebesgue measure, we often denote $\lambda(A)$ by $|A|$, and μ by dx.

Definition 1.3. *A function* $f : E \to G$ *which maps between two measurable spaces* $(E, \mathcal{E})$, $(G, \mathcal{G})$ *is* measurable *if* $f^{-1}(A) \in \mathcal{E}$ *for all* $A \in \mathcal{G}$. *Special cases include:*

a) If $(G,\mathcal{G}) = (\mathbb{R},\mathcal{B}(\mathbb{R}))$*, we simply say* f *is a measurable function on* $(E,\mathcal{E})$*.*

b) If $(G,\mathcal{G}) = ([0,\infty],\mathcal{B}([0,\infty]))$ *we say* f *is a non-negative measurable function on* $(E,\mathcal{E})$*.*

c) If E, G *are topological spaces with their Borel algebras then we say* f *is a Borel function on* E*.*

The class of measurable functions is closed under vector space operations, products and limits.

A *simple function* is a function of the form

$$f = \sum_{k=1}^{N} a_k \mathbb{1}_{A_k}$$

for $A_k \in \mathcal{E}$ and a_k constant (typically in $[0,\infty]$, $\mathbb{R}$ or $\mathbb{C}$). All simple functions are measurable. For a non-negative simple function we define the integral

$$\mu(f) = \int_E f d\mu := \sum_{k=1}^{N} a_k \mu(A_k),$$

where $0 \cdot \infty = 0$ by convention. For a non-negative measurable function we define

$$\mu(f) = \int_E f d\mu := \sup\left\{\mu(g) : \ g \text{ simple and } 0 \leqslant g \leqslant f\right\}.$$

A measurable function is *integrable* if $\mu(|f|) < \infty$, in which case we can write $f = f^+ - f^-$ with $f^\pm$ non-negative and $\mu(f^\pm) < \infty$. Then

$$\mu(f) = \int_E f d\mu := \mu(f^+) - \mu(f^-).$$

The integral satisfies all the usual basic properties (linearity, additivity etc.), and agrees with the Riemann integral when both are defined. We can also state two important theorems for interchanging limits and integrals.

Theorem 1.5 (Monotone convergence)**.** *Let* $(f_n)_{n=1}^{\infty}$ *be an increasing sequence of non-negative measurable functions on a measure space* $(E,\mathcal{E},\mu)$ *which converge to* f*. Then*

$$\lim_{n\to\infty} \int_E f_n d\mu = \int_E f d\mu$$

Theorem 1.6 (Dominated convergence)**.** *Let* $(f_n)_{n=1}^{\infty}$ *be a sequence of measurable functions on a measure space* $(E,\mathcal{E},\mu)$ *such that*

i) $f_n \to f$ *pointwise ae.*

ii) $|f_n| \leqslant g$ *ae for some integrable* g*.*

Then:

$$\lim_{n\to\infty} \int_E f_n d\mu = \int_E f d\mu$$

Associated to each measure space $(E, \mathcal{E}, \mu)$ are scale of Banach spaces.

Definition 1.4. *For $1 \leqslant p < \infty$, and $f : E \to \mathbb{C}$ measurable, we define:*

$$\|f\|_{L^p} = \left(\int_E |f|^p \, d\mu \right)^{\frac{1}{p}}.$$

while for $p = \infty$ we set

$$\|f\|_{L^\infty} = \operatorname*{ess\,sup}_E |f| = \inf\{C : |f| \leqslant C \ \ ae\}.$$

The space $L^p(E, \mu)$ is then defined to be

$$L^p(E, \mu) = \{f : E \to \mathbb{C} \ \textit{measurable} \ : \|f\|_{L^p} < \infty\} / \sim$$

where we quotient by the equivalence relation $f \sim g$ if $f = g$ ae.

If E is a topological space and $\mathcal{B}(E) \subset \mathcal{E}$, we define $L^p_{loc.}(E, \mu)$ to consist of measurable functions (modulo $\sim$) such that $f \mathbb{1}_K \in L^p(E, \mu)$ for all compact K.

When our measure space is $(\mathbb{R}^n, \mathcal{M}, \lambda)$ we will often write $L^p(\mathbb{R}^n) := L^p(\mathbb{R}^n, \lambda)$.

Theorem 1.7. *The space $L^p(E, \mu)$ equipped with the norm $\|\cdot\|_{L^p}$ is a Banach space for $1 \leqslant p \leqslant \infty$.*

It is useful also to note that the set S of complex valued simple functions on E such that

$$\mu(\{x : s(x) \neq 0\}) < \infty$$

is dense in $L^p(E, \mu)$ for $1 \leqslant p < \infty$.

Exercise 1.1. Suppose $f, g : E \to \mathbb{C}$ are measurable functions on some measure space $(E, \mathcal{E}, \mu)$. Show that:

a) $\|fg\|_{L^r} \leqslant \|f\|_{L^p} \|g\|_{L^q}$ where $1 \leqslant p, q, r \leqslant \infty$ satisfy $p^{-1} + q^{-1} = r^{-1}$
[You may wish to first establish the special case $r = 1$.]

b) $\|f + g\|_{L^p} \leqslant \|f\|_{L^p} + \|g\|_{L^p}$ for $1 \leqslant p \leqslant \infty$.

Exercise 1.2. a) Suppose that $\mu(E) < \infty$. Show that if $f \in L^p(E, \mu)$, then $f \in L^q(E, \mu)$ for any $1 \leqslant q \leqslant p$, with

$$\|f\|_{L^q} \leqslant \mu(E)^{\frac{p-q}{qp}} \|f\|_{L^p}.$$

b) Suppose that $f \in L^{p_0}(E,\mu) \cap L^{p_1}(E,\mu)$ with $p_0 < p_1 \leqslant \infty$. For $0 \leqslant \theta \leqslant 1$, define p_θ by

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

Show that $f \in L^{p_\theta}(E,\mu)$ with

$$\|f\|_{L^{p_\theta}} \leqslant \|f\|_{L^{p_0}}^{1-\theta} \|f\|_{L^{p_1}}^{\theta}.$$

c) Show that for $p_1 \neq p_2$ we have $L^{p_1}(\mathbb{R}^n) \not\subset L^{p_2}(\mathbb{R}^n)$. For which p_1, p_2 do we have $L^{p_1}_{loc.}(\mathbb{R}^n) \subset L^{p_2}_{loc.}(\mathbb{R}^n)$?

Exercise 1.3. Let $\mathcal{R}_\mathbb{Q}$ be the set of rectangles of the form $(a_1, b_1] \times \cdots \times (a_n, b_n]$ with $a_i, b_i \in \mathbb{Q}$, and let $S_\mathbb{Q}$ be the set of functions of the form

$$s(x) = \sum_{k=1}^{N} (\alpha_k + i\beta_k) \mathbb{1}_{R_k}$$

for $R_k \in \mathcal{R}_\mathbb{Q}$ and $\alpha_k, \beta_k \in \mathbb{Q}$. For $1 \leqslant p < \infty$ show that $S_\mathbb{Q}$ is dense in $L^p(\mathbb{R}^n)$ and deduce that $L^p(\mathbb{R}^n)$ is separable. Show that $L^\infty(\mathbb{R}^n)$ is not separable. *[Hint: for the last part exhibit an uncountable subset $X \subset L^\infty(\mathbb{R}^n)$ such that $\|f - g\|_{L^\infty(\mathbb{R}^n)} \geqslant 1$ for any $f, g \in X$, $f \neq g$]* .

1.4 Convolution and mollification

In this section, we are going to establish some results concerning mollification of functions in $L^p(\mathbb{R}^n)$. The final result will be to establish the density of $C_c^\infty(\mathbb{R}^n)$ in $L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$. We first establish an important fact about the spaces $L^p(\mathbb{R}^n)$: namely that the translation operator is continuous on these spaces. More concretely, for any $z \in \mathbb{R}^n$ we set $\tau_z f(x) = f(x - z)$. We then show:

Lemma 1.8. *Suppose $p \in [1, \infty)$ and $g \in L^p(\mathbb{R}^n)$. Let $\{z_j\}_{j=1}^\infty \subset \mathbb{R}^n$ be a sequence of points such that $z_j \to 0$ as $j \to \infty$. Then:*

$$\left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)} \to 0.$$

Proof. 1. First, suppose $g = \mathbb{1}_R$, where $R = (a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n]$ is a rectangle, with side-lengths $I_m = b_m - a_m$ for $m = 1, \ldots, n$. Now, since when a box is translated by a vector z_j each side is translated by a distance of at most $|z_j|$, and has area at most $I_{max.}^{n-1}$, where I_{max} is the longest side-length we can crudely estimate

$$\left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)}^p \leqslant 2n\, |z_j|\, I_{\max}^{n-1}.$$

Note that this estimate requires $p < \infty$: it does not hold for $p = \infty$. We conclude that:

$$\lim_{j \to \infty} \left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)} = 0.$$

2. Now suppose $g = \mathbb{1}_A$, where A is a measurable set of finite measure. Fix $\epsilon > 0$. By the Borel regularity of Lebesgue measure, there exists a compact $K \subset A$ and an open $U \supset A$ such that $|U \setminus K| < \epsilon^p$. Since U is open, we can write U as a union of open rectangles:
$$U = \bigcup_{\alpha \in \mathscr{A}} R_\alpha^\circ$$
Since K is compact, it is covered by a finite subset of these:
$$K \subset \bigcup_{i=1}^N R_i := B.$$
Now, note that $K \subset B \subset U$, so the symmetric difference $A \Delta B \subset U \setminus K$. Thus[3] $\|\mathbb{1}_A - \mathbb{1}_B\|_{L^p(\mathbb{R}^n)} = |A \Delta B|^{1/p} < \epsilon$. By the paragraph 1 above, we know that there exists J such that for all $j \geqslant J$ we have:
$$\left\|\tau_{z_j} \mathbb{1}_B - \mathbb{1}_B\right\|_{L^p(\mathbb{R}^n)} < \epsilon$$
Therefore:
$$\begin{aligned}\left\|\tau_{z_j} \mathbb{1}_A - \mathbb{1}_A\right\|_{L^p(\mathbb{R}^n)} &= \left\|\tau_{z_j} \mathbb{1}_A - \tau_{z_j} \mathbb{1}_B + \tau_{z_j} \mathbb{1}_B - \mathbb{1}_B + \mathbb{1}_B - \mathbb{1}_A\right\|_{L^p(\mathbb{R}^n)} \\ &\leqslant \left\|\tau_{z_j} \mathbb{1}_A - \tau_{z_j} \mathbb{1}_B\right\|_{L^p(\mathbb{R}^n)} + \left\|\tau_{z_j} \mathbb{1}_B - \mathbb{1}_B\right\|_{L^p(\mathbb{R}^n)} + \|\mathbb{1}_B - \mathbb{1}_A\|_{L^p(\mathbb{R}^n)} \\ &= 2\|\mathbb{1}_A - \mathbb{1}_B\|_{L^p(\mathbb{R}^n)} + \left\|\tau_{z_j} \mathbb{1}_B - \mathbb{1}_B\right\|_{L^p(\mathbb{R}^n)} \\ &< 3\epsilon\end{aligned}$$
for all $j \geqslant J$. Thus
$$\lim_{j\to\infty} \left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)} = 0.$$

3. Now suppose g is a simple function, $g = \sum_{i=1}^N g_i \mathbb{1}_{A_i}$, where $g_i \in \mathbb{C}$ and A_i are measurable sets of *finite* measure. Then we have:
$$\left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)} \leqslant \sum_{i=1}^N |g_i| \left\|\tau_{z_j} \mathbb{1}_{A_i} - \mathbb{1}_{A_i}\right\|_{L^p(\mathbb{R}^n)}$$
so as $j \to \infty$ we have:
$$\lim_{j\to\infty} \left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)} = 0.$$

4. Now suppose that $g \in L^p(\mathbb{R}^n)$. Fix $\epsilon > 0$. Recall that there exists a simple function $\tilde{g} = \sum_{i=1}^n \tilde{g}_i \mathbb{1}_{A_i}$ with $\tilde{g}_i \in \mathbb{C}$, $|A_i| < \infty$ such that $\|g - \tilde{g}\|_{L^p(\mathbb{R}^n)} < \epsilon$. By the previous part, we can find J such that $\left\|\tau_{z_j}\tilde{g} - \tilde{g}\right\|_{L^p(\mathbb{R}^n)} < \epsilon$ for all $j \geqslant J$. Now:
$$\begin{aligned}\left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)} &= \left\|\tau_{z_j} g - \tau_{z_j}\tilde{g} + \tau_{z_j}\tilde{g} - \tilde{g} + \tilde{g} - g\right\|_{L^p(\mathbb{R}^n)} \\ &\leqslant \left\|\tau_{z_j} g - \tau_{z_j}\tilde{g}\right\|_{L^p(\mathbb{R}^n)} + \left\|\tau_{z_j}\tilde{g} - \tilde{g}\right\|_{L^p(\mathbb{R}^n)} + \|\tilde{g} - g\|_{L^p(\mathbb{R}^n)} \\ &= 2\|g - \tilde{g}\|_{L^p(\mathbb{R}^n)} + \left\|\tau_{z_j}\tilde{g} - g\right\|_{L^p(\mathbb{R}^n)} \\ &< 3\epsilon\end{aligned}$$

[3] This is another point at which $p \neq \infty$ is crucial.

Thus, we conclude that

$$\lim_{j\to\infty} \left\|\tau_{z_j} g - g\right\|_{L^p(\mathbb{R}^n)} = 0.$$

and we're done. □

If f, g are functions mapping $\mathbb{R}^n$ to $\mathbb{C}$, then we define the convolution of f and g to be:

$$(f \star g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy,$$

provided the integral exists. This will happen if (for example) $f \in L^1(\mathbb{R}^n)$ and $g \in L^\infty(\mathbb{R}_n)$.

Lemma 1.9. *Suppose $f, g, h \in C_c^\infty(\mathbb{R}^n)$. Then:*

$$f \star g = g \star f, \qquad f \star (g \star h) = (f \star g) \star h.$$

and

$$\int_{\mathbb{R}^n} (f \star g)(x)dx = \int_{\mathbb{R}^n} f(x)dx \int_{\mathbb{R}^n} g(x)dx.$$

Proof. With the change of variables $y = x - z$, we have[4]

$$(f \star g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy = \int_{\mathbb{R}^n} f(x-z)g(z)dz = (g \star f)(x)$$

Next, we calculate:

$$\begin{aligned}
\left[f \star (g \star h)\right](x) &= \int_{\mathbb{R}^n} f(y)\left(\int_{\mathbb{R}^n} g(z)h(x-y-z)dz\right)dy \\
&= \int_{\mathbb{R}^n} f(y)\left(\int_{\mathbb{R}^n} g(w-y)h(x-w)dw\right)dy \\
&= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(y)g(w-y)dy\right)h(x-w)dw \\
&= \left[(f \star g) \star h\right](x)
\end{aligned}$$

Above we have made the substitution $w = y + z$ to pass from the first to second line, and we have used the fact that $f, g, h \in C_c^\infty(\mathbb{R}^n)$ to invoke Fubini's theorem (Theorem B.30)

[4] If you're worried about a missing minus sign from the change of variables when n is odd, observe:

$$\int_{-\infty}^{\infty} k(x)dx = \int_{\infty}^{-\infty} k(-y)d(-y) = -\int_{\infty}^{-\infty} k(-y)dy = \int_{-\infty}^{\infty} k(-y)dy.$$

when passing from the second to third line. Finally, we calculate:

$$\begin{aligned}\int_{\mathbb{R}^n} (f \star g)(x)dx &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(y)g(x-y)dy \right) dx \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(y)g(x-y)dx \right) dy \\ &= \int_{\mathbb{R}^n} \left(f(y) \int_{\mathbb{R}^n} g(z)dz \right) dy \\ &= \int_{\mathbb{R}^n} f(x)dx \int_{\mathbb{R}^n} g(z)dy.\end{aligned}$$

where again, the fact that $f, g \in C_c^\infty(\mathbb{R}^n)$ allows us to invoke Fubini. □

The assumption that the functions are smooth and compactly supported is certainly overkill in this theorem. It would be enough, for example, to consider functions in $C_c^0(\mathbb{R}^n)$, or even weaker spaces, provided we can justify the application of Fubini's theorem.

Exercise(∗). Suppose that $f, g, h \in C_c^\infty(\mathbb{R}^n)$.

a) Show that for any multi-index α, we have that $D^\alpha f \in L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$, i.e. that

$$\|D^\alpha f\|_{L^p(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n} |D^\alpha f(x)|^p \, dx \right)^{\frac{1}{p}} < \infty.$$

b) Define

$$\begin{aligned} F \quad : \quad & \mathbb{R}^n \times \mathbb{R}^n, \\ & (x, y) \mapsto f(x)g(y-x). \end{aligned}$$

Show that $F \in L^1(\mathbb{R}^n \times \mathbb{R}^n)$.

c) For each $x \in \mathbb{R}^n$, set

$$\begin{aligned} G_x \quad : \quad & \mathbb{R}^n \times \mathbb{R}^n, \\ & (y, z) \mapsto f(y)g(z)h(x-y-z). \end{aligned}$$

Show that $G_x \in L^1(\mathbb{R}^n \times \mathbb{R}^n)$.

1.4.1 Differentiating convolutions

A remarkable property of the convolution is that the regularity of $f \star g$ is determined by the regularity of the *smoother* of f and g. This is a result of the following Lemma:

Theorem 1.10. *Suppose $f \in L^1_{loc.}(\mathbb{R}^n)$ and $g \in C_c^k(\mathbb{R}^n)$ for some $k \geqslant 0$. Then $f \star g \in C^k(\mathbb{R}^n)$ and*

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

for any multiindex with $|\alpha| \leqslant k$.

Before we prove this, it's convenient to prove a technical Lemma which will streamline the proof. We introduce the difference quotient

$$\Delta_i^h f(x) = \frac{f(x+he_i) - f(x)}{h}.$$

Lemma 1.11. *a) Suppose $f \in C^0_c(\mathbb{R}^n)$ and $\{z_i\}_{i=1}^\infty \subset \mathbb{R}^n$ is a sequence with $z_i \to 0$ as $i \to \infty$. Then for any $x \in \mathbb{R}^n$:*

i) $\tau_{z_j} f(x) \to f(x)$ as $j \to \infty$.

ii) $\left|\tau_{z_j} f(x)\right| \leqslant (\sup_{\mathbb{R}^n} |f|)\, \mathbb{1}_{B_R(0)}(x)$, for some $R > 0$ and all j.

b) Suppose $f \in C^1_c(\mathbb{R}^n)$ and $\{h_j\}_{j=1}^\infty \subset \mathbb{R}$ is a sequence with $h_j \to 0$ as $j \to \infty$. Then for any $x \in \mathbb{R}^n$:

i) $\Delta_i^{h_j} f(x) \to D_i f(x)$ as $j \to \infty$.

ii) $\left|\Delta_i^{h_j} f(x)\right| \leqslant (\sup_{\mathbb{R}^n} |D_i f|)\, \mathbb{1}_{B_R(0)}(x)$, for some $R > 0$ and all j.

Proof. a) i) Recall $\tau_{z_j} f(x) = f(x - z_j)$. Clearly since $z_j \to 0$, $f(x - z_j) \to f(x)$ as $j \to \infty$ by the continuity of f.

ii) Since $z_j \to 0$, there exists some $\rho > 0$ such that $z_j \in \overline{B_\rho(0)}$ for all j. Now

$$\operatorname{supp} \tau_{z_j} f = \operatorname{supp} f + z_j \subset \operatorname{supp} f + \overline{B_\rho(0)}.$$

Since the sum of two bounded set is bounded, we conclude that there exists $R > 0$ such that $\operatorname{supp} \tau_{z_j} f \subset B_R(0)$. Thus $\tau_{z_j} f = \tau_{z_j} f \mathbb{1}_{B_R(0)}$ and we estimate:

$$\left|\tau_{z_j} f(x)\right| = \left|\tau_{z_j} f(x)\right| \mathbb{1}_{B_R(0)}(x) \leqslant \sup_{\mathbb{R}^n} |f|\, \mathbb{1}_{B_R(0)}(x).$$

b) Suppose $f \in C^1_c(\mathbb{R}^n)$ and $\{h_j\}_{j=1}^\infty \subset \mathbb{R}$ is a sequence with $h_j \to 0$ as $j \to \infty$. Then for any $x \in \mathbb{R}^n$:

i) From the definition of the difference quotient and of the partial derivative:

$$\Delta_i^{h_j} f(x) = \frac{f(x+h_j e_i) - f(x)}{h} \to D_i f(x), \qquad \text{as } j \to \infty.$$

ii) Since $h_j \to 0$, there is some $\rho > 0$ such that $|h_j| \leqslant \rho$ for all j. We have:

$$\begin{aligned} \operatorname{supp} \Delta_i^{h_j} f &\subset \operatorname{supp} \tau_{-h_j e_i} f \cup \operatorname{supp} f = (\operatorname{supp} f - h_j e_i) \cup \operatorname{supp} f \\ &\subset \left(\operatorname{supp} f + \overline{B_\rho(0)}\right) \cup \operatorname{supp} f \\ &\subset B_R(0) \end{aligned}$$

for some $R > 0$ since the union of two bounded sets is bounded. Thus $\Delta_i^{h_j} f = \Delta_i^{h_j} f \mathbb{1}_{B_R(0)}$. We also observe that by the mean value theorem, for any $h \in \mathbb{R}$, there exists $s \in \mathbb{R}$ with $|s| < |h|$ such that

$$\frac{f(x+he_i) - f(x)}{h} = D_i f(x + s e_i)$$

thus

$$\left|\Delta_i^{h_j} f(x)\right| \leqslant \sup_{\mathbb{R}^n} |D_i f|.$$

Putting these two facts together, we readily find:

$$\left|\Delta_i^{h_j} f(x)\right| = \left|\Delta_i^{h_j} f(x)\right| \mathbb{1}_{B_R(0)}(x) \leqslant \sup_{\mathbb{R}^n} |D_i f| \, \mathbb{1}_{B_R(0)}(x).$$

□

Now, with this technical result in hand we can attack the proof our original theorem.

Proof of Theorem 1.10. 1. First we establish the result for $k = 0$. We need to show that if $f \in L^1_{loc.}(\mathbb{R}^n)$ and $g \in C^0_c(\mathbb{R}^n)$ then $f \star g$ is continuous. To show this, it suffices to show that $f \star g(x - z_j) \to f \star g(x)$ for any $x \in \mathbb{R}^n$ and any sequence $\{z_j\}_{j=1}^\infty$ with $z_j \to 0$. Now, note that

$$f \star g(x - z_j) = \int_{\mathbb{R}^n} f(y) g(x - z_j - y) dy = \int_{\mathbb{R}^n} f(y) \tau_{z_j} g(x - y) dy.$$

Now, sending $j \to \infty$, we are done, so long as we can justify interchanging the limit and the integral. Note that for any fixed x and all j:

$$\left|f(y) \tau_{z_j} g(x - y)\right| \leqslant \sup_{\mathbb{R}^n} |g| \, \mathbb{1}_{B_R(0)}(x - y) \, |f(y)|$$

for some R by the previous Lemma. Since $f \in L^1_{loc.}(\mathbb{R}^n)$ the right hand side is integrable, and so by the dominated convergence theorem:

$$\lim_{j\to\infty} f \star g(x - z_j) = \int_{\mathbb{R}^n} \lim_{j\to\infty} f(y) \tau_{z_j} g(x - y) dy = \int_{\mathbb{R}^n} f(y) g(x - y) dy = f \star g(x).$$

2. Now suppose that $f \in L^1_{loc.}(\mathbb{R}^n)$ and $g \in C^1_c(\mathbb{R}^n)$. Clearly $f \star D_i g$ is continuous by the previous argument. To show $f \star g \in C^1(\mathbb{R}^n)$, it suffices to show that for any $x \in \mathbb{R}^n$ and any sequence $\{h_j\}_{j=1}^\infty \subset \mathbb{R}$ with $h_j \to 0$ we have:

$$\lim_{j\to\infty} \Delta_i^{h_j} f \star g(x) = f \star D_i g(x).$$

Note that

$$\begin{aligned}
\Delta_i^{h_j} f \star g(x) &= \frac{f \star g(x + h_j e_i) - f \star g(x)}{h} \\
&= \int_{\mathbb{R}^n} f(y) \left(\frac{g(x + h_j e_i - y) - g(x - y)}{h} \right) dy \\
&= \int_{\mathbb{R}^n} f(y) \Delta_i^{h_j} g(x - y) dy
\end{aligned}$$

so that again we are done provided we can send $j \to \infty$ and interchange the limit and the integral. An argument precisely analogous to the previous case allows us to invoke the DCT and deduce that:

$$\lim_{j\to\infty} \Delta_i^{h_j} f \star g(x) = \int_{\mathbb{R}^n} \lim_{j\to\infty} f(y) \Delta_i^{h_j} g(x - y) dy = f \star D_i g(x).$$

3. The case where $g \in C_c^k(\mathbb{R}^n)$ with $k > 1$ now follows by a simple induction. □

Exercise(∗). Show that $f \star g \in C^k(\mathbb{R}^n)$ under the hypotheses:

a) $f \in L^1(\mathbb{R}^n)$, $g \in C^k(\mathbb{R}^n)$ with $\sup_{\mathbb{R}^n} |D^\alpha g| < \infty$ for all $|\alpha| \leqslant k$.

b) $f \in L^1(\mathbb{R}^n)$ with $\operatorname{supp} f$ compact, $g \in C^k(\mathbb{R}^n)$.

We have shown that when two functions are convolved, loosely speaking the resulting function is at least as regular as the *better* of the two original functions. It is also important to know how convolution modifies the support of a function.

Lemma 1.12. *Suppose $f \in L^1_{loc.}(\mathbb{R}^n)$ and $g \in C_c^k(\mathbb{R}^n)$ for some $k \geqslant 0$. Then*[5]

$$\operatorname{supp}(f \star g) \subset \operatorname{supp} f + \operatorname{supp} g.$$

Proof. Recall:

$$f \star g(x) = \int_{\mathbb{R}^n} f(y) g(x-y) dy.$$

Clearly, if $f \star g(x) \neq 0$, then there must exist $y \in \mathbb{R}^n$ such that $y \in \operatorname{supp} f$ and $x - y = z \in \operatorname{supp} g$. Thus $x = y + z$ with $y \in \operatorname{supp} f$ and $z \in \operatorname{supp} g$. This tells us that:

$$\{x \in \mathbb{R}^n : f \star g(x) \neq 0\} \subset \operatorname{supp} f + \operatorname{supp} g.$$

Since $\operatorname{supp} f$ is closed and $\operatorname{supp} g$ is compact, we know that $\operatorname{supp} f + \operatorname{supp} g$ is closed, thus

$$\operatorname{supp} f \star g = \overline{\{x \in \mathbb{R}^n : f \star g(x) \neq 0\}} \subset \operatorname{supp} f + \operatorname{supp} g,$$

which is the result we require. □

Exercise(∗). a) Prove the following identities for $r, s > 0$ and $x \in \mathbb{R}^n$:

i) $B_r(x) + B_s(0) = B_{r+s}(x)$
ii) $\overline{B_r(x)} + B_s(0) = B_{r+s}(x)$
iii) $\overline{B_r(x)} + \overline{B_s(0)} = \overline{B_{r+s}(x)}$

Suppose that $A, B \subset \mathbb{R}^n$. Show that:

b) If one of A or B is open, then so is $A + B$.

c) If A and B are both bounded, then so is $A + B$.

d) If A is closed and B is compact, then $A + B$ is closed.

e) If A and B are both compact, then so is $A + B$.

Exercise(∗). Show that if $f \in C_c^k(\mathbb{R}^n)$ and $g \in C_c^l(\mathbb{R}^n)$ then $f \star g \in C_c^{k+l}(\mathbb{R}^n)$. Conclude that $C_c^\infty(\mathbb{R}^n)$ is closed under convolution.

[5] Strictly speaking, we haven't defined the support of a measurable function. We can do this in several ways, but the simplest is to define:

$$\operatorname{supp} f = \bigcap \{E \subset \mathbb{R}^n : E \text{ is closed, and } f = 0 \text{ a.e. on } E^c\}.$$

In other words $\operatorname{supp} f$ is the smallest closed set such that f vanishes almost everywhere on its complement.

1.4.2 Approximation of the identity

An important use of the convolution is to construct smooth approximations to functions in various function spaces. The following theorem is very useful in constructing approximations:

Theorem 1.13. *Suppose $\phi \in C_c^\infty(\mathbb{R}^n)$ satisfies:*

i) $\phi \geqslant 0$

ii) $\operatorname{supp}\phi \subset B_1(0)$

iii) $\int_{\mathbb{R}^n} \phi(x)dx = 1$

Such a ϕ exists by Theorem 1.1. Define:

$$\phi_\epsilon(y) = \frac{1}{\epsilon^n}\phi\left(\frac{y}{\epsilon}\right).$$

Then:

a) If $f \in C_c^k(\mathbb{R}^n)$, then $\phi_\epsilon \star f$ is smooth, and

$$D^\alpha\left(\phi_\epsilon \star f\right) \to D^\alpha f \qquad \text{as } \epsilon \to 0,$$

uniformly on $\mathbb{R}^n$ for any multi-index with $|\alpha| \leqslant k$.

b) If $g \in L^p(\mathbb{R}^n)$ with $1 \leqslant p < \infty$, then $\phi_\epsilon \star g$ is smooth, and

$$\phi_\epsilon \star g \to g \qquad \text{in } L^p(\mathbb{R}^n) \qquad \text{as } \epsilon \to 0.$$

c) Suppose $f \in C^k(\mathbb{R}^n)$ with $\sup_{\mathbb{R}^n} |D^\alpha f| < \infty$ for $|\alpha| \leqslant k$, and suppose $g \in L^1(\mathbb{R}^n)$ with $g \geqslant 0$, $\int_{\mathbb{R}^n} g(x)dx = 1$. Set $g_\epsilon(y) = \epsilon^{-n} g\left(\epsilon^{-1}y\right)$. Then $f \star g_\epsilon \in C^k(\mathbb{R}^n)$, and

$$D^\alpha\left(f \star g_\epsilon\right)(x) \to D^\alpha f(x) \qquad \text{as } \epsilon \to 0,$$

for any $x \in \mathbb{R}^n$ and any multi-index with $|\alpha| \leqslant k$.

Proof. a) Note that the rescaling of ϕ to produce ϕ_ϵ is such that a change of variables gives:

$$\int_{\mathbb{R}^n} \phi_\epsilon(y)dy = 1.$$

By Theorem 1.10, we have that $D^\alpha(\phi_\epsilon \star f) = \phi_\epsilon \star D^\alpha f$ for any $|\alpha| \leqslant k$. Using these two facts, we calculate:

$$\begin{aligned}
D^\alpha(\phi_\epsilon \star f)(x) - D^\alpha f(x) &= \int_{\mathbb{R}^n} \phi_\epsilon(y) D^\alpha f(x-y)dy - D^\alpha f(x)\int_{\mathbb{R}^n} \phi_\epsilon(y)dy \\
&= \int_{\mathbb{R}^n} \phi_\epsilon(y)\left[D^\alpha f(x-y) - D^\alpha f(x)\right]dy \\
&= \int_{B_1(0)} \phi(z)\left[D^\alpha f(x-\epsilon z) - D^\alpha f(x)\right]dz
\end{aligned}$$

where in the last line we made the substitution $y = \epsilon z$, and noted that ϕ has support in $B_1(0)$, so we can restrict the range of integration. Now, since $\phi \geqslant 0$, we can estimate:

$$\begin{aligned} |D^\alpha(\phi_\epsilon \star f)(x) - D^\alpha f(x)| &\leqslant \int_{B_1(0)} \phi(z)\, |D^\alpha f(x - \epsilon z) - D^\alpha f(x)|\, dz \\ &\leqslant \sup_{z \in B_1(0)} |D^\alpha f(x - \epsilon z) - D^\alpha f(x)| \times \int_{B_1(0)} \phi(z) dz \\ &= \sup_{z \in B_1(0)} |D^\alpha f(x - \epsilon z) - D^\alpha f(x)| \end{aligned}$$

since $\int_{\mathbb{R}^n} \phi = 1$. Now, since $D^\alpha f$ is continuous and of compact support, it is uniformly continuous on $\mathbb{R}^n$. Fix $\tilde{\epsilon} > 0$. There exists δ such that for any $v, w \in \mathbb{R}^n$ with $|x - y| < \delta$, we have

$$|D^\alpha f(v) - D^\alpha f(w)| < \tilde{\epsilon}$$

For any $x \in \mathbb{R}^n$, taking $\epsilon < \delta$, and $v = x - \epsilon z$, $w = x$ with $z \in B_1(0)$ we have $|v - w| < \delta$, so:

$$|D^\alpha f(x - \epsilon z) - D^\alpha f(x)| < \tilde{\epsilon}$$

holds for any $x \in \mathbb{R}^n, z \in B_1(0)$. We have therefore shown that for any $\tilde{\epsilon} > 0$, there exists δ such that for any $\epsilon < \delta$ we have:

$$\sup_{x \in \mathbb{R}^n} |D^\alpha(\phi_\epsilon \star f)(x) - D^\alpha f(x)| < \tilde{\epsilon}.$$

This is the statement of uniform convergence on $\mathbb{R}^n$.

b) Noting that $L^p(\mathbb{R}^n) \subset L^1_{loc.}(\mathbb{R}^n)$ by an application of Hölder's inequality (see Exercise 1.2), Theorem 1.10 immediately establishes the smoothness of $\phi_\epsilon \star g$. To establish convergence as $\epsilon \to 0$, we shall require certain measure theoretic results. First we require Minkowski's Integral Identity (see Exercise 1.4). This states[6] that for $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ a measurable function, we have the estimate:

$$\left[\int_{\mathbb{R}^n} \left|\int_{\mathbb{R}^n} F(x, y) dx\right|^p dy\right]^{\frac{1}{p}} \leqslant \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |F(x, y)|^p\, dy\right]^{\frac{1}{p}} dx$$

Now, following the calculation in the previous proof, we readily have that:

$$|(\phi_\epsilon \star g)(x) - g(x)| \leqslant \int_{\mathbb{R}^n} \phi(z)\, |g(x - \epsilon z) - g(x)|\, dz$$

[6] There is more general statement for a map $F : X \times Y \to \mathbb{C}$, which is measurable with respect to the product measure $\mu \times \nu$ where (X, μ) and (Y, ν) are measure spaces.

Integrating and applying Minkowski's integral inequality, we have:

$$\begin{aligned}
\|\phi_\epsilon \star g - g\|_{L^p(\mathbb{R}^n)} &= \left[\int_{\mathbb{R}^n} |(\phi_\epsilon \star g)(x) - g(x)|^p \, dx\right]^{\frac{1}{p}} \\
&\leqslant \left[\int_{\mathbb{R}^n} \left|\int_{\mathbb{R}^n} \phi(z) \, |g(x-\epsilon z) - g(x)| \, dz\right|^p dx\right]^{\frac{1}{p}} \\
&\leqslant \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \phi(z)^p \, |g(x-\epsilon z) - g(x)|^p \, dx\right]^{\frac{1}{p}} dz \\
&= \int_{\mathbb{R}^n} \phi(z) \, \|\tau_{\epsilon z} g - g\|_{L^p(\mathbb{R}^n)} \, dz \qquad (1.3)
\end{aligned}$$

To establish our result it will suffice to set $\epsilon = \epsilon_j$, where $\{\epsilon_j\}_{j=1}^\infty \subset \mathbb{R}$ is any sequence with $\epsilon_j \to 0$, and show that $\left\|\phi_{\epsilon_j} \star g - g\right\|_{L^p(\mathbb{R}^n)} \to 0$. Note that since $\left\|\tau_{\epsilon_j z} g\right\|_{L^p(\mathbb{R}^n)} = \|g\|_{L^p(\mathbb{R}^n)}$ we have:

$$\phi(z) \left\|\tau_{\epsilon_j z} g - g\right\|_{L^p(\mathbb{R}^n)} \leqslant 2\phi(z) \, \|g\|_{L^p(\mathbb{R}^n)}$$

so the integrand is dominated uniformly in j by an integrable function. Now by Lemma 1.8, as y varies, $\tau_y : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ is a continuous family of bounded linear operators. This means that for each $z \in \mathbb{R}^n$ we have:

$$\lim_{j\to\infty} \left\|\tau_{\epsilon_j z} g - g\right\|_{L^p(\mathbb{R}^n)} = 0.$$

Thus we can apply the Dominated Convergence Theorem (Theorem B.26) to the integral on the right hand side of 1.3, and conclude that

$$\lim_{j\to\infty} \left\|\phi_{\epsilon_j} \star g - g\right\|_{L^p(\mathbb{R}^n)} = 0.$$

c) Again, by Theorem 1.10, we have that $D^\alpha(f \star g_\epsilon) = D^\alpha f \star g_\epsilon$ for any $|\alpha| \leqslant k$. By a change of variables, we calculate:

$$D^\alpha(f \star g_\epsilon)(x) = \int_{\mathbb{R}^n} g_\epsilon(y) D^\alpha f(x-y) dy = \int_{\mathbb{R}^n} g(z) D^\alpha f(x - \epsilon z) dz$$

Now, clearly for each fixed $x \in \mathbb{R}^n$:

$$g(z) D^\alpha f(x - \epsilon z) \to g(z) D^\alpha f(x)$$

for $z \in \mathbb{R}^n$ as $\epsilon \to 0$. Furthermore,

$$|g(z) D^\alpha f(x - \epsilon z)| \leqslant g(z) \sup_{\mathbb{R}^n} |D^\alpha f|$$

which is an integrable function of z, so by the Dominated convergence theorem, we conclude:

$$D^\alpha(f \star g_\epsilon)(x) \to D^\alpha f(x) \int_{\mathbb{R}^n} g(z) dz = D^\alpha f(x)$$

as $\epsilon \to 0$.

□

The final application of convolutions is to the construction of cut-off functions. These are often extremely useful for localising a problem to a particular region of interest for some reason or other.

Lemma 1.14. *Suppose $\Omega \subset \mathbb{R}^n$ is open, and $K \subset \Omega$ is compact. Then there exists $\chi \in C_c^\infty(\Omega)$ such that $\chi = 1$ in a neighbourhood of K.*

Proof. By Lemma 1.3, there exists $\epsilon > 0$ such that $d(K, \partial\Omega) > 4\epsilon$. We define $K_\epsilon = K + \overline{B_{2\epsilon}(0)}$. As the sum of two compact sets, K_ϵ is compact. Moreover, $K_\epsilon \subset \Omega$. Suppose ϕ_ϵ is as in Theorem 1.13. Consider:

$$\chi := \phi_\epsilon \star \mathbb{1}_{K_\epsilon}.$$

We have by Theorem 1.10 that $\chi \in C^\infty(\mathbb{R}^n)$ and from Lemma 1.12 we deduce:

$$\operatorname{supp}\chi = K_\epsilon + \operatorname{supp}\phi_\epsilon \subset K + \overline{B_{2\epsilon}(0)} + \overline{B_\epsilon(0)} = K + \overline{B_{3\epsilon}(0)} \subset \Omega.$$

Thus $\chi \in C_c^\infty(\Omega)$. Now, suppose $x \in K + B_\epsilon(0)$. Then $x + B_\epsilon(0) \subset K_\epsilon$ and so:

$$\begin{aligned}\chi(x) &= \int_{\mathbb{R}^n} \phi_\epsilon(y)\mathbb{1}_{K_\epsilon}(x-y)dy \\ &= \int_{B_\epsilon(0)} \phi_\epsilon(y)\mathbb{1}_{K_\epsilon}(x-y)dy \\ &= \int_{B_\epsilon(0)} \phi_\epsilon(y)dy = 1.\end{aligned}$$

Thus $\chi(x) = 1$ for $x \in K + B_\epsilon(0)$, which is a neighbourhood of K. □

The following exercise establishes results required for the proof of Theorem 1.13.

Exercise 1.4. a) Suppose $1 \leqslant p \leqslant \infty$ and let q satisfy $p^{-1} + q^{-1} = 1$. Show that for a measurable function $f : \mathbb{R}^n \to \mathbb{C}$:

$$\|f\|_{L^p} = \sup\left\{\int_{\mathbb{R}^n} |f(x)g(x)|\, dx : g \in L^q(\mathbb{R}^n), \|g\|_{L^q} \leqslant 1\right\}.$$

b) Now suppose $p < \infty$ and assume $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is integrable. Set $G(y) = \int_{\mathbb{R}^n} F(x,y)dx$. Show that if $\|g\|_{L^q} \leqslant 1$ then

$$\int_{\mathbb{R}^n} |G(y)g(y)|\, dy \leqslant \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |F(x,y)|^p\, dy\right]^{\frac{1}{p}} dx.$$

Deduce Minkowski's integral inequality

$$\left[\int_{\mathbb{R}^n} \left|\int_{\mathbb{R}^n} F(x,y)dx\right|^p dy\right]^{\frac{1}{p}} \leqslant \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |F(x,y)|^p\, dy\right]^{\frac{1}{p}} dx.$$

Exercise 1.5. Let $I = (0,1)$ and $1 \leq p < \infty$. Exhibit a sequence $(f_j)_{j=1}^\infty$ with $f_j \in L^p(I)$ such that $f_j \to 0$ in $L^p(I)$, but $f_j(x)$ does not converge for any x. Does such a sequence exist if $p = \infty$?

1.5 Lebesgue differentiation theorem

The fundamental theorem of calculus is a fundamental result in the theory of Riemann integration. It comes in two (related) flavours.

Theorem 1.15 (Fundamental Theorem of Calculus)**.** *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous, and define the function $F : [a, b] \to \mathbb{R}$ by:*

$$F(x) = \int_a^x f(t)dt.$$

Then F is differentiable on (a, b), and:

$$F'(x) = f(x).$$

From this one can deduce the alternative form of the Fundamental Theorem of Calculus, relating the integral of a function to its anti-derivative.

Corollary 1.16. *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and that $F : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b) and satisfies $F'(x) = f(x)$ for all $x \in (a, b)$. Then:*

$$\int_a^b f(t)dt = F(b) - F(a).$$

We seek to generalise Theorem 1.15 in the setting of the Lebesgue integral. First, we note that the result implies

$$\begin{aligned} f(x) &= \lim_{r\to 0} \frac{F(x+r) - F(x-r)}{2r} = \lim_{r\to 0} \frac{1}{2r} \int_{x-r}^{x+r} f(t)dt \\ &= \lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(t)dt \end{aligned}$$

where we have used that the ball of radius r about x is simply $B_r(x) = (x - r, x + r)$ in one dimension. Rearranging, we can further conclude

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} (f(t) - f(x))\, dt = 0.$$

This statement is meaningful in dimensions higher that one. In fact we shall prove something slightly stronger

Theorem 1.17 (Lebesgue differentiation theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{C}$ be integrable. Then for almost every $x \in \mathbb{R}^n$ we have*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy = 0. \tag{1.4}$$

Note that it suffices for f to be defined on an open set $\Omega \subset \mathbb{R}^n$, and we obtain the same differentiability result at almost every $x \in \Omega$ by considering $f\mathbb{1}_\Omega$. We say that a point x such that (1.4) holds is a *Lebesgue point* of f. In order to establish this result, we first introduce a related quantity for which we are able to prove an estimate.

Definition 1.5. *Given an integrable function $f : \mathbb{R}^n \to \mathbb{C}$, the Hardy–Littlewood Maximal function Mf is defined to be*

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\, dy.$$

The Hardy–Littlewood Maximal function and its generalisations are of use in many contexts in mathematics, in particular in harmonic anaysis. For our purposes, a key result is that it satisfies a *weak* L^1-bound.

Lemma 1.18 (Weak L^1-bound for Mf)**.** *Suppose $f \in L^1(\mathbb{R}^n)$ for $n \geqslant 1$. Then Mf is measurable, finite almost everywhere, and there exists a constant C_n, depending only on n such that:*

$$|\{x : Mf(x) > \lambda\}| \leqslant \frac{C_n}{\lambda} \|f\|_{L^1} \tag{1.5}$$

for all $\lambda > 0$.

Proof. Let $A_\lambda = \{x : Mf(x) > \lambda\}$. Then for each $x \in A_\lambda$ there exists a radius r_x such that

$$\frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)|\, dy > \lambda. \tag{1.6}$$

We claim A_λ is open, which implies Mf is measurable. To see this, suppose $x \in A_\lambda$ with corresponding r_x and let $(x_m)_{m=1}^\infty$ be a sequence with $x_m \to x$ and $x_m \notin A_\lambda$. Then by the dominated convergence theorem we have

$$\frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x_m)} |f(y)|\, dy \to \frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)|\, dy,$$

however, by assumption

$$\frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x_m)} |f(y)|\, dy \leqslant \lambda, \qquad \frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)|\, dy > \lambda,$$

a contradiction.

Fix $K \subset A_\lambda$ a compact set. Since K is covered by $\cup_{x \in A_\lambda} B_{r_x}(x)$, we can pick a finite subcover of K, say $K \subset \cup_{i=1}^N B_i$, where $B_i = B_{r_x}(x)$ for some x. By Wiener's covering Lemma (see Exercise 1.8) there is a disjoint subcollection $B_{i_1}, \ldots, B_{i_k}$ such that

$$|K| \leqslant \left| \bigcup_{i=1}^N B_i \right| \leqslant 3^n \sum_{j=1}^n \left| B_{i_j} \right|.$$

Now, each B_{i_j} satisfies (1.6), so we have

$$|K| \leqslant \frac{3^n}{\lambda} \sum_{j=1}^n \int_{B_{i_j}} |f(y)|\, dy \leqslant \frac{3^n}{\lambda} \|f\|_{L^1}$$

Where in the final inequality we use that the B_{i_j} are disjoint. Since this holds for all compact $K \subset A_\lambda$, (1.5) follows. Finally, note that $\{Mf = \infty\} \subset \{Mf > \lambda\}$, which implies $|\{Mf = \infty\}| < C/\lambda$ for all λ, thus $|\{Mf = \infty\}| = 0$. □

With this result in hand, we are now ready to establish the Lebesgue differentiation theorem.

Proof of Theorem 1.17. For each $\lambda > 0$ define:

$$A_\lambda = \left\{ x \in \mathbb{R}^n : \limsup_{r \to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy > 2\lambda \right\}$$

If we can show that $|A_\lambda| = 0$ for all λ, then we will be done, as the set of $x \in \mathbb{R}^n$ which are not Lebesgue points for f is precisely $\cup_{n=1}^{\infty} A_{\frac{1}{n}}$.

Fix $\epsilon > 0$. We can find $g \in C_c^\infty(\mathbb{R}^n)$ such that $\|f - g\|_{L^1} < \epsilon$. We estimate:

$$\begin{aligned} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy &\leqslant \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - g(y)|\, dy \\ &\quad + \frac{1}{|B_r(x)|} \int_{B_r(x)} |g(y) - g(x)|\, dy + |g(x) - f(x)| \end{aligned}$$

We can bound the first term by

$$\begin{aligned} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - g(y)|\, dy &\leqslant \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - g(y)|\, dy \\ &= M[f-g](x) \end{aligned}$$

Now, since g is continuous, we have

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

hence

$$\limsup_{r \to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy \leqslant M[f-g](x) + |f(x) - g(x)|\,.$$

Now, if $x \in A_\lambda$, then we must either have $M[f-g](x) > \lambda$ or $|f(x) - g(x)| > \lambda$. By Lemma 1.18 we know

$$|\{x : M[f-g](x) > \lambda\}| \leqslant \frac{C_n}{\lambda} \|f - g\|_{L^1} \leqslant \frac{C_n \epsilon}{\lambda},$$

and by Tchebychev's inequality we know

$$|\{x : |f(x) - g(x)| > \lambda\}| \leqslant \frac{\|f - g\|_{L^1}}{\lambda} \leqslant \frac{\epsilon}{\lambda}$$

we conclude that

$$|A_\lambda| \leqslant \frac{1 + C_n}{\lambda} \epsilon.$$

Since ϵ was arbitrary, we conclude $|A_\lambda| = 0$, and we're done. □

Exercise 1.6. Suppose $1 \leqslant p < \infty$.

a) Suppose $f \in L^p(\mathbb{R}^n)$. Show that

$$|\{x : |f(x)| > \lambda\}| \leqslant \frac{\|f\|_{L^p}^p}{\lambda^p}.$$

This is known as Tchebychev's inequality, the $p = 1$ case is Markov's inequality.

b) We say that a measurable $f : \mathbb{R}^n \to \mathbb{C}$ is in *weak-$L^p(\mathbb{R}^n)$*, written $f \in L^{p,w}(\mathbb{R}^n)$ if there exists a constant C such that

$$|\{x : |f(x)| > \lambda\}| \leqslant \frac{C^p}{\lambda^p}.$$

Show that $L^p(\mathbb{R}^n) \subset L^{p,w}(\mathbb{R}^n)$, and that the inclusion is proper.

Exercise 1.7. Suppose that $f \in L^r(\mathbb{R}^n)$ for some $1 \leqslant r < \infty$. Show that $\|f\|_{L^\infty} = \lim_{p\to\infty} \|f\|_{L^p}$.
[Hint: you may find the estimates in Exercises 1.2 b), 1.6 a) useful.]

Exercise 1.8. a) Let $B_1, \ldots, B_N$ be a finite collection of open balls in $\mathbb{R}^n$. Show that there exists a subcollection $B_{i_1}, \ldots, B_{i_k}$ of *disjoint* balls such that

$$\bigcup_{i=1}^{N} B_i \subset \bigcup_{j=1}^{k} (3B_{i_j}),$$

where $3B$ is the ball with the same centre as B but three times the radius. Deduce

$$\left|\bigcup_{i=1}^{N} B_i\right| \leqslant 3^n \sum_{j=1}^{k} \left|B_{i_j}\right|.$$

b) (*) Suppose $\{B_j : j \in J\}$ is an arbitrary collection of balls in $\mathbb{R}^n$ such that each ball has radius at most R. Show that there exists a countable subcollection $\{B_j : j \in J'\}$, $J' \subset J$ of disjoint balls such that

$$\bigcup_{i\in J} B_i \subset \bigcup_{i\in J'} (5B_i).$$

These are Wiener and Vitali's covering Lemmas, respectively.

Exercise 1.9. Suppose $f : \mathbb{R} \to \mathbb{C}$ is integrable and let $F(x) = \int_{-\infty}^{x} f(t)dt$. Show that F is differentiable with $F'(x) = f(x)$ at each Lebesgue point $x \in \mathbb{R}$. Deduce that F is differentiable almost everywhere.

Exercise 1.10. Suppose $\phi \in L^\infty(\mathbb{R}^n)$ satisfies $\phi \geqslant 0$, supp $\phi \subset B_1(0)$, and $\int_{\mathbb{R}^n} \phi\, dx = 1$. Set $\phi_\epsilon(x) = \epsilon^{-n}\phi(\epsilon^{-1}x)$. Show that if $f \in L^1(\mathbb{R}^n)$, and x is a Lebesgue point of f,

$$\phi_\epsilon \star f(x) \to f(x), \qquad \text{as } \epsilon \to 0.$$

1.6 Littlewood's principles: Egorov's Theorem and Lusin's Theorem

In his 1944 "Lectures on the Theory of Functions", J. E. 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 of these results may be stated more precisely in our language as follows:

Lemma 1.19. *Suppose* $|A| < \infty$*. Then for any* $\epsilon > 0$ *there exists a set* B*, which is a finite union of rectangles, such that*

$$|A \Delta B| < \epsilon.$$

This follows straightforwardly from the basic properties of Lebesgue measurable sets. The third of Littlewood's principles follows from

Theorem 1.20 (Egorov's Theorem)**.** *Suppose* $(f_k)_{k=1}^{\infty}$ *is a sequence of functions defined on a set* $E \subset \mathbb{R}^n$ *with* $|E| < \infty$*, and suppose that* $f_k \to f$ *almost everywhere on* E*. Then given* $\epsilon > 0$ *we can find a closed set* $A_\epsilon \subset E$ *such that* $|E - A_\epsilon| \leqslant \epsilon$ *and* $f_k \to f$ uniformly *on* A_ϵ*.*

Proof. By discarding a set of measure zero if necessary, we can assume without loss of generality that $f_k(x) \to f(x)$ for all $x \in E$. For each $n, k \in \mathbb{N}$ let

$$E_k^n = \left\{ x \in E : |f_j(x) - f(x)| < \frac{1}{n}, \text{ for all } j > k \right\}.$$

Fixing n, we note that $E_k^n \subset E_{k+1}^n$ and that $\cup_{k=1}^{\infty} E_k^n = E$. By countable additivity, we have $|E_k^n|$ is an increasing sequence, with $|E_k^n| \to |E|$ as $k \to \infty$. Pick k_n such that $\left|E \setminus E_{k_n}^n\right| < 2^{-n}$. By construction we have:

$$|f_j(x) - f(x)| < \frac{1}{n}, \qquad \text{for all } j > k_n \text{ and } x \in E_{k_n}^n.$$

Now pick N such that $\sum_{n=N}^{\infty} 2^{-n} < \epsilon/2$ and let

$$A'_\epsilon = \bigcap_{n=N}^{\infty} E_{k_n}^n.$$

Now we observe

$$\left|E \setminus A'_\epsilon\right| \leqslant \sum_{n=N}^{\infty} \left|E \setminus E_{k_n}^n\right| < \epsilon/2.$$

Next. suppose that $\delta > 0$. Pick $n \geqslant N$ such that $1/n < \delta$, and note that $x \in A'_\epsilon$ implies $x \in E_{k_n}^n$. We deduce that $|f_j(x) - f(x)| < \delta$ for all $j > k_n$ and hence $f_j \to f$ uniformly on A'_ϵ. Finally, we can pick a closed set $A_\epsilon \subset A'_\epsilon$ such that $|A'_\epsilon \setminus A_\epsilon| < \epsilon/2$ and hence $|E - A_\epsilon| \leqslant \epsilon$. □

The final of Littlewood's principles is given flesh by

Theorem 1.21 (Lusin's Theorem)**.** *Suppose f is measurable and finite valued on E, where $E \subset \mathbb{R}^n$ with $|E| < \infty$. Then given $\epsilon > 0$ we can find a closed set $F_\epsilon \subset E$ with $|E \setminus F_\epsilon| < \epsilon$ such that $f|_{F_\epsilon}$ is continuous.*

Proof. Suppose first f is a simple function

$$f = \sum_{k=1}^{m} a_k \mathbb{1}_{A_k},$$

where $|A_k| < \infty$ and the A_k are disjoint with $E = \cup_{k=1}^m A_k$ (if necessary, we add the term $0\mathbb{1}_{f^{-1}(0)}$ to arrange this). For any $\epsilon > 0$, we can pick compact sets $K_k \subset A_k$ with

$$|A_k \setminus K_k| < \frac{\epsilon}{m}.$$

Let $B = \cup_{k=1}^m K_k$. Then $|E \setminus B| < \epsilon$. Since the sets K_k are compact and disjoint (hence $\min_{i,j} \operatorname{dist}(K_i, K_j) > 0$), and f is constant on each K_k, we have that f is continuous on B.

Now let f_n be a sequence of simple functions such that $f_n \to f$ ae. Then we can find C_n such that $|C| < 2^{-n}$ and f_n is continuous outside C_n. By Egorov's theorem, we can find a set $A_{\epsilon/3}$ such that $f_n \to f$ uniformly on $A_{\epsilon/3}$ and $\left|E \setminus A_{\epsilon/3}\right| < \epsilon/3$. Let N be sufficiently large that $\sum_{n=N}^{\infty} 2^{-n} < \epsilon/3$.

$$F'_\epsilon = A_{\epsilon/3} \setminus \bigcup_{n=N}^{\infty} C_n$$

Now, $|E \setminus F'_\epsilon| < 2\epsilon/3$ and moreover, for $n > N$ the functions f_n are continuous on F'_ϵ, so since they converge uniformly to f, we have that f is continuous on F'_ϵ. Finally, picking $F_\epsilon \subset F'_\epsilon$ closed with $|F'_\epsilon \setminus F_\epsilon| < \epsilon/3$ we're done. □

Remark. *Note that Lusin's Theorem asserts that $f|_{F_\epsilon}$ is continuous, which means that f is continuous if we think of it as defined only at points of F_ϵ. This is not the same as the statement that f (defined on E) is continuous at points of F_ϵ. For example if $f = \mathbb{1}_{\mathbb{Q}}$, then $f|_{\mathbb{R}\setminus\mathbb{Q}} = 0$ is continuous, however f is nowhere continuous.*