Part A: Integration

Alison Etheridge

1 Introduction

In Mods you learned how to integrate step functions and continuous functions on closed bounded intervals. We begin by recalling some of that theory.

Definition 1.1 (Step functions and their integrals)**.** *A function ϕ defined on the interval I (with endpoints a, b) is a* step function *if there is a partition $(a_i)_{i=0}^k$ of the interval (that is $a = a_0 < a_1 < a_2 < \ldots < a_k = b$) so that ϕ is constant on each interval (a_{j-1}, a_j), $j \leq k$.*

If $\phi(x) = c_i$ for $x \in (a_{i-1}, a_i)$, then

$$\int_I \phi \equiv \int_I \phi(x)dx = \sum_{i=1}^{k} c_i(a_i - a_{i-1}).$$

Write $L^{\text{step}}[a, b]$ for the space of step functions on $[a, b]$.

Definition 1.2 (Lower and Upper Sums)**.** *For a real-valued function on $[a, b]$ define*

$$\underline{\int_a^b} f = \sup\left\{\int_a^b \phi : \phi \in L^{\text{step}}[a, b], \phi \leq f\right\}, \quad \overline{\int_a^b} f = \inf\left\{\int_a^b \phi : \phi \in L^{\text{step}}[a, b], \phi \geq f\right\}$$

The fundamental result that you proved was the following.

Theorem 1.3 (Integrating continuous functions over closed bounded intervals)**.** *For a continuous function f on the closed bounded interval $[a, b]$*

$$\underline{\int_a^b} f = \overline{\int_a^b} f. \tag{1}$$

The quantity in equation (1) is then defined to be the integral of f over $[a, b]$. We write $C[a, b]$ for continuous functions on $[a, b]$.

Later in the course you considered sequences of continuous functions:

Theorem 1.4 (Exchanging integrals and limits under *uniform* convergence)**.** *Let $\{f_n\}_{n\geq 1}$ be a sequence of functions in $C[a, b]$ for which $f_n \to f$* uniformly *(and so in particular $f \in C[a, b]$) then*

$$\int_a^b f_n \to \int_a^b f \quad \text{as } n \to \infty.$$

You then exploited this to interchange summation and integration for power series: Suppose that the power series $f(t) = \sum_{k=0}^{\infty} a_k t^k$ has radius of convergence R. Write $f_n(t) = \sum_{k=0}^{n} a_k t^k$. Since for $|x| < R$, $f_n(t) \to f(t)$ uniformly on $[0, x]$ you deduced that

$$\begin{aligned} \int_0^x f(t)dt &= \lim_{n\to\infty} \int_0^x f_n(t)dt \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} a_k \frac{x^{k+1}}{k+1} \\ &= \sum_{k=0}^{\infty} a_k \frac{x^{k+1}}{k+1}, \end{aligned}$$

and this proved to be a powerful tool in establishing deep properties of exponential and trigonometric functions.

Originally integration was introduced as the inverse of differentiation. The goal of Newton (1642–1727) and Leibniz (1646 – 1716) was to solve the *problem of primitives*:

Find the functions $F(x)$ which have derivative a given function $f(x)$.

At that time the notion of function was not really well defined, but usually it meant associating a quantity y to a variable x through an equation involving arithmetic operations (addition, subtraction, multiplication, division, taking roots), trigonometric operations (sin, cos, tan and their inverses) and logarithms and exponentials.

With the geometric understanding of integration as area under a curve came the generalisation of this idea of function, but still *continuous* functions were viewed as the 'real' functions and the notion of integration from Mods sufficed. The version of the Fundamental Theorem of Calculus that you proved in Mods showed that for continuous functions f on closed bounded intervals, the problem of primitives was solved by the indefinite integral of f. But then along came Fourier (1768–1830). He showed that trigonometric series which can be used to represent continuous functions on bounded intervals can also be used to represent *discontinuous* ones. In particular, the function

$$f(x) = \begin{cases} 0 & 0 \le x \le \pi, \\ 1 & \pi < x \le 2\pi, \end{cases}$$

can be represented by a convergent trigonometric series. The notion of function had to be extended and with it the notion of integral.

For a finite number of discontinuities, one can piece together the portions on which the function is continuous and your methods easily extend - and in fact you can even go a bit further provided the points of discontinuity are 'exceptional' in a sense that we'll make precise later. But worse was to come. Dirichlet (1805–1859) came across the following function which now takes his name:

Definition 1.5 (Dirichlet function)**.** *The function*

$$\chi(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}, \end{cases} \tag{2}$$

is known as the Dirichlet function*.*

In fact Dirichlet obtained χ as a limit:

$$\chi(x) = \lim_{m\to\infty} \lim_{n\to\infty} (\cos(m!\pi x))^{2n}.$$

The Dirichlet function is *nowhere* continuous and yet Dirichlet constructed it as a double pointwise limit of a sequence of continuous functions.

Evidently our theory of integration from Mods really *cannot* handle the Dirichlet function. Any step function ϕ on $[0,1]$ satsifying $\phi \le \chi$ also satisfies $\phi \le 0$ on $[0,1]$ and the constant step function 0 on $[0,1]$ satisfies $0 \le \chi$ and so $\underline{\int_0^1}\chi(x)dx = 0$. Likewise, the constant function step function 1 on $[0,1]$ satisfies $1 \ge \chi$ and any step function ϕ on $[0,1]$ satisfying $\chi \le \phi$ also satisfies $\phi \ge 1$ on $[0,1]$, so $\overline{\int_0^1}\chi(x)dx = 1$. Thus

$$0 = \underline{\int_0^1}\chi(x)dx < \overline{\int_0^1}\chi(x)dx = 1.$$

Of course there are many other less perverse ways than Dirichlet's to obtain the Dirichlet function as a pointwise limit of functions that are integrable in the sense of Mods.

Example 1.6. *Let $\{r_n\}_{n\ge1}$ be an enumeration of the rationals in $[0,1]$. For each $k \in \mathbb{N}$ define $f_k : [0,1] \to \mathbb{R}$ by*

$$f_k(x) = \begin{cases} 1 & x = r_n, 1 \le n \le k, \\ 0 & \textit{otherwise.} \end{cases}$$

Then $\int f_k(x)dx = 0$ for each k, but $\int \lim_{k\to\infty} f_k(x)dx$ is not (yet) defined.

The Lebesgue integral, introduced by Lebesgue in a very short paper of 1901 but fully explained in a beautiful set of lecture notes published in 1904 (from a course delivered in 1902-3) is an *extension* of the integral that you developed in Mods that behaves well under passage to the limit.

So when will it be the case that $f_n \to f$ implies $\int f_n \to \int f$ in our new theory? The next example shows that the answer should certainly be *not always*.

Example 1.7. *Define $f_k : [0,1] \to \mathbb{R}$ by*

$$f_k(x) = \begin{cases} k(1-kx) & 0 \le x \le 1/k, \\ 0 & \textit{otherwise.} \end{cases}$$

Each f_k is continuous and $f_k(x) \to 0$ as $k \to \infty$ for all $x \in (0,1]$. So $f_k \to f \equiv 0$. But $\int_0^1 f_k(x)dx = 1/2$ for all k. So

$$0 = \int_0^1 \lim_{k\to\infty} f_k(x)dx < \lim_{k\to\infty}\int_0^1 f_k(x)dx = \frac{1}{2}.$$

The difficulty from the point of view of the integration theory of Mods is that the convergence is *not uniform*. It will turn out that we'll get convergence theorems for our new integral under other conditions. For example, suppose that we restrict ourselves to functions with $|f_k(x)| \le M$ for all k and all $x \in [0,1]$. A theorem with this hypothesis is called a *bounded convergence theorem*. Or, because we don't want to restrict ourselves to bounded functions on bounded intervals, we could try taking a monotone sequence $f_k \uparrow f$ (by which we mean that $f_k \le f_{k+1}$ for all k and $f_k \to f$) or $f_k \downarrow f$. A theorem with this hypothesis is called a *monotone convergence theorem*. Example 1.6 shows that neither of these conditions is enough to rescue the integral of Mods, but they will help us out here. The functions of Example 1.7 don't satisfy either of our hypotheses and we agreed that for them we would not *expect* a convergence theorem. It was natural that the integral of the limit was strictly less than the limit of the integrals. We'll see that for positive functions an inequality a bit like this always holds. The limit of the integrals might not exist and so we replace it with a new concept that we'll define later called the lim inf. When the limit does exist then it is equal to the lim inf. Probably the most important result of our theory, *Fatou's Lemma*, will tell us that for *positive* functions f_k converging to f

$$\int \lim_{k\to\infty} f_k(x)dx \leq \liminf_{k\to\infty} \int f_k(x)dx.$$

So what should the integral of the Dirichlet function of Definition 1.5 over $[0,1]$ be?

The Dirichlet function is just the *indicator* function of the rationals. For an interval I we define the integral of its indicator function over $[0,1]$ to be the length of $I \cap [0,1]$. Similarly for a set S which is a finite union of disjoint intervals, the integral over $[0,1]$ of the indicator function of S is the total length of the intervals making up $S \cap [0,1]$. For an interval $[a,b]$ we'll write $m([a,b]) = b - a$ for its length. It would make sense then for the integral of the Dirichlet function over $[0,1]$ to be the 'length' of the set $\mathbb{Q} \cap [0,1]$, $m(\mathbb{Q} \cap [0,1])$. So now we have to work out how to assign a length to this set.

For two sets $A \subseteq B$ our intuition says that the length of A should be less than or equal to that of B, that is $m(A) \leq m(B)$. Now, given $\epsilon > 0$, let $\{r_n\}_{n\geq 1}$ be an enumeration of the rationals in $[0,1]$ and let

$$I_n = (r_n - \frac{\epsilon}{2^{n+1}}, r_n + \frac{\epsilon}{2^{n+1}}),$$

then $\mathbb{Q} \cap [0,1] \subseteq \bigcup_{n=1}^{\infty} I_n$ and (again using a bit of intuition) the total length of the right hand side is less than or equal to $\sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon$. Since ϵ was arbitrary we conclude that in a reasonable world we'd better take the length of $\mathbb{Q} \cap [0,1]$ to be zero. We'll make that more precise later, but then with this notion of length,

$$\int_0^1 \chi(x)dx = 0$$

and our problem with the functions $\{f_k\}_{k\geq 1}$ of Example 1.6 goes away,

$$0 = \int_0^1 \chi(x)dx = \int_0^1 \lim_{k\to\infty} f_k(x)dx = \lim_{k\to\infty} \int_0^1 f_k(x)dx.$$

What we've done in this example, instead of trying to approximate χ by step functions based on intervals, is we've asked what is the 'length' or (as we'll usually say) *measure* of the set where χ is equal to one. Lebesgue's key idea was to extend this.

To approximate a function by *step functions* as you did in Mods, you divide the *domain* of the function.

To approximate a function by *simple functions*, as we'll do here, we divide the *range* of the function.

The *simple function* in the right hand picture is defined to be

$$\sum_{k=1}^{4} y_{k-1} \mathbf{1}_{S_k}(x)$$

where $S_k = \{x \in \mathbb{R} : y_{k-1} \leq f(x) < y_k\}$ and $\mathbf{1}_S$ is the indicator function of the set S. Of course in this picture the simple function *is* a step function. But in general it won't be, as we saw in our Dirichlet function example where it was the indicator function of $\mathbb{Q} \cap [0,1]$. (See problem sheet 1 for an example of a *continuous* function for which Lebesgue's recipe does not always give a step function.)

Exercise 1.8. *As another example try approximating* $\sin\frac{1}{x}$ *on* $(0,1]$ *by simple functions.*

Here then is Lebesgue's recipe for integrating a bounded function over an interval $[a,b]$:

1. Subdivide the y-axis by a series of points

$$\min_{x\in[a,b]} f(x) \geq y_0 < y_1 < \cdots < y_n > \max_{x\in[a,b]} f(x).$$

2. Form the sum

$$\sum_{k=1}^{n} y_{k-1} \times \mathrm{measure}\left(\{\mathrm{x} \in [\mathrm{a}, \mathrm{b}] : \mathrm{y_{k-1}} \leq \mathrm{f(x)} < \mathrm{y_k}\}\right).$$

3. Let the number of points $\{y_n\}$ go to infinity in such a way that $\max(y_k - y_{k-1}) \to 0$.

The limit of the sequence of sums in 2 is $\int_a^b f(x)dx$.

In fact we'll see that for bounded functions on $[a, b]$

$$\int_a^b f(x)dx = \inf\{\int_a^b \psi(x)dx : \psi \text{ simple }, \psi \geq f\} = \sup\{\int_a^b \phi(x)dx : \phi \text{ simple }, \phi \leq f\},$$

so when a function is integrable in the sense of Mods we'll get exactly the same answer as before. But because

$$\text{step functions} \subsetneq \text{simple functions}$$

we can define the integral for a wider class of functions this way. Moreover, our recipe just needs us to be able to assign a value to

$$\mathrm{measure}\left(\{\mathrm{x} : \mathrm{y_{k-1}} \leq \mathrm{f(x)} < \mathrm{y_k}\}\right)$$

so now we're not restricted to integrating functions $f : \mathbb{R} \to \mathbb{R}$ over intervals. We can integrate functions $f : \Omega \to \mathbb{R}$ over any subset of Ω provided we have a suitable notion of size of sets where f takes particular values. In particular, Ω could be a probability space or $\mathbb{Z}$ or $\mathbb{N}$ or a sphere or a torus or ....

To see the power (and limitations) of the theory though we need to know which sets we can measure. So our first task is going to be to develop a theory of measure.

BUT FIRST:

Please remember that you can still integrate everything that you already know how to integrate. Lebesgue integration extends the theory of Mods so that you can integrate more *functions than before.*

In particular, you can still integrate continous functions over closed bounded intervals and technniques that you justified in Mods like integration by parts and substitution can still be used for those functions. Thus you may freely use the following results from your Analysis III notes:

Theorem 1.9 (Integration by parts). *If u, v are differentiable on $[a, b]$ and u', v' are in $C[a, b]$ then*

$$\int_a^b (u'v) + \int_a^b (uv') = u(b)v(b) - u(a)v(a).$$

Theorem 1.10 (Substitution). *Suppose that $g \in C([c, d])$ is strictly monotone increasing and continuously differentiable (so $g' \in C[c, d]$ and $g' > 0$) and suppose that $g(c) = a$, $g(d) = b$. Then for $f \in C[a, b]$*

$$\int_c^d f(g(x))\, g'(x)dx = \int_a^b f(x)dx.$$

2 Lebesgue measurable sets in $\mathbb{R}$

In this section we concentrate on the problem that motivated Lebesgue. Here's the problem that he proposes:

> We wish to attach a non-negative real number which we shall call the measure $m(E)$ to each bounded subset of $\mathbb{R}$ in such a way that the following conditions are satisfied:
>
> 1. If E' is a translation of E then $m(E') = m(E)$.
> 2. If $\{E_i\}_i$ is a finite or countably infinite collection of sets with $E_i \cap E_j = \emptyset$ whenever $i \neq j$ then
> $$m\left(\bigcup_i E_i\right) = \sum_i m(E_i).$$
> 3. $m((0,1)) = 1$.

Let's pause to think about 2. It seems reasonable that if I split a set into two disjoint pieces $E = E_1 \cup E_2$ with $E_1 \cap E_2 = \emptyset$ say, then $m(E) = m(E_1) + m(E_2)$. But we're asking for something a bit stronger. If $\{E_i\}_{i=1}^{\infty}$ is a countably infinite family of sets with $E_i \cap E_j = \emptyset$ for $i \neq j$, then writing

$$E = \bigcup_{i=1}^{\infty} E_i$$

and

$$B_n = E \backslash \bigcup_{i=1}^{n} E_i$$

then if we are going to have any hope of developing a theory that behaves well under passage to the limit, it is reasonable to ask for a kind of 'continuity' property, namely that if $m(E) < \infty$ then $m(B_n) \to 0$ as $n \to \infty$. If we assume that 2 holds for *finite* collections of sets then this becomes

$$m(B_n) = m(E) - m\left(\bigcup_{i=1}^{n} E_i\right) = m(E) - \sum_{i=1}^{n} m(E_i) \to 0 \quad \text{as } n \to \infty.$$

That is

$$m(E) = \lim_{n\to\infty} m(\bigcup_{i=1}^{n} E_i) = \lim_{n\to\infty} \sum_{i=1}^{n} m(E_i) = \sum_{i=1}^{\infty} m(E_i),$$

which is what 2 says for countable disjoint unions.

A consequence of Lebesgue's three conditions is (see Problem Sheet 1) that for any interval $I = [a, b]$ (or $(a, b]$, $[a, b)$, (a, b)), the measure $m(I) = b - a$. So Lebesgue measure is required to be an extension of our notion of length of an interval. It turns out (but it is beyond our scope in this course) that if we were to weaken 2 by only requiring this property for *finite* collections, then one can define a notion of length for *all* subsets of $\mathbb{R}$. But we are interested in countable operations (it was interchange of limits and integration that motivated us) so we need this stronger version and then one can construct sets to which it is not possible to assign a measure. This was first achieved by Vitali and we'll see (a variant of) his famous set in Example 2.12. On the other hand it turns out that one has to work hard to construct a non-measurable set and so we shall not be greatly inconvenienced by their existence.

Lebesgue's approach to constructing his measure owes a lot to the upper and lower sums of the integration theory from Mods. The key idea is the one that we used to suggest that the 'length' of $\mathbb{Q} \cap [0, 1]$ (that arose in integrating the Dirichlet function over $[0, 1]$) should be zero.

Recall that Lebesgue's conditions already tell us that the measure of an interval with endpoints a and b (with $a < b$ say) is $b - a$. (We also allow intervals to be empty in which case they have length zero.)

Definition 2.1 (Outer measure)**.** *For a set $E \subseteq \mathbb{R}$, the* outer measure *of E is*

$$m^*(E) = \inf\left\{\sum_{n=1}^{\infty} m(I_n) : E \subseteq \bigcup_{n=1}^{\infty} I_n, I_n \text{ intervals}\right\}.$$

What we showed in §1 was that $m^*\left(\mathbb{Q} \cap [0,1]\right) = 0$.

Definition 2.2 (Null sets)**.** *A set $E \subseteq \mathbb{R}$ is* null *(with respect to Lebesgue measure) if $m^*(E) = 0$.*

Lemma 2.3 (Some properties of null sets)**.** *1. Any subset of a null set is null.*

2. A countable union of null sets is null.

WARNING: NOT ALL NULL SETS ARE COUNTABLE.

The proof of Lemma 2.3 and an example of an *uncountable* null set are both on Problem Sheet 1.

Lemma 2.4 (Some properties of outer measure)**.** *1. If E' is a translation of E then $m^*(E') = m^*(E)$.*

2. If $\{E_i\}_i$ is a finite or countably infinite collection of sets then

$$m^*(\bigcup_i E_i) \leq \sum_i m^*(E_i).$$

3. $m^*([a,b]) = b - a$.

The difficulty is 2. It tells us that m^* is countably *sub*additive, but we cannot deduce that it is countably additive which was Lebesgue's condition 2.

Exercise 2.5. *Prove 2. If you get stuck look in, for example, Capinski & Kopp.*

The outer measure is really playing the rôle of the upper sum in our old definition of the integral if we try to define the integral of the indicator function of a set E. To define *measurable* sets we need an analogue of the lower sum. Lebesgue provided this with the *inner measure*.

Definition 2.6 (Inner measure)**.** *For a subset E of the interval $[a,b]$, the* inner measure *of E is given by*

$$\begin{aligned} m_*(E) &= m^*\left([a,b]\right) - m^*\left([a,b]\backslash E\right) \\ &= b - a - m^*\left([a,b]\backslash E\right). \end{aligned}$$

Definition 2.7 (Lebesgue measurable sets)**.** *A subset E of $\mathbb{R}$ is Lebesgue measurable if for each bounded interval I*

$$m^*(E \cap I) + m^*(E^c \cap I) = m^*(I).$$

Its Lebesgue measure is then equal to its outer measure. When a set E is Lebesgue measurable we denote its Lebesgue measure by $m(E)$.

Notice in particular that all bounded intervals are Lebesgue measurable. We can deduce from this definition that a set E is Lebesgue measurable if and only if, for every finite interval I, the inner and outer measures of the set $E \cap I$ agree.

Theorem 2.8. *The outer measure m^* restricted to the Lebesgue measurable sets satisfies Lebesgue's conditions 1, 2 and 3.*

This Theorem tells us that by restricting to the Lebesgue measurable sets we have overcome the difficulties of outer measure and recovered countable additivity. This is only going to be a good theory if the class of Lebesgue measurable sets is sufficiently large. So which sets are Lebesgue measurable?

Theorem 2.9. *The class of Lebesgue measurable sets, denoted $\mathcal{M}_{\text{Leb}}$, has the following properties:*

1. $\mathbb{R} \in \mathcal{M}_{\text{Leb}}$.
2. *If $E \in \mathcal{M}_{\text{Leb}}$ then $E^c \in \mathcal{M}_{\text{Leb}}$.*
3. *If $E_n \in \mathcal{M}_{\text{Leb}}$ for all $n \in \mathbb{N}$ then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}_{\text{Leb}}$.*

This Theorem says that $\mathcal{M}_{\text{Leb}}$ forms what is called a *σ-algebra.* It is closed under finite or countable set operations. In particular you can check, using de Morgan's laws, that the intersection of a countable collection of Lebesgue measurable sets is also Lebesgue measurable.

Proofs of Theorems 2.8 and 2.9 can be found, for example, in Temple (1971), Chapter 7. There is also a proof in Capinski & Kopp, but they take a different definition of Lebesgue measurable sets. They assume what is known as the Carathéodory condition, that

$$m^*(E \cap A) + m^*(E^c \cap A) = m^*(A)$$

for all $A \subseteq \mathbb{R}$ and not just for bounded intervals. Their definition is equivalent to ours, but that is not a very easy thing to show.

As we already remarked, all *intervals* are Lebesgue measurable and so any set obtained by taking finite or countable unions or intersections of intervals is Lebesgue measurable. In particular,

Lemma 2.10. *All open sets $E \subseteq \mathbb{R}$ are Lebesgue measurable.*

The class of Lebesgue measurable sets is beginning to look rather big. It is actually slightly larger still because all *null* sets are Lebesgue measurable (and have Lebesgue measure zero). In fact the Lebesgue measurable sets are precisely those that can be obtained from intervals and null sets by a finite or countable sequence of set operations. There are a lot of those sets.

Definition 2.11. *The collection of sets that can be obtained by a finite or countable sequence of set operations from the intervals of $\mathbb{R}$ is called the σ-algebra generated by the intervals, or the* Borel *σ-algebra on $\mathbb{R}$. Supplementing these by taking set operations with sets E with $m^*(E) = 0$ is known as taking the* completion *with respect to the (Lebesgue) outer measure and yields the Lebesgue measurable sets.*

The next (non-examinable) example illustrates how hard one has to work to construct a non-measurable set. It is a variant of an example due to Vitali. It is convenient to work with the unit circle S^1. The definition of measurable set translates in an obvious way since S^1 can be thought of as $[0, 2\pi)$.

Example 2.12 (A non-measurable set)**.** *Let*

$$S^1 = \{e^{i\theta} : \theta \in \mathbb{R}\}.$$

Define an equivalence relation on S^1 *by writing*

$$z \sim w \text{ if } \exists \alpha, \beta \in \mathbb{R} \text{ s.t. } z = e^{i\alpha}, w = e^{i\beta}, \text{ and } \alpha - \beta \in \mathbb{Q}.$$

Let A *be a set obtained by choosing exactly one representative from each equivalence class. Define*

$$A_q = e^{iq} \cdot A = \{e^{iq}z : z \in A\}.$$

Then $\{A_q\}_{q\in\mathbb{Q}}$ *is a collection of pairwise disjoint sets, each obtained from* A *by rotation and* $S^1 = \bigcup_{q\in\mathbb{Q}} A_q$.

Suppose that A *is measurable, then so is* A_q *and* $m(A_q) = m(A)$ *for all* $q \in \mathbb{Q}$. *Then by countable additivity,*

$$2\pi = m(S^1) = \sum_{q\in\mathbb{Q}} m(A_q) = \begin{cases} 0 & \text{if } m(A) = 0 \\ \infty & \text{if } m(A) > 0, \end{cases}$$

which yields a contradiction.

Thus we cannot *assign a measure to* A.

It turns out that to construct a non-measurable set one has to invoke the 'axiom of choice' which is what allowed us to choose exactly one representative from each of the uncountably infinite number of equivalence classes.

In higher dimensions things are even worse. In three dimensions one can cut a solid ball into a finite number of non-overlapping sets which can then be reasssembled to yield two identical copies of the original ball. Do a web search on Banach-Tarski paradox to find out more.

3 Other measure spaces

In §2 we saw that we could define a notion of 'length' or 'measure' for a wide class of subsets of $\mathbb{R}$ called the Lebesgue measurable sets. A fancy way to say this is to say that $(\mathbb{R}, \mathcal{M}_{\text{Leb}}, m)$ is a *measure space.*

Definition 3.1 (Measure space)**.** *We say that* $(\Omega, \mathcal{F}, \mu)$ *is a* measure space *if* Ω *is an abstractly given set,* $\mathcal{F}$ *is a collection of subsets of* Ω *which includes* $\emptyset$ *and is closed under complements and finite or countable unions (recall that this says that* $\mathcal{F}$ *is a* σ*-algebra) and* $\mu : \mathcal{F} \to [0, \infty]$ *is a* countably additive set function*; that is if* $\{E_n\}_{n\in\mathbb{N}}$ *is a countable collection of sets from* $\mathcal{F}$ *with* $E_i \cap E_j = \emptyset$ *whenever* $i \neq j$ *then*

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

There are lots of measure spaces out there, several of which you are already familiar with.

Example 3.2 (Discrete measure theory)**.** *Let* Ω *be a countable set and* $\mathcal{F}$ *the* power set *of* Ω *(that is the collection of* all *subsets of* Ω*). A* mass function *is any function* $\bar{\mu} : \Omega \to [0, \infty]$. *We can then define a measure on* Ω *by* $\mu(\{x\}) = \bar{\mu}(x)$ *and extend to arbitrary subsets of* Ω *using Property 2.*

Equally given a measure on Ω *we can define a mass function.*

In particular,

1. *If* $\sum_{x\in\Omega} \bar{\mu}(x) = \sum_{x\in\Omega} \mu(\{x\}) = 1$ *then* μ *is a probability measure and* $\bar{\mu}$ *is the corresponding probability mass function. For* $A \subseteq \Omega$,

$$\mu(A) = \sum_{x\in A} \mu(\{x\}).$$

The null sets correspond in the terminology of Mods probability to events that have probability zero.

2. *If $\Omega = \mathbb{N}$ and $\bar{\mu}(k) = 1, \forall k \in \mathbb{N}$ then we get* counting measure. *For $A \subseteq \mathbb{N}$,*

$$\mu(A) = \sharp\{n : n \in A\}.$$

In this case the only null set is the empty set.

These discrete measure spaces provide a 'toy' version of the general theory where each of the results we prove for general measure spaces reduces to some straightforward fact about convergence of series. This is all we need to do (for example) discrete probability and so that topic is generally introduced without discussing measure theory. But for more general probability and most of analysis (both pure and applied) this is not enough.

Example 3.3. *Consider the problem from 'continuous' probability of picking a point uniformly from $[0, 1]$. We face making sense of the 'probability' that a given number is chosen or that the number chosen falls within a particular set. Write U for the (random) point picked. Evidently $\mathbb{P}[U = x] = 0$ if all points are equally likely, for otherwise $\sum_{x\in[0,1]} \mathbb{P}[U = x] = \infty$. This means that we won't be able to distinguish the probabilities of two sets that differ by just one point, or indeed that differ by a finite number of points. On the other hand we'd expect to be able to define for an interval $I = [a, b] \subseteq [0, 1]$*

$$\mathbb{P}[U \in I] = b - a.$$

The Lebesgue measurable subsets of $[0, 1]$ are precisely those to which we can assign $\mathbb{P}[U \in A]$ in a consistent way.

The measure that we defined on the Lebesgue measurable subsets of $\mathbb{R}$ was certainly not the *only* measure on $(\mathbb{R}, \mathcal{M}_{\text{Leb}})$. The form that it took was forced upon us by our insistence that $m([a, b]) = b-a$.

Example 3.4. 1. *Define*

$$\mu([a, b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Then μ extends to a measure on all the Lebesgue measurable subsets of $\mathbb{R}$. This measure is of central importance in statistics.

2. *For $[a, b] \subseteq [0, \infty)$ define*

$$\mu_t([a, b]) = \int_a^b e^{-tx} dx.$$

Then for each $t > 0$, μ_t extends to a measure on all the Lebesgue measurable subsets of $[0, \infty)$. This family *of measures is very important in the study of differential equations.*

These examples are all of either discrete measures or measures which are *absolutely continuous* with respect to Lebesgue measure. We'll see the origin of that term in §7. But of course one can have mixtures of the two (see Problem Sheet 2 for an example) and in fact in §7 we'll see that there are more exotic measures still.

Although mostly we'll concentrate on Lebesgue measure, always keep in mind that our theory is *much* more general.

4 Measurable functions

We now turn our attention to the objects at the heart of integration theory: the measurable functions.

Last term in the Analysis course you saw that a function $f : X \to Y$ (where X and Y are subsets of Euclidean space) is *continuous* if and only if

$$A \text{ open in } Y \implies f^{-1}(A) \text{ open in } X.$$

If you are studying topology then you'll know that this definition applies more generally. For the definition of measurable functions we'll replace 'open' by 'measurable'. The advantage of measurable functions over continuous functions is that they are much more 'stable'. In particular, they behave well under limits: the pointwise limit of a sequence of measurable functions is measurable - and this 'closure' of the space of measurable functions is at the heart of the convergence theorems.

Definition 4.1 (Measurable function)**.** *Let $(\Omega, \mathcal{F}, \mu)$ and $(\Lambda, \mathcal{G}, \nu)$ be measure spaces. A function $f : \Omega \to \Lambda$ is* measurable *(with respect to $\mathcal{F}$, $\mathcal{G}$) if and only if*

$$G \in \mathcal{G} \implies f^{-1}(G) \in \mathcal{F}.$$

For concreteness, mostly we're going to work in the special setting of Lebesgue measurable functions on $\mathbb{R}$. But the arguments that we use will be insensitive to replacing the Lebesgue measurable sets, $\mathcal{M}_{\mathrm{Leb}}$, by $\mathcal{F}$ of Definition 4.1.

Definition 4.2 (Lebesgue measurable function)**.** *A function $f : \mathbb{R} \to \mathbb{R}$ is* Lebesgue measurable *if and only if for every interval $I \subseteq \mathbb{R}$, $f^{-1}(I) = \{x \in \mathbb{R} : f(x) \in I\}$ is a Lebesgue measurable set.*

Similarly, if E is a measurable subset of $\mathbb{R}$, then $f : E \to \mathbb{R}$ is Lebesgue measurable if for each interval I, $f^{-1}(I)$ is Lebesgue measurable.

In what follows, if we say that $f : \mathbb{R} \to \mathbb{R}$ is measurable, without further qualification, then we mean Lebesgue measurable.

The definition is *equivalent* to requiring that for all *Borel sets* A, $f^{-1}(A)$ is *Lebesgue* measurable. The point is that it is enough to check just for intervals and indeed we don't need to check it for *all* intervals I. It is convenient to know the following result.

Theorem 4.3. *Let E be a measurable subset of $\mathbb{R}$ and $f : E \to \mathbb{R}$ be a function. The following are equivalent:*

1. *f is measurable.*
2. *For all a, $f^{-1}((a, \infty))$ is measurable.*
3. *For all a, $f^{-1}([a, \infty))$ is measurable.*
4. *For all a, $f^{-1}(-\infty, a))$ is measurable.*
5. *For all a, $f^{-1}(-\infty, a])$ is measurable.*

Proof. Let's just prove one of these. Obviously (1) implies all the others. We prove here that (2) $\implies$ (1).

We must show that for *any* interval I, $f^{-1}(I) \in \mathcal{M}_{\mathrm{Leb}}$. We already have that $f^{-1}((a, \infty)) \in \mathcal{M}_{\mathrm{Leb}}$. Now suppose that $I = (-\infty, a]$.

$$f^{-1}((-\infty, a]) = f^{-1}(\mathbb{R} \backslash (a, \infty)) = E \backslash f^{-1}((a, \infty)) \in \mathcal{M}_{\mathrm{Leb}}$$

since measurable sets are closed under taking complements.

Now note that

$$\begin{aligned} f^{-1}((-\infty,b)) &= f^{-1}\left(\bigcup_{n=1}^{\infty}(-\infty,b-\frac{1}{n}]\right) \\ &= \bigcup_{n=1}^{\infty} f^{-1}((-\infty,b-\frac{1}{n}]) \in \mathcal{M}_{\mathrm{Leb}} \end{aligned}$$

since countable unions of measurable sets are measurable. Taking complements we also have $f^{-1}(([b,\infty)) \in \mathcal{M}_{\mathrm{Leb}}$.

Now let $I = (a,b)$.

$$f^{-1}((a,b)) = f^{-1}((-\infty,b)) \cap f^{-1}((a,\infty)) \in \mathcal{M}_{\mathrm{Leb}}.$$

Similarly,

$$f^{-1}([a,b]) = f^{-1}((-\infty,b]) \cap f^{-1}([a,\infty)) \in \mathcal{M}_{\mathrm{Leb}}.$$

The same reasoning gives the half open intervals. □

In practice, most functions are measurable.

Example 4.4. *1. Constant functions are measurable.*

2. More generally, continuous functions are measurable.

3. The indicator function of the set A defined by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise,} \end{cases}$$

is measurable if and only if A is (Lebesgue) measurable.

Proof. We just prove the first two assertions (the third is on Problem Sheet 2).

First suppose that $f(x) \equiv c$. Then

$$f^{-1}((a,\infty)) = \begin{cases} \mathbb{R} & \text{if } a < c \\ \emptyset & \text{otherwise} \end{cases}$$

and in both cases we have Lebesgue measurable sets.

Now suppose that f is continuous. Note that (a,∞) is an open set and, recalling that for a continuous function the inverse image of an open set is open, $f^{-1}((a,\infty))$ must be an open set. All open sets are measurable so f is measurable as required. □

Theorem 4.5. *If f and g are measurable functions then so are $f+g$ and fg. In particular, taking $g(x) \equiv c$ (a constant) cf is measurable.*

This result tells us that the measurable functions form a vector space which is closed under multiplication.

An elegant proof of Theorem 4.5 uses the following useful Lemma.

Lemma 4.6. *Suppose that $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a continuous function. If f and g are measurable, then $h(x) = F(f(x), g(x))$ is also measurable.*

Theorem 4.5 will follow upon setting $F(u,v) = u + v$ and $F(u,v) = uv$.

Proof of Lemma 4.6. The proof uses the two-dimensional analogue of the fact that open sets in $\mathbb{R}$ can be expressed as finite or countable unions of open intervals. In two dimensions every open set is a finite or countable union of open *rectangles*.

So for any real a, put $G_a = F^{-1}(a,\infty)$, that is $\{x : h(x) > a\} = \{x : (f(x), g(x)) \in G_a\}$. Then G_a is open, since F is continuous, and so can be written as

$$G_a = \bigcup_{n=1}^{n} R_n$$

where the R_n are open rectangles, $R_n = (a_n, b_n) \times (c_n, d_n)$ say. So

$$h^{-1}(a,\infty) = \{x : h(x) > a\} = \bigcup_{n=1}^{n} \{x : f(x) \in (a_n, b_n)\} \cap \{x : g(x) \in (c_n, d_n)\}$$

which is measurable by measurability of f and g and since $\mathcal{M}_{\text{Leb}}$ is closed under finite or countable set operations. $\square$

The following results are on Problem Sheet 2.

Proposition 4.7. *Let E be a measurable subset of $\mathbb{R}$ and $f : E \to \mathbb{R}$ a function. Let*

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) > 0 \\ 0 & \text{if } f(x) \le 0, \end{cases}$$

and

$$f^-(x) = \begin{cases} 0 & \text{if } f(x) > 0 \\ -f(x) & \text{if } f(x) \le 0, \end{cases}$$

Then

1. *f is measurable if and only if f^+ and f^- are measurable.*
2. *If f is measurable then so is $|f|$, but the converse is false.*
3. *If f is measurable then*

$$f^a(x) = \begin{cases} a & \text{if } f(x) > a \\ f(x) & \text{if } f(x) \le a, \end{cases}$$

is also measurable.

Before proving our key result about measurable functions - stability under limits - we need to introduce a new concept which generalises the idea of limit.

Definition 4.8 (limsup and liminf)**.** *For a sequence of real numbers $(x_n)_{n\in\mathbb{N}}$ we define*

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup_{m\ge n} x_m,$$

and

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf_{m\ge n} x_m.$$

To understand this a little better, suppose first that the sequence $\{x_n\}_{n\geq 1}$ is *bounded*. Now notice that $S_n \equiv \sup_{m\geq n} x_m$ *decreases* as n increases (since we are taking the supremum over a smaller set). Moreover, since the sequence $\{x_n\}_{n\geq 1}$ is bounded, $\{S_n\}_{n\geq 1}$ is also bounded and a monotone decreasing sequence of real numbers which is bounded below necessarily *converges*. The limit, $\lim_{n\to\infty} S_n$, is denoted $\limsup_{n\to\infty} x_n$. Literally it is 'the limit of the suprema'.

Similarly, $I_n = \inf_{m\geq n} x_m$ defines a monotone *increasing* sequence of real numbers which is bounded above and hence convergent. Then $\lim_{n\to\infty} I_n$ is denoted $\liminf_{n\to\infty} x_n$.

Since the infimum of a decreasing sequence is its limit, one sometimes writes

$$\limsup_{n\to\infty} x_n = \inf_{n\geq 1}\{\sup_{m\geq n} x_m\}.$$

Similarly

$$\liminf_{n\to\infty} x_n = \sup_{n\geq 1}\{\inf_{m\geq n} x_m\}.$$

We will use this in the proof of Theorem 4.12 below.

The $\limsup$ and $\liminf$ of a bounded sequence $\{x_n\}_{n\geq 1}$ are equal if and only if the sequence $\{x_n\}_{n\geq 1}$ *converges*, then $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = \lim_{n\to\infty} x_n$. In general the $\liminf$ and $\limsup$ of a sequence will be different.

Example 4.9. *1. Let* $x_n = \sin\left(n\frac{\pi}{2}\right)$. *Then*

$$\liminf_{n\to\infty} x_n = -1, \qquad \limsup_{n\to\infty} x_n = 1.$$

2. Let $x_n = e^{\sin(n\pi/2)}$. *Then*

$$\liminf_{n\to\infty} x_n = e^{-1}, \qquad \limsup_{n\to\infty} x_n = e^{1}.$$

We can use the same definition for *unbounded* sequences where we use the convention that the supremum of a set which is not bounded above is $+\infty$ and the infimum of a set which is not bounded below is $-\infty$. In general then $\liminf$ and $\limsup$ can take the values $\pm\infty$.

Example 4.10. *1. Let* $x_n = (-1)^n n$. *Then*

$$\liminf_{n\to\infty} x_n = -\infty, \qquad \limsup_{n\to\infty} x_n = \infty.$$

2. Let $x_n = e^{(-1)^n n^2}$. *Then*

$$\liminf_{n\to\infty} x_n = 0, \qquad \limsup_{n\to\infty} x_n = \infty.$$

For both bounded and unbounded sequences the idea is that

$$\text{if } z > \limsup_{n\to\infty} x_n \text{ then } x_n < z \textit{ eventually},$$

that is for sufficiently large n. Moreover,

$$\text{if } z < \limsup_{n\to\infty} x_n \text{ then } x_n > z \textit{ infinitely often,}$$

that is given N there is $n > N$ so that $x_n > z$.

Similarly

$$\text{if } v < \liminf_{n\to\infty} x_n \text{ then } x_n > v \textit{ eventually},$$

and

$$\text{if } v > \liminf_{n\to\infty} x_n \text{ then } x_n < v \textit{ infinitely often.}$$

For any $\epsilon > 0$ the sequence is eventually trapped between $\liminf_{n\to\infty} x_n - \epsilon$ and $\limsup_{n\to\infty} x_n + \epsilon$.

In the picture the sequence is not converging, but it is becoming trapped between LI and LS.

Example 4.11. *1. Let* $x_n = (-1)^n \frac{n^2+2n+16}{n^2+79}$ *then* $\liminf_{n\to\infty} x_n = -1$, $\limsup_{n\to\infty} x_n = 1$.

2. Let

$$x_n = \exp\left(1 + \frac{n}{n+2}\sin(n\pi/2)\right),$$

then

$$\liminf_{n\to\infty} x_n = 1, \quad \limsup_{n\to\infty} x_n = e^2.$$

Now here is the key result which makes working with measurable functions so flexible.

Theorem 4.12. *Let* $\{f_n\}_{n\in\mathbb{N}}$ *be a sequence of measurable functions defined on the (measurable) set* $E \subseteq \mathbb{R}$. *Then the following are also measurable:*

$$\max_{n\le k} f_n, \quad \min_{n\le k} f_n, \quad \sup_{n\in\mathbb{N}} f_n, \quad \inf_{n\in\mathbb{N}} f_n, \quad \limsup_{n\to\infty} f_n, \quad \liminf_{n\to\infty} f_n.$$

Proof. First recall that if f is measurable then so is $-f$ and so since

$$\min_{n\le k} f_n = -\max_{n\le k}\{-f_n\}, \quad \inf_{n\in\mathbb{N}} f_n = -\sup_{n\in\mathbb{N}}\{-f_n\} \quad \text{and } \liminf_{n\to\infty} f_n = -\limsup_{n\to\infty}\{-f_n\}$$

we only have to consider three of the statements.

Now note that

$$\{x : \max_{n\le k} f_n(x) > a\} = \bigcup_{n=1}^{k}\{x : f_n(x) > a\}$$

and each of the sets on the right hand side is measurable (since f_n is). Since the union of a finite or countable collection of measurable sets is measurable we have shown that $\max_{n\le k} f_n$ is measurable.

$$\{x : \sup_{n\ge k} f_n(x) > a\} = \bigcup_{n=k}^{\infty}\{x : f_n(x) > a\}$$

so in the same way we have that $\sup_{n\ge k} f_n$ is measurable. In particular the case $k = 0$ gives that $\sup_{n\in\mathbb{N}} f_n = 0$.

Finally,

$$\limsup_{n\to\infty} f_n = \inf_{n\ge 1}\{\sup_{m\ge n} f_m\}$$

and we have shown that $h_n = \sup_{m\ge n} f_m$ is measurable and so $\inf_{n\ge 1} h_n$ is measurable and the result is proved. $\square$

Corollary 4.13. *If a sequence of measurable functions* $\{f_n\}_{n\ge 1}$ *converges pointwise to a limit* f *then the limit is a measurable function.*

Proof. This is immediate since for a *convergent* sequence

$$\lim_{n\to\infty} f_n = \limsup_{n\to\infty} f_n \quad \left(= \liminf_{n\to\infty} f_n\right)$$

which Theorem 4.12 tells us is measurable. □

The null sets play a very important rôle in Lebesgue integration.

Definition 4.14 (Almost everywhere)**.** *We say that a property holds* almost everywhere*, written a.e., if it holds except on a set of measure zero.*

In particular, for two functions f and g defined on a set $E \subseteq \mathbb{R}$ we say that $f = g$ a.e. if $\{x \in E : f(x) \neq g(x)\}$ has measure zero.

Similarly, if $\{f_n\}_{n\geq 1}$ is a sequence of measurable functions and $f_n(x) \to f(x)$ except on a set of measure zero then we say that f_n converges to f a.e. or for a.e. x.

Remark 4.15 (Null sets)**.** *Always remember that which sets are of measure zero is determined by which measure you are using. For example, the null sets for Lebesgue measure are not the same as for counting measure. If it is not completely clear from the context then specify which measure you are using.*

Example 4.16. *The Dirichlet function of Definition 1.5 is equal to 1 a.e. with respect to Lebesgue measure (since the rationals are a Lebesgue null set).*

Although measurable sets and measurable functions are new concepts, it is worth remembering three principles aptly summarised by Littlewood:

1. Every Lebesgue measurable set of $\mathbb{R}$ of finite measure is 'nearly' a finite union of intervals.
2. Every measurable function on a set of finite measure is 'nearly' continuous.
3. Every convergent sequence of measurable functions defined on a set of finite measure is 'nearly' uniformly convergent.

The catch of course is the word 'nearly'. An exact formulation of 3 is the following important result.

Theorem 4.17 (Egorov's Theorem)**.** *Suppose that $\{f_k\}_{k=1}^{\infty}$ is a sequence of measurable functions defined on a measurable set E with $m(E) < \infty$ and assume that $f_k \to f$ a.e. on E.*

Given $\epsilon > 0$ we can find a closed set $A_\epsilon \subseteq E$ such that $m(E \backslash A_\epsilon) < \epsilon$ and $f_k \to f$ uniformly *on A_ϵ.*

Our proof requires a lemma.

Lemma 4.18. *Let E be a measurable set with $m(E) < \infty$. Let $\{E_i\}_{i\in\mathbb{N}}$ be a sequence of measurable sets with $E_i \subseteq E_{i+1}$ and $E = \bigcup_i E_i$. Then given $\epsilon > 0$, $\exists N$ such that $n \geq N$ implies $m(E\backslash E_n) < \epsilon$.*

Proof. This is really a rehash of an argument at the beginning of §2 that we used to motivate Lebesgue's condition of countable additivity except that now we are not taking the sets E_i to be pairwise disjoint and so we define a new sequence of sets F_i by $F_1 = E_1$ and $F_{i+1} = E_{i+1}\backslash E_i$ for $i \geq 1$. The sets F_i *are* pairwise disjoint and $E = \bigcup_i F_i$ so by countable additivity of Lebesgue measure

$$m(E) = \sum_i m(F_i). \tag{3}$$

Now $m(E) < \infty$ and so the sum on the right hand side of equation (3) converges which implies that given $\epsilon > 0$, $\exists N$ so that $n \geq N$ implies

$$m(\bigcup_{i \geq n+1} F_i) = \sum_{i=n+1}^{\infty} m(F_i) < \epsilon$$

which on rewriting in terms of our original sets E_i becomes $m(E \backslash E_n) < \epsilon$ whenever $n \geq N$ as required. □

Note. The assumption that $m(E) < \infty$ was *essential* to this proof.

Proof of Egorov's Theorem.

We don't use the extra information that A_ϵ is closed in the proofs that follow and so we restrict ourselves to showing that there is a measurable set A_ϵ with $f_k \to f$ uniformly on A_ϵ and $m(E \backslash A_\epsilon) < \epsilon$. (The fact that we can take A_ϵ closed is then a consequence of the first of Littlewood's three principles.)

Without loss of generality we can assume that $f_k(x) \to f(x)$ for *every* $x \in E$ (otherwise remove the set of measure zero where it fails from E).

For every pair of non-negative integers n and k let

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

Now *fix* n. Note that $E_k^n \subseteq E_{k+1}^n$ and $E = \bigcup_k E_k^n$, so by Lemma 4.18, $\exists k_n$ such that $m(E \backslash E_{k_n}^n) < 1/2^n$. Notice that by construction

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

Now choose N so that $\sum_{n=N}^{\infty} 2^{-n} < \epsilon$ and let

$$A_\epsilon = \bigcap_{n \geq N} E_{k_n}^n.$$

First note that

$$m(E \backslash A_\epsilon) \leq \sum_{n=N}^{\infty} m(E \backslash E_{k_n}^n) < \epsilon.$$

Next, if $\delta > 0$, choose $n \geq N$ such that $1/n < \delta$ and note that $x \in A_\epsilon$ implies $x \in E_{k_n}^n$ so that $|f_j(x) - f(x)| < \delta$ whenever $j > k_n$. Hence $f_k \to f$ uniformly on A_ϵ as required. □

Note that just as for the proof of Lemma 4.18, it is crucial that $m(E) < \infty$.

Example 4.19. *Let $E = \mathbb{R}$ and let*

$$f_k(x) = \begin{cases} 1 & \text{if } x \in (-k, k) \\ 0 & \text{otherwise.} \end{cases}$$

Then $f_k(x) \to 1$ as $k \to \infty$ for all x, but the convergence is not uniform on sets whose complements have finite measure.

5 Integration and the convergence theorems

Finally we are in a position to follow Lebesgue's recipe and define the *integral.* We proceed in stages by progressively integrating

1. 'Simple' functions,
2. Bounded functions supported on a set of finite measure,
3. Non-negative functions,
4. Integrable functions (the general case).

STEP 1. Simple functions.

Definition 5.1. *A* simple function *is a finite sum*

$$\phi(x) = \sum_{k=1}^{N} a_k \mathbf{1}_{E_k}(x) \tag{4}$$

where each E_k is a measurable set of finite measure and the a_k are constants.

Step functions are simple functions, but there are simple functions that are not step functions.

Example 5.2. *The function*

$$\phi(x) = \mathbf{1}_{\mathbb{Q}\cap[0,1]}(x)$$

is a simple function, but it is not *a step function.*

A complication (that mirrors what you observed for step functions in Mods) is that any given simple function can be written in a multitude of different ways as a finite linear combination of indicator functions. As a trivial example observe that $0 = \mathbf{1}_A - \mathbf{1}_A$ for any measurable set A. Fortunately there is a *natural* representation which can be described unambiguously.

Definition 5.3. *The* canonical form *of a simple function ϕ is the unique decomposition as in (4) where the numbers a_k are distinct and the sets E_k are disjoint.*

Finding the canonical form is easy. ϕ can take on only finitely many values $c_1, \ldots, c_M$ say. Set $F_k = \{x : \phi(x) = c_k\}$. Evidently these are disjoint so set

$$\phi(x) = \sum_{k=1}^{M} c_k \mathbf{1}_{F_k}(x).$$

Definition 5.4. *If ϕ is a simple function with canonical form*

$$\phi(x) = \sum_{k=1}^{M} c_k \mathbf{1}_{F_k}(x)$$

then we define the Lebesgue integral of ϕ by

$$\int_{\mathbb{R}} \phi(x)dx = \sum_{k=1}^{M} c_k m(F_k).$$

If E is a subset of $\mathbb{R}$ of finite measure then $\phi(x)\mathbf{1}_E(x)$ is also a simple function and we define

$$\int_E \phi(x)dx = \int \phi(x)\mathbf{1}_E(x)dx.$$

Remark 5.5. *Notice that there is really nothing special about Lebesgue measure here. If* $\{E_k\}_{k=1}^M$ *are measurable with respect to a measure* μ *then we can define*

$$\int \phi(x)\mu(dx) = \sum_{k=1}^{M} c_k \mu(E_k).$$

Proposition 5.6. *The integral of simple functions defined in this way has the following properties:*

1. Independence of the representation. *If* $\phi(x) = \sum_{k=1}^N a_k \mathbf{1}_{E_k}(x)$ *is any representation of* ϕ *then*
$$\int \phi(x)dx = \sum_{k=1}^{N} a_k m(E_k).$$
2. Linearity. *If* ϕ *and* ψ *are simple functions and* $a, b \in \mathbb{R}$ *then*
$$\int (a\phi + b\psi)(x)dx = a\int \phi(x)dx + b\int \psi(x)dx.$$
3. Additivity. *If* E *and* F *are* disjoint *subsets of* $\mathbb{R}$ *with finite measure then*
$$\int_{E\cup F} \phi(x)dx = \int_E \phi(x)dx + \int_F \phi(x)dx.$$
4. Monotonicity. *If* $\phi \leq \psi$ *are simple functions then*
$$\int \phi(x)dx \leq \int \psi(x)dx.$$
5. Triangle inequality. *If* ϕ *is a simple function then so is* $|\phi|$ *and*
$$\left|\int \phi(x)dx\right| \leq \int |\phi(x)|dx.$$

The only slightly tricky one of these assertions to check is the first. For that one, if the E_k are disjoint then for each distinct value taken by ϕ, take the union of the sets E_k on which ϕ takes that value to reduce to canonical form and then use that for $E_j \cap E_k = \emptyset$, $m(E_j \cup E_k) = m(E_j) + m(E_k)$. If the E_k's are *not* disjoint, then first 'refine' the sets so that they are disjoint (in much the same way as the corresponding proof for step functions in Mods refined the partition) and thus reduce to the previous case. □

Notice that if f and g are simple functions that agree almost everywhere then $\int f(x)dx = \int g(x)dx$.

STEP 2. Bounded functions supported on a set of finite measure.

Definition 5.7. *The* support *of a measurable function* f *is defined to be the set of points where* f *does not vanish*

$$\mathrm{supp(f)} = \{\mathrm{x} : \mathrm{f(x)} \neq 0\}.$$

We shall say that f *is* supported on the set E *if* $f(x) = 0$ *whenever* $x \notin E$.

Since f is measurable, so is the set supp(f). The next goal is to integrate bounded functions for which $m(\mathrm{supp(f)}) < \infty$. The key step is to approximate such f above and below by simple functions in much the same way as in Mods we approximated continuous functions above and below by step functions.

Proposition 5.8. *Let f be a bounded function on a measurable set E with $m(E) < \infty$. In order that*

$$\inf\{\int_E \psi(x)dx : \psi \textit{ simple }, \psi \geq f\} = \sup\{\int_E \phi(x)dx : \phi \textit{ simple }, \phi \leq f\} \tag{5}$$

it is necessary and sufficient that f be measurable.

We mainly care about the *sufficiency* here because then we'll *define* the integral to be the sup (or equivalently inf) in (5).

Proof of sufficiency. Suppose that f is measurable and $|f| < M$ say. We'll follow Lebesgue's prescription to construct two sequences of simple functions $\{\phi_n(x)\}_{n\geq 1}$ and $\{\psi_n(x)\}_{n\geq 1}$ so that $\phi_n \uparrow f$, $\psi_n \downarrow f$ and $\int_E \psi_n - \int_E \phi_n \to 0$ as $n \to \infty$. The 'inf' is clearly less than or equal to the 'sup' in (5) and so this will prove the equality.

Recall then Lebesgue's recipe from §1 (where now $[a, b]$ is replaced by the measurable set E):

1. Subdivide the y-axis by a series of points

$$\min_{x\in E} f(x) \geq y_0 < y_1 < \cdots < y_N > \max_{x\in E} f(x).$$

2. Let $F_k = \{x \in E : y_{k-1} \leq f(x) < y_k\}$ and consider the simple function

$$\sum_{k=1}^{N} y_{k-1}\mathbf{1}_{F_k}(x) \tag{6}$$

with integral

$$\sum_{k=1}^{N} y_{k-1}m(F_k).$$

3. Let the number of points $\{y_k\}$ go to infinity in such a way that $\max(y_k - y_{k-1}) \to 0$.

The sum in (6) gives a candidate for the ϕ_n's. Modifying it to

$$\sum_{k=1}^{N} y_k\mathbf{1}_{F_k}$$

gives a candidate for the ψ_n's.

All we need is a concrete choice for the points $y_i \in \mathbb{R}$. Since $-M < f(x) < M$ it is natural to choose $y_0 = -M$ and $y_N = M$. We take $N = 2n$ and subdivide $[-M, M]$ equally to obtain the other points. It is convenient to write

$$E_k = \left\{x : \frac{(k-1)M}{n} \leq f(x) < \frac{kM}{n}\right\}, \quad -n+1 \leq k \leq n.$$

(This is just F_{k+n} in our previous notation.) Now the sets $\{E_k\}_{k=-n+1}^{n}$ are disjoint and measurable (since f is) and $\sum_{k=-n+1}^{n} m(E_k) = m(E)$.

Define

$$\psi_n(x) = \sum_{k=-n+1}^{n} \frac{kM}{n}\mathbf{1}_{E_k}(x), \quad \phi_n(x) = \sum_{k=-n+1}^{n} \frac{(k-1)M}{n}\mathbf{1}_{E_k}(x).$$

Then

$$\phi_n(x) \leq f(x) \leq \psi_n(x).$$

Thus

$$\inf\left\{\int_E \psi(x)dx : \psi \text{ simple }, \psi \geq f\right\} \leq \int_E \psi_n(x)dx = \sum_{k=-n+1}^{n} \frac{kM}{n} m(E_k)$$

and

$$\sup\left\{\int_E \phi(x)dx : \phi \text{ simple }, \phi \leq f\right\} \geq \int \phi_n(x)dx = \sum_{k=-n+1}^{n} \frac{(k-1)M}{n} m(E_k),$$

from which

$$\begin{aligned}
0 &\leq \inf\left\{\int_E \psi(x)dx : \psi \text{ simple }, \psi \geq f\right\} - \sup\left\{\int_E \phi(x)dx : \phi \text{ simple }, \phi \leq f\right\} \\
&\leq \sum_{k=-n+1}^{n} \left\{\frac{kM}{n} - \frac{(k-1)M}{n}\right\} m(E_k) \\
&= \frac{M}{n} \sum_{k=-n+1}^{n} m(E_k) = \frac{M}{n} m(E).
\end{aligned}$$

Since $m(E) < \infty$, $M < \infty$ and n is arbitrary we have that

$$\inf\left\{\int_E \psi(x)dx : \psi \text{ simple }, \psi \geq f\right\} = \sup\left\{\int_E \phi(x)dx : \phi \text{ simple }, \phi \leq f\right\}$$

as required. $\square$

The proof of necessity can be found, for example, in Royden.

Definition 5.9. *We define the quantity in (5) to be the Lebesgue integral of f over E and denote it by*

$$\int_E f(x)dx.$$

If $E = [a, b]$ we'll write this as $\int_a^b f(x)dx$ and for $a > b$ define $\int_a^b = -\int_b^a$.

Notice that Proposition 5.8 means that we can integrate bounded measurable functions over sets of finite measure with *no* extra conditions. This is already a big step up from what we could do with the technology of Mods. But is this theory a genuine *extension* of that of Mods? That is, for functions that are integrable in the sense of Mods will we get the same answer from our modified recipe?

Recall the notation

$$\underline{\int_a^b} f(x)dx = \sup\left\{\int_a^b \phi : \phi \in L^{\text{step}}[a, b], \phi \leq f\right\}, \quad \overline{\int_a^b} f(x)dx = \inf\left\{\int_a^b \phi : \phi \in L^{\text{step}}[a, b], \phi \geq f\right\}$$

Proposition 5.10. *Let f be a bounded measurable function on $[a, b]$. If it is integrable in the sense of Mods then it is also Lebesgue integrable and the value of the two integrals agree.*

Proof. Since every step function is a simple function

$$\underline{\int_a^b} f(x)dx \leq \sup\{\int_E \phi(x)dx : \phi \text{ simple }, \phi \leq f\} \leq \inf\{\int_E \psi(x)dx : \psi \text{ simple }, \psi \geq f\} \leq \overline{\int_a^b} f(x)dx$$

and the result follows. $\square$

In fact Lebesgue was able to characterise exactly which functions are integrable in the sense of Mods.

Theorem 5.11 (Lebesgue)**.** *Let $f : [a, b] \to \mathbb{R}$ be bounded.*

$$\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx$$

if and only if the set of points at which f is discontinuous has Lebesgue measure zero.

Example 5.12. *1. The Dirichlet function of Definition 1.5 is Lebesgue integrable over $[0, 1]$, but* not *integrable in the sense of Mods.*

2. The popcorn function (also known as Thomae's function or Riemann's function or ...) defined by

$$f(x) = \begin{cases} 1/q & x \in \mathbb{Q}, x = p/q, (p, q) = 1 \\ 0 & x \notin \mathbb{Q}, \end{cases}$$

is integrable over $[0, 1]$ in the sense of Mods and hence also in the sense of Lebesgue.

Life is often made easier by the following (intuitively clear) result.

Lemma 5.13. *Let f be a bounded measurable function supported on a measurable set E with $m(E) < \infty$. If $\{\phi_n\}_{n\geq 1}$ is any sequence of simple functions, bounded by M, supported by E and such that $\phi_n(x) \to f(x)$ a.e. x then $\lim_{n\to\infty} \int_E \phi_n(x)dx$ exists and*

$$\lim_{n\to\infty} \int_E \phi_n(x)dx = \int_E f(x)dx.$$

Proof. Let $\{\psi_n\}_{n\geq 1}$ be as in the proof of Proposition 5.8. Note that $\{\phi_n - \psi_n\}_{n\geq 1}$ is a sequence of bounded simple functions on E with $\phi_n - \psi_n \to 0$ a.e.. By linearity of the integral for simple functions it is enough therefore to show that if $\{\phi_n\}_{n\geq 1}$ is a sequence of *bounded* simple functions with $\phi_n \to 0$ a.e. then $\int_E \phi_n(x)dx \to 0$.

Fix $\epsilon > 0$, then Egorov's Theorem guarantees the existence of a (closed) measurable set A_ϵ with $m(E\backslash A_\epsilon) < \epsilon$ and such that $\phi_n \to 0$ *uniformly* on A_ϵ. Then

$$\begin{aligned} \left|\int_E \phi_n(x)dx\right| &\leq \int_E |\phi_n(x)|dx \\ &= \int_{A_\epsilon} |\phi_n(x)|dx + \int_{E\backslash A_\epsilon} |\phi_n(x)|dx. \end{aligned}$$

Now $\phi_n \to 0$ uniformly on A_ϵ so $\exists N$ such that $|\phi_n(x)| < \epsilon$ for all $x \in A_\epsilon$ whenever $n \geq N$. Then for $n \geq N$ we have

$$\begin{aligned} \left|\int_E \phi_n(x)dx\right| &\leq \epsilon m(A_\epsilon) + Mm(E\backslash A_\epsilon) \\ &\leq \epsilon m(E) + \epsilon M. \end{aligned}$$

Since ϵ was arbitrary and M, $m(E)$ are finite the result follows. □

Lemma 5.13 is often combined with the following result.

Lemma 5.14 (Approximation by Step Functions)**.** *Suppose that f is a measurable function supported on a set $E \subset \mathbb{R}$ of bounded Lebesgue measure with $|f(x)| \leq M$ for all $x \in E$. Then there exists a sequence of step functions $\{\phi_n\}_{n\geq 1}$, supported on E with $|\phi_n(x)| \leq M$ for all $n \geq 1$ and all $x \in E$ and such that $\phi_n(x) \to f(x)$ as $n \to \infty$ for a.e. $x \in E$.*

The proof, which combines Lebesgue's prescription for a sequence of simple functions approximating f with approximation of measurable sets by unions of disjoint intervals (Littlewood's first principle) can be found, for example, in Kurtz & Swartz. Of course in general measure spaces we don't have a natural analogue of step functions and so this result does not generalise.

The following proposition is straightforward.

Proposition 5.15. *Suppose that f and g are bounded measurable functions supported on a set of finite measure. Then the following properties hold:*

1. Linearity. *If $a, b \in \mathbb{R}$ then*
$$\int (af + bg)(x)dx = a\int f(x)dx + b\int g(x)dx.$$
2. Additivity. *If E and F are* disjoint *subsets of $\mathbb{R}$ with finite measure then*
$$\int_{E\cup F} f(x)dx = \int_E f(x)dx + \int_F f(x)dx.$$
3. Monotonicity. *If $f \leq g$ then*
$$\int f(x)dx \leq \int g(x)dx.$$
4. Triangle inequality.
$$\left|\int f(x)dx\right| \leq \int |f(x)|dx.$$

Now we have our integral for bounded measurable functions supported on sets of bounded measure. Before moving on to Step 3, integration of non-negative measurable functions, we prove our first convergence theorem.

Theorem 5.16 (The Bounded Convergence Theorem)**.** *Let $\{f_n\}_{n\geq 1}$ be a sequence of measurable functions defined on a set E of finite measure and suppose that there is a real number M such that $|f_n(x)| \leq M$ for all n and all $x \in E$. If $f(x) = \lim_{n\to\infty} f_n(x)$ for each $x \in E$ then*
$$\int_E f(x)dx = \lim_{n\to\infty}\int_E f_n(x)dx.$$

Proof. Note that f is measurable (why?) and so the statement of the result makes sense. The conclusion of the Theorem would be trivial if f_n were to converge *uniformly* to f, but we don't have that. On the other hand, Egorov's Theorem says that it 'nearly does'. So given $\epsilon > 0$, $\exists A_\epsilon$ with $m(E\backslash A_\epsilon) < \epsilon$ and such that $f_n \to f$ uniformly on A_ϵ.

Now choose N large enough that for all $x \in A_\epsilon$ and $n \geq N$, $|f_n(x) - f(x)| < \epsilon$. Then for $n \geq N$
$$\begin{aligned}
\left|\int_E f_n(x)dx - \int_E f(x)dx\right| &= \left|\int_E (f_n(x) - f(x))dx\right| \\
&\leq \int_E |f_n(x) - f(x)|dx \\
&= \int_{A_\epsilon} |f_n(x) - f(x)|dx + \int_{E\backslash A_\epsilon} |f_n(x) - f(x)|dx \\
&< \epsilon m(A_\epsilon) + 2Mm(E\backslash A_\epsilon) \\
&\leq \epsilon\left(m(E) + 2M\right).
\end{aligned}$$

Since ϵ was arbitrary we see that

$$\left| \int_E f_n(x)dx - \int_E f(x)dx \right| \to 0 \quad \text{as } n \to \infty$$

as required. $\square$

Exercise 5.17. *Show that we could relax the condition* $f(x) = \lim_{n\to\infty} f_n(x)$ *in Theorem 5.16 to* $f(x) = \lim_{n\to\infty} f(x)$ *for a.e.* $x \in E$.

WARNING: Crucial to our proof was that $\{f_n\}$ were *bounded*. The function in Example 1.7 shows that the result can fail without this condition.

We can already make a fair amount of progress with the Bounded Convergence Theorem.

Example 5.18. *Find*

$$\sum_{k=1}^{\infty} \frac{1}{k(k+2)}.$$

Solution. *This is an illustration of the technique, it is far from the easiest way to calculate the sum.*

We're going to proceed formally and then go back to justify each step.

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{1}{k(k+2)} &= \frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{1}{k} - \frac{1}{k+2}\right) \\
&= \frac{1}{2}\sum_{k=1}^{\infty}\int_0^1 \left(x^{k-1} - x^{k+1}\right) dx \\
&= \frac{1}{2}\sum_{k=1}^{\infty}\int_0^1 x^{k-1}\left(1 - x^2\right) dx \\
&= \frac{1}{2}\int_0^1 \sum_{k=1}^{\infty} x^{k-1}\left(1 - x^2\right) dx \\
&= \frac{1}{2}\int_0^1 \frac{(1-x^2)}{(1-x)} dx \\
&= \frac{1}{2}\int_0^1 (1+x)dx = \frac{3}{4}.
\end{aligned}$$

To do this properly we hvae to justify the interchange of integration and summation in passing from the third to the fourth line. So set

$$f_n(x) = \sum_{k=1}^{n} x^{k-1}(1-x^2).$$

This defines a sequence of bounded measurable functions converging to $(1+x)$ on $[0,1]$ as $n \to \infty$ so the Bounded Convergence Theorem gives the required justification. $\square$

Example 5.19 (Euler's constant)**.** *Let*

$$u_n(x) = \frac{x}{n(x+n)}, \qquad x \in [0,1].$$

By considering $\int \sum_{n=1}^{N} u_n(x)dx$ *deduce the existence of*

$$\lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{n}-\log(N+1)\right).$$

The limit is called Euler's constant.

Solution. First note that

$$\sum_{n=1}^{N}\frac{x}{n(x+n)}\leq\sum_{n=1}^{N}\frac{1}{n^2}\leq\sum_{n=1}^{\infty}\frac{1}{n^2},\quad\text{for all } x\in[0,1]$$

and since $\displaystyle\sum_{n=1}^{N}\frac{x}{n(x+n)}$ is monotone increasing as N increases (it is a sum of positive terms) and bounded above it must converge as $N\to\infty$ to a limit. Now the Bounded Convergence Theorem tells us that

$$\int_0^1\lim_{N\to\infty}\sum_{n=1}^{N}u_n(x)dx=\lim_{N\to\infty}\sum_{n=1}^{N}\int_0^1 u_n(x)dx.$$

The left hand side is finite. We manipulate the right hand side:

$$u_n(x)=\frac{x}{n(x+n)}=\frac{1}{n}-\frac{1}{x+n}$$

and so

$$\int_0^1 u_n(x)dx=\frac{1}{n}-[\log(x+n)]_0^1=\frac{1}{n}-\log\left(\frac{n+1}{n}\right)$$

and

$$\sum_{n=1}^{N}\int_0^1 u_n(x)dx=\sum_{n=1}^{N}\left(\frac{1}{n}-\log\left(\frac{n+1}{n}\right)\right)=\sum_{n=1}^{N}\frac{1}{n}-\log\left(\prod_{n=1}^{N}\frac{n+1}{n}\right)=\sum_{n=1}^{N}\frac{1}{n}-\log(N+1).$$

Now apply the Bounded Convergence Theorem to recover the result. □

Our Example 1.7 with an *unbounded*, albeit convergent, sequence of measurable functions on $[0,1]$ tells us that we mustn't cheat - we crucially used boundedness of both the function and the measure of its support in our proof. Now we'd like to go beyond this to integrate functions where one or both of these assumptions is relaxed.

STEP 3. Integration of non-negative measurable functions.

We're going to consider non-negative functions, but we allow them to be 'extended-real-valued', that is we allow the value $+\infty$, provided that the set where they are infinite is measurable. Recall the convention that the supremum of an unbounded set is $+\infty$.

Definition 5.20 (Extended Lebesgue integral)**.** *For a non-negative measurable function f on a measurable subset $E\subseteq\mathbb{R}$ we define its* extended *Lebesgue integral by*

$$\int_E f(x)dx=\sup\left\{\int_E g(x)dx : g \textit{ bounded, measurable, } m(\mathrm{supp(g)})<\infty, 0\leq \mathrm{g}\leq \mathrm{f}\right\}.$$

If the supremum is finite then we say that f is Lebesgue integrable *over E and we write $f\in L^1(E)$ or $f\in L^1(E,m)$.*

Remark 5.21 (Advanced Warning)**.** *We'll talk about $L^1(E)$ in a bit more detail in §8. Because two integrable functions which differ only on a set of measure zero are indistinguishable from the point of view of integration - no matter which subset of E we integrate them over, the integrals will be equal - when we talk about points in $L^1(E)$ we don't really mean a function, but rather an* equivalence class *of functions which differ from one another only on a set of measure zero. However, we shall loosely use the notation $f \in L^1(E)$ to mean that the function f is Lebesgue integrable over E.*

Example 5.22. *Let*

$$F_a(x) = \frac{1}{1+|x|^a}, \quad x \in \mathbb{R}.$$

Then $F_a \in L^1(\mathbb{R})$ if and only if $a > 1$.

Proof. Consider

$$g_n(x) = \frac{1}{1+|x|^a}\mathbf{1}_{[-n,n]}(x).$$

g_n is bounded, measurable and its support has bounded measure. Moreover, $0 \le g_n(x) \le F_a(x)$ for all x, so

$$\int F_a(x)dx \ge \sup_{n\in\mathbb{N}} \int g_n(x)dx.$$

In particular, if the right hand side is infinite then so is the left. On the other hand, for any g with g bounded and measurable with support of bounded measure and $g \le F_a$, then given $\epsilon > 0$, since $F_a(x) \to 0$ as $|x| \to \infty$, there exists n so that $\int_{[-n,n]^c} g(x) < \epsilon$ and then, for this n, $\int g(x)dx \le \int g_n(x)dx + \epsilon$. Thus rather than considering all such g we find that

$$\int F_a(x)dx = \sup_{n\in\mathbb{N}} \int g_n(x)dx.$$

Suppose then that $a \neq 1$. For $a > 1$ note that

$$\begin{aligned}
\int_{-n}^{n} \frac{1}{1+|x|^a}dx &= 2\int_0^n \frac{1}{1+x^a}dx \\
&\le 2\int_0^1 1dx + 2\int_1^n \frac{1}{x^a}dx \\
&= 2 + 2\left[-\frac{1}{a-1}\frac{1}{x^{a-1}}\right]_1^n \\
&= 2 + \frac{2}{a-1} - \frac{2}{a-1}\frac{1}{n^{a-1}}.
\end{aligned}$$

Taking the supremum over the right hand side we obtain $2a/(a-1) < \infty$ so F_a is integrable if $a > 1$.

Now suppose that $a < 1$.

$$\begin{aligned}
\int_{-n}^{n} \frac{1}{1+|x|^a}dx &= 2\int_0^n \frac{1}{1+x^a}dx \\
&\ge 2\int_0^1 \frac{1}{2}dx + 2\int_1^n \frac{1}{2x^a}dx \\
&= 1 + \left[-\frac{1}{a-1}\frac{1}{x^{a-1}}\right]_1^n \\
&= 1 + \frac{1}{a-1} - \frac{1}{a-1}\frac{1}{n^{a-1}}.
\end{aligned}$$

This time, since $a < 1$, the right hand side diverges to infinity as $n \to \infty$ so F_a is *not* integrable over $\mathbb{R}$ for $a < 1$.

For $a = 1$,

$$\begin{aligned}\int_{-n}^{n} \frac{1}{1+|x|}dx &= 2\int_{0}^{n} \frac{1}{1+x}dx \\ &= 2\left[\log(1+x)\right]_0^n = 2\log(n+1) \to \infty \text{ as } n \to \infty.\end{aligned}$$

So F_1 is also not integrable over $\mathbb{R}$. □

Proposition 5.23. *Suppose that f and g are non-negative measurable functions. Then the integral has the following properties:*

1. Linearity. *If $a, b \in \mathbb{R}$ then*

$$\int (af + bg)(x)dx = a\int f(x)dx + b\int g(x)dx.$$

2. Additivity. *If E and F are* disjoint *subsets of $\mathbb{R}$ with finite measure then*

$$\int_{E\cup F} f(x)dx = \int_E f(x)dx + \int_F f(x)dx.$$

3. Monotonicity. *If $f \leq g$ then*

$$\int f(x)dx \leq \int g(x)dx.$$

4. Comparison. *If g is integrable and f is measurable with $0 \leq f \leq g$ then f is integrable.*

5. *If f is integrable then $f(x) < \infty$ for a.e. x.*

6. *If f is non-negative, measurable and $\int f(x)dx = 0$ then $f(x) = 0$ for a.e. x.*

Of these 4 is the one you end up using all the time. To decide whether or not a function is integrable we often try to dominate by something that is integrable or use it to dominate something that is not.

Example 5.24. *Is the function*

$$\frac{x - \sin x}{x(1+x^2)}$$

integrable over $[0, \infty)$?

Solution. Yes, since

$$0 \leq \frac{x - \sin x}{x(1+x^2)} \leq \frac{2}{1+x^2}$$

and the function on the right is integrable. □

Remark 5.25. *As always, there was nothing special about Lebesgue measure in Proposition 5.23. If instead we integrated against the counting measure, we'd recover familiar results from Mods analysis about series of non-negative terms.*

Another approach to determining whether or not a function is integrable is to apply one of the celebrated convergence theorems and we now turn to versions of those that apply to non-negative measurable functions.

Recall that in Example 1.7 we had a sequence of positive measurable functions f_n on $[0,1]$ converging a.e. to zero, but for which $\lim_{n\to\infty}\int f_n(x)dx = 1/2$ so that $\int \lim_{n\to\infty} f_n(x)dx < \lim_{n\to\infty}\int f_n(x)dx$. This is a special case of the cornerstone of the convergence theorems, the result rather modestly known as Fatou's Lemma. Everything will follow easily from this.

Recall from Definition 4.8 that for a sequence of real numbers $\{x_n\}_{n\in\mathbb{N}}$,

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty}\sup_{m\geq n} x_m, \quad \liminf_{n\to\infty} x_n = \lim_{n\to\infty}\inf_{m\geq n} x_m$$

and that the idea is that if $z < \liminf_{n\to\infty} x_n$ then 'eventually' $x_n > z$ and $\liminf_{n\to\infty} x_n$ is the largest possible value consistent with this so that if $z > \liminf_{n\to\infty} x_n$, then for any N there is an $n > N$ such that $x_n > z$.

Theorem 5.26 (Fatou's Lemma)**.** *Suppose that $\{f_n\}_{n\geq 1}$ is a sequence of measurable functions with $f_n \geq 0$. If $\lim_{n\to\infty} f_n(x) = f(x)$ for a.e. x then*

$$\int f(x)dx \leq \liminf_{n\to\infty}\int f_n(x)dx.$$

Proof. Suppose that $0 \leq g \leq f$ where g is bounded and supported on a set E of finite measure. Set

$$g_n(x) = \min\left(g(x), f_n(x)\right),$$

then g_n is measurable, supported on E and $g_n(x) \to g(x)$ a.e., so by the Bounded Convergence Theorem

$$\int g_n(x)dx \to \int g(x)dx.$$

By construction we also have that $g_n \leq f_n$ so that

$$\int g_n(x)dx \leq \int f_n(x)dx. \tag{7}$$

Now for a convergent sequence the limit and the liminf are the same, so

$$\int g(x)dx = \lim_{n\to\infty}\int g_n(x)dx = \liminf_{n\to\infty}\int g_n(x)dx$$

and so using equation (7) we obtain

$$\int g(x)dx \leq \liminf_{n\to\infty}\int f_n(x)dx.$$

Now by Definition 5.20 $\int f(x)dx$ is the supremum over g of the left hand side and the result follows. □

We can immediately deduce some important corollaries.

NOTE. In Corollary 5.27 we allow $\int f(x)dx = \infty$.

Corollary 5.27. *Suppose that f is a non-negative measurable function and $\{f_n\}_{n\geq 1}$ is a sequence of non-negative measurable functions with $f_n(x) \leq f(x)$ and $f_n(x) \to f(x)$ for a.e. x. Then*

$$\lim_{n\to\infty}\int f_n(x)dx = \int f(x)dx.$$

Proof. Since $f_n(x) \leq f(x)$ a.e. x we have that $\int f_n(x)dx \leq \int f(x)dx$ for all n and hence

$$\limsup_{n\to\infty} \int f_n(x)dx \leq \int f(x)dx.$$

But by Fatou's Lemma,

$$\int f(x)dx \leq \liminf_{n\to\infty} \int f_n(x)dx.$$

So combining the two

$$\liminf_{n\to\infty} f_n(x)dx \geq \int f(x)dx \geq \limsup_{n\to\infty} \int f_n(x)dx \geq \liminf_{n\to\infty} f_n(x)dx$$

so that in fact $\lim_{n\to\infty} \int f_n(x)dx$ exists and equals $\int f(x)dx$ as required. □

Corollary 5.28 (Monotone Convergence Theorem)**.** *Suppose that $\{f_n\}_{n\geq 1}$ is a non-decreasing sequence of non-negative measurable functions and M a constant independent of n for which*

$$\int f_n(x)dx < M < \infty \quad \forall n.$$

Then $\{f_n\}_{n\geq 1}$ converges a.e. to an integrable function f and

$$\int f(x)dx = \lim_{n\to\infty} \int f_n(x)dx.$$

Proof.

We *define* $f(x) = \lim_{n\to\infty} f_n(x)$. This function takes its values in the extended reals, but is well defined since $\{f_n(x)\}_{n\geq 1}$ is monotone increasing and f is measurable by Theorem 4.12.

By Corollary 5.27,

$$\int f_n(x)dx \to \int f(x)dx,$$

and since $\int f_n(x)dx < M$ for all n we conclude that $\int f(x)dx \leq M$. Thus by Proposition 5.23 (4), $f(x) < \infty$ for a.e. x. Changing f on a null set doesn't change the integral so define $f(x) = 0$ on $\{x : \lim_{n\to\infty} f_n(x) = \infty\}$ to complete the proof. □

From now on MCT=Monotone Convergence Theorem.

There are *lots* of useful consequences of the MCT. First it allows us to (rigorously and painlessly) extend our armoury of integrable and non-integrable functions with which to make comparisons.

Example 5.29. *For which values of a is $f(x) = 1/x^a$ integrable over $[0,1]$?*

Solution. Let

$$f_n(x) = \begin{cases} 1/x^a & x \in [1/n, 1] \\ 0 & \text{otherwise.} \end{cases}$$

Evidently $f_n \uparrow f$ as $n \to \infty$.

For $a < 1$,

$$\begin{aligned} \int_0^1 f_n(x)dx &= \left[-\frac{1}{a-1}\frac{1}{x^{a-1}}\right]_{1/n}^1 \\ &= \frac{1}{a-1}\left(n^{a-1} - 1\right) \to \frac{1}{1-a} \quad \text{as } n \to \infty. \end{aligned}$$

So by the MCT, $f(x) \in L^1([0,1])$ for $a < 1$.

For $a = 1$,

$$\int_0^1 f_n(x)dx = \log n \to \infty \quad \text{as } n \to \infty.$$

Since $\int f(x)dx \geq \int f_n(x)dx$ for all n we see that $f \notin L^1([0,1])$.

Similarly, for $a > 1$, $\int f(x)dx \geq \int f_n(x)dx \to \infty$ as $n \to \infty$ and so $f \notin L^1([0,1])$. $\square$

Corollary 5.30. *Consider the series $\sum_{k=1}^{\infty} a_k(x)$ where $a_k(x) \geq 0$ is measurable for every $k \geq 1$. Then*

$$\int \sum_{k=1}^{\infty} a_k(x)dx = \sum_{k=1}^{\infty} \int a_k(x)dx.$$

Moreover, if

$$\sum_{k=1}^{\infty} \int a_k(x)dx$$

is finite, then $\sum_{k=1}^{\infty} a_k(x)$ converges for a.e. x.

Proof.

Let $f_n(x) = \sum_{k=1}^{n} a_k(x)$ and $f(x) = \sum_{k=1}^{\infty} a_k(x)$. The functions f_n are measurable (since the a_k's are), $f_n(x) \leq f_{n+1}(x)$ (since the a_k's are non-negative) and $f_n(x) \to f(x)$ as $n \to \infty$. Since $\int f_n(x)dx = \sum_{k=1}^{n} \int a_k(x)dx$ Corollary 5.27 gives

$$\sum_{k=1}^{\infty} \int a_k(x)dx = \int \sum_{k=1}^{\infty} a_k(x)dx$$

and so if $\sum_{k=1}^{\infty} \int a_k(x)dx < \infty$ then $\sum_{k=1}^{\infty} a_k(x)$ is integrable and thus finite a.e. (*c.f* the proof of the MCT). $\square$

Example 5.31. *Evaluate*

$$\int_0^1 \left(\frac{\log x}{1-x}\right)^2 dx$$

using the technology of Corollary 5.30.

Solution. First note that

$$\frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x}\right)$$

and so writing $1/(1-x) = \sum_{k=0}^{\infty} x^k$ for $0 < x < 1$ and using Mods Analysis to justify differentiating this power series term by term (for $0 < x < 1$) we obtain

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}.$$

Set

$$a_k(x) = kx^{k-1}(\log x)^2,$$

then $a_k(x)$ is nonnegative and measurable and

$$\begin{aligned}\int_0^1 a_k(x)dx &= \left[x^k(\log x)^2\right]_0^1 - \int_0^1 2x^k\frac{\log x}{x}dx \\ &= -\int_0^1 2x^{k-1}\log x dx \\ &= \left[-\frac{2}{k}x^k\log x\right]_0^1 + \int_0^1 \frac{2}{k}x^{k-1}dx \\ &= \frac{2}{k^2}.\end{aligned}$$

So

$$\sum_{k=1}^{\infty}\int_0^1 a_k(x)dx = \sum_{k=1}^{\infty}\frac{2}{k^2} = \frac{\pi^2}{3}$$

and by Corollary 5.30

$$\int_0^1\left(\frac{\log x}{1-x}\right)^2 dx = \frac{\pi^2}{3}.$$

□

Remark 5.32. *Actually we cheated a bit. To calculate* $\int_0^1 a_k(x)dx$ *we integrated by parts with* $u(x) = (\log x)^2$ *and* $v(x) = x^k$*. The difficulty is that Theorem 1.9 requires* u *amd* v *to be continuously differentiable on the whole interval* $[0,1]$*, but our* $u(x)$ *is badly behaved at* $x = 0$ *so we don't have that here. It is an easy application of the MCT to justify our calculation (exercise).*

Example 5.33. *Evaluate*

$$\int_0^1 \log\left(\frac{1}{1-x}\right)dx.$$

Solution.

$$\log\left(\frac{1}{1-x}\right) = -\log(1-x) = \sum_{k=1}^{\infty}\frac{x^k}{k},$$

so set $a_k(x) = x^k/k$.

$$\int_0^1 a_k(x)dx = \frac{1}{k(k+1)},$$

so

$$\sum_{k=1}^{\infty}\int_0^1 a_k(x)dx = \sum_{k=1}^{\infty}\frac{1}{k(k+1)}.$$

To evaluate the right hand side, notice that

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$$

and so

$$\sum_{k=1}^{N}\frac{1}{k(k+1)} = \sum_{k=1}^{N}\left(\frac{1}{k} - \frac{1}{k+1}\right) = 1 - \frac{1}{N+1} \to 1 \quad \text{as } N \to \infty.$$

Thus Corollary 5.30 gives

$$\int_0^1 \log\left(\frac{1}{1-x}\right)dx = 1.$$

$\square$

We are still restricted to integrating non-negative functions, so now we rectify that.

STEP 4. The general case.

Definition 5.34 (Lebesgue integrable function)**.** *If f is* any *measurable real-valued function on $\mathbb{R}$, we say that f is* Lebesgue integrable *if the (measurable) non-negative function $|f|$ is Lebesgue integrable. Writing*

$$f^+(x) = \max(f(x), 0), \quad f^-(x) = \max(-f(x), 0)$$

we see that $0 \le f^+, f^- \le |f|$ and so when f is integrable these functions too are integrable and we define

$$\int f(x)dx = \int f^+(x)dx - \int f^-(x)dx.$$

The following result is readily deduced from the corresponding result for non-negative functions.

Proposition 5.35. *For Lebesgue integrable functions, the integral is linear, additive, monotonic and satisfies the triangle inequality.*

Notation. In keeping with what we did for non-negative functions, write $f \in L^1(\mathbb{R})$ if f is Lebesgue integrable over $\mathbb{R}$. Note that $f \in L^1(\mathbb{R})$ if and only if $|f| \in L^1(\mathbb{R})$. Similarly for a function defined on a measurable set $E \subseteq \mathbb{R}$ write $f \in L^1(E)$ if and only if $\int_E |f(x)|dx < \infty$.

Example 5.36. *1. Is*

$$f(x) = \frac{\sin x}{x(\log x)^2}$$

in $L^1([2,\infty))$?

2. Is

$$g(x) = \frac{\sin x}{x}$$

in $L^1([1,\infty))$?

Solution. Since $f \in L^1([2,\infty))$ if and only if $|f| \in L^1([2,\infty))$, we consider $|f|$. Now $|\sin x| \le 1$ and

$$\int_2^N \frac{1}{x(\log x)^2}dx = \left[-\frac{1}{\log x}\right]_2^N = \frac{1}{\log 2} - \frac{1}{\log N} \le \frac{1}{\log 2}.$$

So the MCT gives $1/(x(\log x)^2) \in L^1([2,\infty))$. Since $|f| \le 1/(x(\log x)^2)$ we have (by comparison) that $f \in L^1([2,\infty))$.

Now consider g. Again $g \in L^1([1,\infty))$ if and only if $|g|$ is and $|g| = |\sin x|/x$. Now

$$|\sin x| > \frac{1}{2} \text{ for } x \in \left(\frac{\pi}{3}, \frac{2\pi}{3}\right), \dots \left(n\pi + \frac{\pi}{3}, n\pi + \frac{2\pi}{3}\right), \quad \forall n \in \mathbb{N}$$

so

$$|g(x)| \ge \frac{1}{2}\sum_{n=1}^{\infty} \frac{1}{n\pi + 2\pi/3}\mathbf{1}_{(n\pi+\pi/3, n\pi+2\pi/3)}(x).$$

This function in turn dominates

$$\frac{1}{2}\sum_{n=1}^{N} \frac{1}{n\pi + 2\pi/3}\mathbf{1}_{(n\pi+\pi/3, n\pi+2\pi/3)}(x)$$

for all N which is integrable (as a simple function) with integral

$$\sum_{n=1}^{N} \frac{1}{\pi(n+2/3)} \frac{\pi}{3} \to \infty \quad \text{as } N \to \infty$$

(by comparison with $\sum 1/n$).

Thus $g \notin L^1([1,\infty))$. □

Now for the most important convergence theorem of them all.

Theorem 5.37 (The Dominated Convergence Theorem)**.** *Let g be integrable and let $\{f_n\}_{n\geq 1}$ be a sequence of measurable functions such that $|f_n| \leq g$. If $f(x) = \lim_{n\to\infty} f_n(x)$ for a.e. x then f is integrable with*

$$\int f(x)dx = \lim_{n\to\infty} \int f_n(x)dx.$$

Proof. We have that $g - f_n \geq 0$ and so by Fatou's Lemma

$$\int (g-f)(x) \leq \liminf_{n\to\infty} \int (g-f_n)(x)dx.$$

Since $|f| \leq g$, f is integrable and so we may use linearity of the integral (and the fact that for a sequence of points x_n, $\liminf_{n\to\infty}(-x_n) = -\limsup_{n\to\infty} x_n$) to rewrite this as

$$\int g(x)dx - \int f(x)dx \leq \int g(x)dx - \limsup_{n\to\infty} \int f_n(x)dx.$$

In other words

$$\int f(x)dx \geq \limsup_{n\to\infty} \int f_n(x)dx. \tag{8}$$

Similarly, $g + f_n \geq 0$ and so again Fatou's Lemma gives

$$\int (g+f)(x)dx \leq \liminf_{n\to\infty} \int (g+f_n)(x)dx$$

which upon rearrangement gives

$$\int f(x)dx \leq \liminf_{n\to\infty} \int f_n(x)dx. \tag{9}$$

Combining (8) and (9) we obtain

$$\limsup_{n\to\infty} \int f_n(x)dx \leq \int f(x)dx \leq \liminf_{n\to\infty} \int f_n(x)dx.$$

But $\liminf \leq \limsup$ and so in fact $\lim_{n\to\infty} \int f_n(x)dx$ exists and equals $\int f(x)dx$ as required. □

From now on DCT= Dominated Convergence Theorem.

We have now established very general conditions under which we can exchange the limit and the integral, but beware, we can't cheat, it is essential to check the conditions of one of the convergence theorems.

Example 5.38. *Let* $a_k(x) = x^{k-1} - 2x^{2k-1}$ *and* $f_n(x) = \sum_{k=1}^n a_k(x)$. *Then*

$$\int_0^1 a_k(x)dx = \int_0^1 (x^{k-1} - 2x^{2k-1})dx = \frac{1}{k} - \frac{2}{2k} = 0.$$

But

$$f(x) = \sum_{k=1}^{\infty}(x^{k-1} - 2x^{2k-1}) = \frac{1}{1-x} - \frac{2x}{1-x^2} = \frac{1}{1+x}.$$

So $\int_0^1 f_n(x)dx = 0$ *for all* n, *but* $\int_0^1 f(x)dx = \log 2$.

What went wrong? We didn't find a dominating function.

$$f_n(x) = \sum_{k=1}^{n}(x^{k-1} - 2x^{2k-1}) = \frac{1-x^n}{1-x} - 2x\frac{1-x^{2n}}{1-x^2}$$

and to apply the DCT we need to dominate $|f_n(x)|$ by an integrable function. Here f_n changes sign on $[0,1]$ and it is not hard to show (exercise) that $\int |f_n(x)|dx \to \infty$ as $n \to \infty$. The DCT does not apply.

On the other hand, we have the following corollary to the DCT.

Corollary 5.39 (Beppo Levi)**.** *Suppose that*

$$\sum_{k=1}^{\infty} \int |a_k(x)|dx < \infty.$$

Then $\sum_{k=1}^{\infty} a_k(x)$ *converges for a.e.* x, *its sum is integrable and*

$$\int \sum_{k=1}^{\infty} a_k(x)dx = \sum_{k=1}^{\infty} \int a_k(x)dx.$$

Proof. First apply Corollary 5.30 to the MCT to the positive functions $|a_k(x)|$ to see that $\sum_{k=1}^{\infty} |a_k(x)|$ is integrable and now apply the DCT with dominating function $g(x) = \sum_{k=1}^{\infty} |a_k(x)|$. □

Example 5.40. *For which values of* $a \in \mathbb{R}$ *does the power series* $\sum_{k=1}^{\infty} k^a x^k$ *define an integrable function on* $[-1,1]$?

Solution.

$$\int_{-1}^1 |k^a x^k|dx = k^a \int_{-1}^1 |x|^k dx = 2k^a \int_0^1 x^k dx = \frac{2k^a}{k+1}.$$

If $a < 0$ the series $\sum 2k^a/(k+1)$ converges by comparison with $\sum k^{a-1}$ so, by Corollary 5.39, $\sum_{k=0}^{\infty} k^a x^k$ is integrable.

If $a = 0$, the series sums to $x/(1-x)$ which is not integrable over $[-1,1]$, since if it were then $|x/(1-x)|$ would be integrable over $[-1,1]$, but then for each $N \geq 1$

$$\int_{-1}^1 \left|\frac{x}{1-x}\right| dx \geq \int_{1/2}^{1-1/N} \frac{x}{1-x}dx \geq \frac{1}{2}\int_{1/2}^{1-1/N} \frac{1}{1-x}dx = \frac{1}{2}[-\log(1-x)]_{1/2}^{1-1/N}$$
$$= \frac{1}{2}\log(\frac{N}{2}) \to \infty \text{ as } N \to \infty.$$

Finally, for $a > 0$, for $x \in [0,1]$, $\sum_{k=1}^{\infty} k^a x^k > \sum_{k=1}^{\infty} x^k$ and so by comparison with the case $a = 0$ the series is not integrable. □

Example 5.41. *Show that*

$$\int_0^\infty \frac{x}{e^x - 1} dx = \frac{\pi^2}{6}.$$

Solution. Write

$$\frac{x}{e^x - 1} = \frac{xe^{-x}}{1 - e^{-x}} = x \sum_{k=1}^{\infty} e^{-kx}.$$

Set $a_k(x) = xe^{-kx}$ in Corollary 5.30 to the MCT.

$$\int_0^\infty a_k(x) dx = \int_0^\infty xe^{-kx} dx = \left[-\frac{x}{k} e^{-kx}\right]_0^\infty + \int_0^\infty \frac{1}{k} e^{-kx} dx = \frac{1}{k^2}.$$

And since $\sum_{k=1}^\infty 1/k^2 < \infty$ the Corollary gives

$$\int_0^\infty \frac{x}{e^x - 1} dx = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$$

as required. □

Of course if I were really pedantic I'd have checked that $a_k \in L^1([0, \infty))$ for each k. To do this note that xe^{-x} is integrable over $[0, N]$ for all N and apply the MCT.

Example 5.42. *Show that if* $p > 0$ *then*

$$\int_1^\infty \frac{\log x}{x^p(x - 1)} dx = \sum_{k=0}^{\infty} \frac{1}{(k + p)^2}.$$

Solution. For $x > 1$ we may write

$$\frac{\log x}{x^p(x - 1)} = \frac{\log x}{x^{p+1}(1 - 1/x)} = \frac{\log x}{x^{p+1}} \sum_{k=0}^{\infty} \frac{1}{x^k}.$$

We'd like to set

$$a_k(x) = \frac{\log x}{x^{p+k+1}} \mathbf{1}_{[1,\infty)}(x)$$

in Corollary 5.30 to the MCT. Evidently these functions are positive, but we must check that they are integrable. For each $N \geq 1$ and $p, k \geq 0$ we have

$$\begin{aligned}
\int_1^N \frac{\log x}{x^{p+k+1}} dx &= \left[-\frac{1}{p+k} \frac{\log x}{x^{p+k}}\right]_1^N + \int_1^N \frac{1}{p+k} \frac{1}{x^{p+k+1}} dx \\
&= -\frac{1}{p+k} \frac{\log N}{N^{p+k}} + \frac{1}{(p+k)^2} \left(1 - \frac{1}{N^{p+k}}\right) \\
&\to \frac{1}{(p+k)^2} \text{ as } N \to \infty.
\end{aligned}$$

So using that $a_k(x)$ is non-negative on $[1, \infty)$ the MCT tells us that it is $\in L^1([1, \infty))$ and then Corollary 5.30 gives us that

$$\int_1^\infty \frac{\log x}{x^p(x - 1)} dx = \sum_{k=0}^{\infty} \frac{1}{(p + k)^2}$$

as required. □

Aside: Complex-valued functions.

If f is a complex-valued funtion on $\mathbb{R}^d$ we may write it as

$$f(x) = u(x) + iv(x)$$

where u and v are real-valued functions. We say that f is Lebesgue integrable if the real-valued function $|f(x)| = \left(u(x)^2 + v(x)^2\right)^{1/2}$ is Lebesgue integrable.

Since $(a+b)^{1/2} \leq a^{1/2} + b^{1/2}$ for $a, b \geq 0$ we have $|f(x)| \leq |u(x)| + |v(x)|$ so to check that f is Lebesgue integrable it is enough to check that each of u and v is Lebesgue integrable and then

$$\int f(x)dx = \int u(x)dx + i \int v(x)dx.$$

This can be useful in calculations involving trigonometric functions.

Example 5.43. *Show that for each $a \in \mathbb{R}$, $e^{-x}\sin(ax) \in L^1([0,\infty))$ and establish the value of $\int_0^\infty e^{-x}\sin(ax)dx$.*

Solution. Since $|e^{-x}\sin(ax)| \leq e^{-x}$ comparison immediately gives that $e^{-x}\sin(ax)$ is in $L^1([0,\infty))$.

Now write $\sin(ax) = \operatorname{Im} \mathrm{e}^{\mathrm{iax}}$. Then

$$\int_0^\infty e^{-x}\sin(ax)dx = \operatorname{Im}\int_0^\infty \mathrm{e}^{-\mathrm{x}+\mathrm{iax}}\mathrm{dx} = \operatorname{Im}\int_0^\infty \mathrm{e}^{-\mathrm{x}(1-\mathrm{ia})}\mathrm{dx} = \operatorname{Im}\frac{1}{1-\mathrm{ia}} = \operatorname{Im}\frac{1+\mathrm{ia}}{1+\mathrm{a}^2} = \frac{\mathrm{a}}{1+\mathrm{a}^2}.$$

Similarly,

$$\int_0^\infty e^{-x}\cos(ax)dx = \frac{1}{1+a^2}.$$

□

TAKING STOCK

We've covered a lot of ground so let's take stock of what we have learned so far.

We have extended our notion of integral from Mods so that for a *measurable* function we can now decide whether or not it is integrable. Key tools are:

1. Everything you learned in Mods (and school) about integrating bounded continuous functions over bounded intervals and manipulation of these for example through substitution or integration by parts.

2. Comparison.

3. The Monotone Convergence Theorem.

4. The Dominated Convergence Theorem.

Note that the bounded convergence theorem is a special case of the DCT where the dominating function is a constant multiple of the indicator function of a set of finite measure.

In the exams:

You may assume that you can integrate a *bounded* continuous function over a *bounded* interval without comment.

You may assume the results of Mods that allow you to make substitutions and integrate by parts for continuous functions over bounded intervals.

If you use the MCT/DCT *state them clearly and make it clear that you have checked that the conditions of the Theorem are satisfied.*

Be sure that you have at your disposal a good stock of functions whose integrability/non-integrability you know about for use in comparisons. For example e^{-x}, x^a, $1/(1+|x|^a)$ over various intervals.

It is also useful to have a couple of sequences of functions to try as 'counterexamples'. Standard ones are the functions from Example 1.7 and the sequence

$$g_n(x) = \begin{cases} 0 & x \leq n, \\ 2(x-n) & n < x \leq n+1/2, \\ 2(1-(x-n)) & n+1/2 < x \leq n+1 \\ 0 & x > n = 1. \end{cases}$$

NOTE. Very little of what we have done is specific to the Lebesgue integral on $\mathbb{R}$. Any other measure space would have done equally well, but of course the *null sets* will depend crucially on the measure. For example, sets consisting of single points are not null for the counting measure.

Improper integrals.

It is important to remember that $f \in L^1$ if and *only* if $|f| \in L^1$. (This definition ensures that L^1 functions form what is called a *lattice* - an important concept in more advanced analysis courses.) It is analogous to asking that a series be *absolutely* convergent. As you saw in Mods there are many advantages to restricting our attention to absolutely convergent series, but there were other series for which we had a notion of convergence too. For example $\sum_{n=1}^{\infty}(-1)^n/n$ is convergent, but not *absolutely* convergent.

It is natural to ask if one can define a theory of integration that drops the 'absolute'. The answer is yes. *Improper Riemann integrals* have this property. Two types of improper integral arise. The first is relevant when a function has asymptotes on a bounded interval. For example if the function f has a singularity at a but is well-defined on $(a, b]$ one defines

$$\mathcal{R}\int_a^b f(x)dx = \lim_{\epsilon \downarrow 0} \int_{a+\epsilon}^b f(x)dx$$

(with a similar definition if the singularity is at the other limit of integration). The second is to deal with integration over unbounded regions. We define

$$\mathcal{R}\int_{-\infty}^{\infty} f(x)dx = \lim_{a \to -\infty, b \to \infty} \int_a^b f(x)dx \tag{10}$$

if the limit exists.

You have probably been implicitly using this kind of integral since school. But some words of warning are in order.

Example 5.44. *Let*

$$f(x) = \begin{cases} \frac{(-1)^n}{n+1} & x \in [n, n+1), n \in \mathbb{N} \\ 0 & x < 0. \end{cases}$$

Then

$$\mathcal{R}\int_0^{\infty} f(x)dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$$

clearly exists, but

$$\int_0^{\infty} |f(x)|dx$$

does not *exist.*

If

$$\mathcal{R}\int_{-\infty}^{\infty}|f(x)|dx$$

does exist, then in fact $f \in L^1(\mathbb{R})$ and

$$\mathcal{R}\int_{-\infty}^{\infty}f(x)dx = \int_{-\infty}^{\infty}f(x)dx.$$

There's a weaker form still of the integral, called the *Principal Value* of the integral and defined by

$$PV\int_{-\infty}^{\infty}f(x)dx = \lim_{a\to\infty}\int_{-a}^{a}f(x)dx.$$

This integral can exist even when the improper Riemann integral does not.

Example 5.45. *Let $f(x) = x$. Then*

$$PV\int_{-\infty}^{\infty}f(x)dx = 0$$

but

$$\int_{a}^{b}f(x)dx = b^2 - a^2$$

which can take any value and so in particular the limit in (10) does not exist.

When you calculate integrals using the familiar semicircular contours of complex analysis, it is the *principal value* that you are calculating. However, provided you can check that the function really is in L^1 then the Principal Value and the Lebesgue integral will agree. Complex Analysis remains a powerful tool for calculating integrals, but you must check that the function you are integrating really is in L^1 or you may just be calculating $PV\int$ or $\mathcal{R}\int$.

Of course alternating series are 'improper' integrals when we take the counting measure instead of Lebesgue measure. This suggests that we must be very careful in manipulating improper integrals - if we 'shuffle' an alternating series which is convergent but not absolutely convergent then we can make it sum to any given value. Our next example (which introduces our next topic) illustrates this in a rather different context.

6 The Theorems of Fubini and Tonelli

Our next topic is *multiple integrals*. In elementary calculus, integrals of continuous functions of severable variables are often calculated by iterating one-dimensional integrals. But one has to be careful

Example 6.1. *Consider*

$$\int_0^1\left(\int_0^1\frac{x-y}{(x+y)^3}dy\right)dx = \int_0^1\left(\int_0^1\frac{y-x}{(y+x)^3}dx\right)dy = -\int_0^1\left(\int_0^1\frac{x-y}{(x+y)^3}dx\right)dy$$

and so if the value of the double integral is to be independent of the order of integration, it must be zero. But

$$\begin{aligned}
\int_0^1 \left(\int_0^1 \frac{x-y}{(x+y)^3} dx \right) dy &= \int_0^1 \left(\int_0^1 \frac{1}{(x+y)^2} - \frac{2y}{(x+y)^3} dx \right) dy \\
&= \int_0^1 \left[-\frac{1}{x+y} + \frac{y}{(x+y)^2} \right]_{x=0}^{x=1} dy \\
&= \int_0^1 \left(-\frac{1}{1+y} + \frac{y}{(1+y)^2} \right) dy \\
&= \int_0^1 -\frac{1}{(1+y)^2} dy = \left[\frac{1}{1+y} \right]_0^1 = -\frac{1}{2}.
\end{aligned}$$

So in this example, we get different *answers depending upon the order in which we perform the integrals.*

What has gone wrong? Notice that the integrand $f(x,y) = (x-y)/(x+y)^3$ is unbounded near the origin and moreover it changes sign. In fact (exercise) if we try to integrate $|f|$ then we get infinity. The function is *not* Lebesgue integrable and although the change in signs results in a finite value when we perform the double integrals, much like an alternating series it is very sensitive to manipulation.

Of course to define the Lebesgue integral for real-valued functions on $\mathbb{R}^2$ (or indeed $\mathbb{R}^d$ for any $d \geq 2$) all we need is a definition for the Lebesgue measure on $\mathbb{R}^d$. (Our definition of the integral depended only on having a definition of the measure against which we were integrating.) To define such a measure we simply mimic what we did for $\mathbb{R}$ except that now our outer measure will be based not on covers of sets with *intervals* but rather with products of intervals (in other words *rectangles*).

The Lebesgue measurable sets will be those obtained from products of intervals by countable numbers of set operations (complements, unions, intersections) plus their unions with null sets. And, as you'd expect, the Lebesgue measure of the box $[a,b] \times [c,d]$ will be $(b-a) \cdot (d-c)$. Integrals of *non-negative* functions can be found by iterated integration. This can be checked by a passage to the limit. The most practical way to check for integrability (or otherwise) of functions on $\mathbb{R}^2$ is provided by the theorems of Fubini and Tonelli. We write $\mathbb{R}^*$ for the extended reals.

Theorem 6.2 (Fubini)**.** *Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^*$ be Lebesgue integrable. Then*

1. *The 'x-slice' function $f_x : \mathbb{R} \to \mathbb{R}$ defined by $f_x(y) = f(x,y)$ is Lebesgue integrable for a.e. $x \in \mathbb{R}$. Similarly, the y-slice function defined by $f_y(x) = f(x,y) \in L^1(\mathbb{R})$ for a.e. y.*

2. *The functions $x \mapsto \int_{\mathbb{R}} f_x(y)dy = \int_{\mathbb{R}} f(x,y)dy$ and $y \mapsto \int_{\mathbb{R}} f_y(x)dx = \int_{\mathbb{R}} f(x,y)dx$ are Lebesgue integrable over $\mathbb{R}$.*

3. $$\int_{\mathbb{R}\times\mathbb{R}} f(x,y)dxdy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f_x(y)dy \right) dx = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)dy \right) dx = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)dx \right) dy.$$

Fubini's Theorem can be used to show that a function is *not* integrable. On the other hand Tonelli's Theorem goes the other way.

Theorem 6.3 (Tonelli)**.** *Let $f(x,y)$ be measurable on $\mathbb{R}^2$. Then if one of*

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x,y)|dy \right) dx, \quad \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x,y)|dx \right) dy$$

exists, then so does the other, $f \in L^1(\mathbb{R}^2)$ and Fubini's Theorem applies.

So now we can check that a function *is* integrable.

Both Fubini's Theorem and Tonelli's Theorem are *much* more general. They not only extend to $\mathbb{R}^d$ (in the obvious way) but also to integration over products of other measure spaces. We don't pursue this here.

We are not going to prove either of these theorems. The basic approach is to approximate via simple functions based on 'rectangles'. Instead we'll see that reversing the order of integration can have some useful implications.

Example 6.4. *Consider*

$$f(x,y) = \begin{cases} xe^{-x^2(1+y^2)} & x, y \geq 0, \\ 0 & \textit{otherwise.} \end{cases}$$

This function is positive, so to decide whether or not it is integrable I can just evaluate one of the repeated integrals $\int_0^\infty \left(\int_0^\infty f(x,y)dx\right) dy$, $\int_0^\infty \left(\int_0^\infty f(x,y)dy\right) dx$. Notice that

$$\begin{aligned} \int_0^\infty f(x,y)dx &= \int_0^\infty xe^{-x^2(1+y^2)}dx \\ &= \left[-\frac{1}{2(1+y^2)}e^{-x^2(1+y^2)}\right]_{x=0}^\infty \\ &= \frac{1}{2(1+y^2)}. \end{aligned}$$

Thus

$$\int_0^\infty \left(\int_0^\infty |f(x,y)|dx\right) dy = \int_0^\infty \frac{1}{2(1+y^2)}dy = \frac{\pi}{4}.$$

So certainly $f \in L^1(\mathbb{R}^2)$ by Tonelli's Theorem and by Fubini we must get the same answer if we integrate in the other order. Now

$$\begin{aligned} \int_0^\infty xe^{-x^2(1+y^2)}dy &= xe^{-x^2}\int_0^\infty e^{-x^2y^2}dy \\ &= xe^{-x^2}\int_0^\infty e^{-u^2}\frac{du}{x} \\ &= e^{-x^2}\int_0^\infty e^{-u^2}du. \end{aligned}$$

Thus

$$\begin{aligned} \frac{\pi}{4} &= \int_0^\infty \int_0^\infty f(x,y)dydx \\ &= \int_0^\infty \left(\int_0^\infty xe^{-x^2(1+y^2)}dy\right) dx \\ &= \int_0^\infty \left(e^{-x^2}\int_0^\infty e^{-u^2}du\right) dx \\ &= \int_0^\infty e^{-u^2}du \int_0^\infty e^{-x^2}dx = \left(\int_0^\infty e^{-x^2}dx\right)^2. \end{aligned}$$

Thus

$$\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}.$$

□

Sometimes we can spot clever ways to turn a single integral into a double integral and reverse the order of integration to obtain useful information.

Example 6.5. *Suppose that $b > a > 0$. Evaluate*

$$\int_0^\infty \frac{e^{-ax} - e^{-bx}}{x} dx.$$

Solution. First let's check that the integral *exists*. Since $b > a$, $f(x) > 0$ on $x > 0$. Now on $[1, \infty)$, $0 \leq f(x) \leq e^{-ax}$, so $f \in L^1([1, \infty))$ by comparison. On the other hand on $[0, 1]$, $f(x)$ is bounded and continuous (notice that $(e^{-ax} - e^{-bx})/x \approx (b - a)$ as $x \to 0$) and so $f \in L^1([0, 1])$. Combining these we have that $f \in L^1([0, \infty))$.

Now we rewrite the single integral as a double integral. Observe that

$$\int_a^b e^{-xy} dy = \left[-\frac{1}{x}e^{-xy}\right]_a^b = \frac{e^{-ax} - e^{-bx}}{x}.$$

Thus

$$\int_0^\infty \frac{e^{-ax} - e^{-bx}}{x} dx = \int_0^\infty \left(\int_a^b e^{-xy} dy\right) dx.$$

We have checked that the double integral is finite and since $e^{-xy} \geq 0$ Tonelli's Theorem tells us that e^{-xy} is in $L^1([0, \infty) \times [a, b])$ and so we may apply Fubini's Theorem to reverse the order of integration to obtain

$$\begin{aligned}
\int_0^\infty \frac{e^{-ax} - e^{-bx}}{x} dx &= \int_a^b \left(\int_0^\infty e^{-xy} dx\right) dy \\
&= \int_a^b \frac{1}{y} dy = \log(b/a).
\end{aligned}$$

□

Example 6.6. *Some of the most important so-called* special functions *are the Bessel functions (introduced by Daniel Bernoulli). They arise as solutions to Bessel's differential equation*

$$x^2 \frac{d^2 y}{dx^2} + x\frac{dy}{dx} + (x^2 - \alpha^2)y = 0$$

where the parameter α can be real or complex. The most common case is when α is an integer, in which case it is refered to as the order *of the Bessel function. Bessel functions of the first kind behave well at the origin, those of the second don't.*

The Bessel function of the first kind of order zero can be written in many different ways, but Bessel wrote it as

$$J_0(x) = \frac{2}{\pi} \int_0^{\pi/2} \cos(x \cos\theta) d\theta.$$

Notice that this defines a bounded continuous function, but of course the equation is not very explicit. Often we recognise it through its Laplace transform *defined to be*

$$\overline{J}_0(a) = \int_0^\infty J_0(x) e^{-ax} dx \quad a > 0.$$

Calculate this.

Solution. Since $|J_0(x)| \leq 1$ (and J_0 is continuous, so measurable) the Laplace transform obviously *exists* for $a > 0$ by comparison with e^{-ax}, that is

$$\int_0^\infty \left(\int_0^{\pi/2} \frac{2}{\pi} |\cos(x\cos\theta) e^{-ax}| d\theta \right) dx < \infty$$

and so we can apply Tonelli and consequently Fubini's Theorem to obtain:

$$\begin{aligned}
\int_0^\infty J_0(x) e^{-ax} dx &= \int_0^\infty \left(\int_0^{\pi/2} \frac{2}{\pi} \cos(x\cos\theta) e^{-ax} d\theta \right) dx \\
&= \int_0^{\pi/2} \left(\int_0^\infty \frac{2}{\pi} \cos(x\cos\theta) e^{-ax} dx \right) d\theta \\
&= \int_0^{\pi/2} \left(\mathrm{Re} \int_0^\infty \frac{2}{\pi} e^{-ax+ix\cos\theta} dx \right) d\theta \\
&= \int_0^{\pi/2} \left(\frac{2}{\pi} \mathrm{Re} \frac{1}{a - i\cos\theta} \right) d\theta \\
&= \int_0^{\pi/2} \frac{2}{\pi} \frac{a}{a^2 + \cos^2\theta} d\theta.
\end{aligned}$$

Now let $t = \tan\theta$ so that $dt = (1+t^2)d\theta$ and $\cos^2\theta = 1/(1+t^2)$. Then the integral becomes

$$\begin{aligned}
\int_0^\infty \frac{2}{\pi} \frac{a}{a^2 + 1/(1+t^2)} \frac{1}{1+t^2} dt &= \frac{2a}{\pi} \int_0^\infty \frac{1}{a^2 + 1 + a^2 t^2} dt \\
&= \frac{2a}{\pi} \frac{1}{a^2+1} \int_0^\infty \frac{1}{1 + (a^2/(a^2+1)t^2} dt.
\end{aligned}$$

Finally set $u = \sqrt{a^2/(a^2+1)}t$ to obtain

$$\overline{J}_0(a) = \frac{2}{\pi} \int_0^\infty \frac{1}{1+u^2} du \frac{a}{a^2+1} \sqrt{\frac{a^2+1}{a^2}} = \frac{1}{\sqrt{a^2+1}}.$$

□

Fubini's Theorem also provides an alternative approach to integration by parts.

Example 6.7 (Integration by parts revisited)**.** *Suppose that $f, g \in L^1([a,b])$ and for an arbitrary fixed value of $c \in [a,b]$ define*

$$F(x) = \int_c^x f(t)dt, \quad G(x) = \int_c^x g(t)dt.$$

By integrating $f(s)g(t)$ over the region $a \leq s \leq t \leq b$ recover the integration by parts formula.

Solution.
We are asked to integrate $f(s)g(t)$ over $a \leq s \leq t \leq b$ (the shaded region in the picture). So write $k(s,t) = \mathbf{1}_{a \leq s \leq t \leq b}$ and consider $f(s)g(t)k(s,t)$ on $[a,b] \times [a,b]$.
First notice that

$$\begin{aligned}
\int_a^b \left(\int_a^b |f(s)g(t)k(s,t)| ds \right) dt &\leq \int_a^b \left(\int_a^b |f(s)| \cdot |g(t)| ds \right) dt \\
&= \int_a^b |f(s)| ds \int_a^b |g(t)| dt
\end{aligned}$$

which is finite since $f, g \in L^1([a,b])$ and so by Tonelli's Theorem $f(s)g(t)k(s,t) \in L^1([a,b]\times[a,b])$ and by Fubini's Theorem

$$\int_a^b \left(\int_a^b f(s)g(t)k(s,t)ds\right) dt = \int_a^b \left(\int_a^b f(s)g(t)k(s,t)dt\right) ds.$$

Writing this in more familiar form by taking k into the limits we obtain

$$\begin{aligned}\int_a^b \left(\int_a^t f(s)g(t)ds\right) dt &= \int_a^b \left(\int_s^b f(s)g(t)dt\right) ds \\ &= \int_a^b \left(\int_a^b f(s)g(t)dt\right) ds - \int_a^b \left(\int_a^s f(s)g(t)dt\right) ds \\ &= \int_a^b f(s)ds \int_a^b g(t)dt - \int_a^b f(s)\left(\int_a^s g(t)dt\right) ds.\end{aligned}$$

Substituting for F and G gives

$$\int_a^b (F(t)-F(a))\, g(t)dt = (F(b)-F(a))\,(G(b)-G(a)) - \int_a^b f(s)\,(G(s)-G(a))\, ds. \qquad (11)$$

The left hand side of this equation is

$$\int_a^b F(t)g(t)dt - F(a)\,(G(b)-G(a))$$

and

$$\int_a^b f(s)\,(G(s)-G(a))\, ds = \int_a^b f(s)G(s)ds - (F(b)-F(a))\, G(a),$$

so substituting in equation (11) we obtain

$$\int_a^b F(t)g(t)dt = F(b)G(b) - F(a)G(a) - \int_a^b f(s)G(s)ds$$

which is the usual integration by parts formula. □

As a last example, here is a useful observation for performing calculations in probability theory.

Example 6.8. *Recall that for a* continuous *real-valued random variable X, the* density function *g is defined by*

$$\mathbb{P}[X \in A] = \int_A g(x)dx$$

for all measurable sets A. The density function is necessarily a non-negative function whose integral over $\mathbb{R}$ is one. The expected value *of X is*

$$\mathbb{E}[X] = \int_{\mathbb{R}} xg(x)dx$$

(if it exists). Often, the quantity that we actually model is $\mathbb{P}[X > x]$ for real values of x. (This is one minus the cumulative distribution function.*) Let us write*

$$G(x) = \mathbb{P}[X > x] = \int_x^\infty g(y)dy.$$

Now suppose that X is non-negative. *Then we have*

$$\begin{aligned}
\int_0^\infty xg(x)dx &= \int_0^\infty \left(\int_0^\infty \mathbf{1}_{y\leq x} g(x)dy\right)dx \\
&= \int_0^\infty \left(\int_0^\infty \mathbf{1}_{y\leq x} g(x)dx\right)dy \\
&= \int_0^\infty \left(\int_y^\infty g(x)dx\right)dy = \int_0^\infty G(y)dy.
\end{aligned}$$

Exercise 6.9. *Why did we take X non-negative?*

Remark 6.10. *Because Fubini works equally well for other measure spaces we can perform similar calculations for random variables that are not necessarily continuous. For example, for a* discrete *random variable again taking only* non-negative *values, this time non-negative integers,*

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^\infty k\mathbb{P}[X=k] \\
&= \sum_{k=1}^\infty \left(\sum_{j=1}^k 1\right)\mathbb{P}[X=k] \\
&= \sum_{k=1}^\infty \left(\sum_{j=1}^\infty \mathbf{1}_{j\leq k}\mathbb{P}[X=k]\right) \\
&= \sum_{j=1}^\infty \left(\sum_{k=1}^\infty \mathbf{1}_{j\leq k}\mathbb{P}[X=k]\right) \\
&= \sum_{j=1}^\infty \mathbb{P}[X\geq j] = \sum_{k=0}^\infty \mathbb{P}[X>k].
\end{aligned}$$

Before turning back to some more abstract mathematics, let's record and illustrate a couple more corollaries of the DCT which can be useful in performing calculations.

Integrals that depend on a parameter.

Often one is interested in integrals that depend on a parameter and sometimes examination of the dependence on that parameter can be the key to evaluating the integral. We need two results.

Theorem 6.11 (Continuity Lemma)**.** *Let $E\subseteq\mathbb{R}$ be measurable and let $(a,b)\subseteq\mathbb{R}$ be a non-empty open interval and $u:(a,b)\times E\to\mathbb{R}$ be a function satisfying*

1. *$x\mapsto u(t,x)\in L^1(E)$ for every fixed $t\in(a,b)$.*
2. *$t\mapsto u(t,x)$ is continuous for each fixed $x\in E$.*
3. *$|u(t,x)|\leq w(x)$ for all $(t,x)\in(a,b)\times E$ where $w\in L^1(E)$.*

Then the function $v:(a,b)\to\mathbb{R}$ given by

$$t\mapsto v(t)=\int_E u(t,x)dx$$

is continuous.

Proof. First notice that $v(t)$ is well-defined by the first assumption. We are going to show that for any $t \in (a,b)$ and every sequence $\{t_j\}_{j\in\mathbb{N}}$ of points in (a,b) with $\lim_{j\to\infty} t_j = t$ we have

$$\lim_{j\to\infty} v(t_j) = v(t).$$

This is equivalent to continuity of v at the point t.

Because of the second assumption of the Theorem, $u(\cdot, x)$ is continuous and so

$$u_j(x) \equiv u(t_j, x) \to u(t,x) \text{ as } j \to \infty$$

and $|u_j(x)| \le w(x)$ for all $x \in E$. So by the DCT

$$\begin{aligned} \lim_{j\to\infty} v(t_j) &= \lim_{j\to\infty} \int_E u(t_j, x)dx \\ &= \int_E \lim_{j\to\infty} u(t_j,x)dx \\ &= \int u(t,x)dx = v(t) \end{aligned}$$

as required. □

A very similar argument will yield the following.

Theorem 6.12 (Differentiability Lemma)**.** *Let $(a,b) \subseteq \mathbb{R}$ be a non-empty open interval and $E \subseteq \mathbb{R}$ be a measurable set. Let $u : (a,b) \times E \to \mathbb{R}$ be a function satisfying*

1. *$x \mapsto u(t,x)$ is in $L^1(E)$ for every fixed $t \in (a,b)$.*
2. *$t \mapsto u(t,x)$ is differentiable for every fixed $x \in E$.*
3. $$\left|\frac{\partial}{\partial t} u(t,x)\right| \le w(x) \text{ for all } (t,x) \in (a,b) \times E$$

   for some $w \in L^1(E)$.

Then the function $v : (a,b) \to \mathbb{R}$ given by

$$t \mapsto v(t) = \int_E u(t,x)dx$$

is differentiable *and its derivative is*

$$\frac{dv}{dt}(t) = \int_E \frac{\partial u}{\partial t}(t,x)dx.$$

Proof. Let $t \in (a,b)$ and fix some sequence $\{t_j\}_{j\in\mathbb{N}}$ in (a,b) be such that $t_j \ne t$ for all j and $\lim_{j\to\infty} t_j = t$. Set

$$u_j(x) = \frac{u(t_j,x) - u(t,x)}{t_j - t} \to \frac{\partial u}{\partial t}(t,x) \quad \text{as } j \to \infty.$$

(Notice in particular that this shows that $\partial u / \partial t$ is measurable.)

By the Mean Value Theorem of Mods differential calculus and assumption 3 we have that for some $\theta_j(x) \in (a,b)$

$$|u_j(x)| = \left|\frac{\partial u}{\partial t}(t,x)\Big|_{t=\theta_j}\right| \le w(x) \quad \text{for all } j \in \mathbb{N}.$$

Thus $u_j \in L^1(E)$ and the sequence $\{u_j\}_{j\in\mathbb{N}}$ satisfies the conditions of the DCT and so

$$\begin{aligned}
\lim_{j\to\infty} \frac{v(t_j)-v(t)}{t_j-t} &= \lim_{j\to\infty} \int \frac{u(t_j,x)-u(t,x)}{t_j-t} dx \\
&= \lim_{j\to\infty} \int u_j(x) dx \\
&= \int \lim_{j\to\infty} u_j(x) dx \\
&= \int \frac{\partial u}{\partial t}(t,x) dx.
\end{aligned}$$

The limit is independent of the choice of the sequence $\{t_j\}_{j\in\mathbb{N}}$ and so $v(t)$ is differentiable with derivative

$$\int \frac{\partial u}{\partial t}(t,x) dx$$

as required. $\square$

Example 6.13. *Let a be a positive real number and f be defined by*

$$f(x,y) = e^{-xy} \frac{\sin x}{x}, \qquad x > 0, y > a.$$

Show that for each $y > a$ the function $x \mapsto f(x,y)$ is Lebesgue integrable over $(0,\infty)$.

Define the function F on (a,∞) by

$$F(y) = \int_0^\infty e^{-xy} \frac{\sin x}{x} dx.$$

Stating clearly any Theorem that is used, show that F is differentiable on (a,∞) and that

$$F'(y) = -\int_0^\infty e^{-xy} \sin x dx.$$

Hence or otherwise, show that for $y > 0$

$$\int_0^\infty e^{-xy} \frac{\sin x}{x} dx = \frac{\pi}{2} - \tan^{-1} y.$$

Solution. Since $|\sin x| \le |x|$, for $y > a$ we have $|f(x,y)| \le e^{-ax}$ for $x > 0$ and so $f \in L^1((0,\infty))$ by comparison with e^{-ax}.

Now you'd need to state our differentiability lemma.

Now check that the conditions of the lemma are satisfied. We already did the first one. Evidently $f(x,y)$ is differentiable and

$$\frac{\partial f}{\partial y}(x,y) = -e^{-xy} \sin x.$$

Finally, since $|\sin x| \le 1$,

$$\left| \frac{\partial f}{\partial y}(x,y) \right| \le e^{-ax} \qquad \text{for } x > 0, y > a$$

and so setting $w(x) = e^{-ax} \mathbf{1}_{x\ge 0}$ gives the final condition.

Applying the differentiability lemma then gives

$$\begin{aligned}\frac{dF}{dy} &= \int_0^\infty -e^{-xy}\sin x dx\\ &= -\mathrm{Im}\int_0^\infty e^{-xy+ix}dx\\ &= -\mathrm{Im}\frac{1}{y-i} = -\frac{1}{y^2+1}.\end{aligned}$$

Solving the differential equation,

$$F(y) = -\tan^{-1}(y) + C$$

where C is a constant. All that remains is to find C, but notice that (by the DCT) $F(y) \to 0$ as $y \to \infty$ and so $-\pi/2 + C = 0$ which gives

$$F(y) = \frac{\pi}{2} - \tan^{-1} y$$

as required. □

Example 6.14 (The Gamma Function)**.** *The Gamma Function is defined by*

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx.$$

Use the differentiation lemma to find a quick proof that $\Gamma(n+1) = n!$*.*

Solution. Let

$$F(t) = \int_0^\infty e^{-tx}dx = \frac{1}{t} \qquad \text{for } t > \alpha > 0.$$

Differentiating the integrand n times gives $(-1)^n x^n e^{-tx}$. Now $|(-1)^n x^n e^{-tx}| \le x^n e^{-\alpha x}$ which is in $L^1([0,\infty))$ so

$$F^{(n)}(t) = \int_0^\infty (-1)^n x^n e^{-tx}dx.$$

But using that $F(t) = 1/t$, this must also equal $(-1)^n n!/t^{n+1}$. Setting $t = 1$ and equating the two expressions for $F^{(n)}(t)$ we obtain

$$\int_0^\infty x^n e^{-x}dx = n!$$

as required. □

7 Differentiation and integration

As we said right at the beginning of the course, originally integration was introduced as the inverse of differentiation. The goal was to find the functions $F(x)$ which have derivative $f(x)$ for a given function $f(x)$. So with our more general theory of integration, two questions naturally arise.

1. Suppose $f \in L^1([a,b])$ and F is its indefinite integral, that is $F(x) = \int_a^x f(y)dy$. Does this imply that F is differentiable (at least for a.e. x) and that $F' = f$?

2. What conditions on a function F on $[a,b]$ guarantee that $F'(x)$ exists for a.e. x, that F' is integrable and that moreover $F(b) - F(a) = \int_a^b F'(x)dx$?

In Mods you showed that if f is *continuous* on $[a, b]$ then $F' = f$ (Fundamental Theorem of Calculus). If x is a point of continuity of f then the same result will hold true. In fact it is not too hard to show that $F' = f$ for a.e. x.

Theorem 7.1. *Let $f \in L^1([a, b])$ and suppose that*

$$F(x) = F(a) + \int_a^x f(t)dt, \qquad \textit{for } x \in [a, b].$$

Then $F'(x) = f(x)$ for almost all $x \in [a, b]$ (that is everywhere except on a set of measure zero).

So the original problem that motivated Newton and Leibnitz is solved for all Lebesgue integrable functions (at least with our 'a.e.' caveat). But the second problem is more subtle.

First some obvious conditions.

Example 7.2. *Consider the following function on $[0, 1]$.*

$$F(x) = \begin{cases} 0 & x \in [0, 1/2), \\ 1 & x \in [1/2, 1]. \end{cases}$$

Then $F'(x) = 0$ except at $x = 1/2$ where it is not defined and evidently

$$F(x) \neq \int_0^x F'(x)dx \qquad \textit{for } x > 1/2.$$

So points of discontinuity are problematic. Somehow the '$\infty \times 0$' corresponding to integrating F' across $x = 1/2$ cannot be neglected.

But is continuity enough?

Example 7.3 (The Cantor-Lebesgue function or Devil's Staircase)**.** *Recall the construction of the Cantor set from the problem sheet:*

1. *Start with the interval $[0, 1]$ and remove the 'middle third' to obtain*

$$C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1].$$

2. *Remove the middle third of each of the two intervals making up C_1 to give C_2 consisting of four intervals each of length $1/9$ and so on.*

3. *At the nth stage we have a set C_n consisting of 2^n disjoint closed intervals, each of length $1/3^n$.*

$$C = \bigcap_{n=1}^{\infty} C_n$$

is the Cantor Set.

Now we build a continuous function on $[0, 1]$ based on this construction. First define

$$F(x) = \frac{1}{2} \qquad \textit{for } x \in (\frac{1}{3}, \frac{2}{3}).$$

Now set

$$\textit{For } x \in (\frac{1}{9}, \frac{2}{9}), \quad F(x) = \frac{1}{4},$$

$$\textit{For } x \in (\frac{7}{9}, \frac{8}{9}), \quad F(x) = \frac{3}{4},$$

and so on.

If $x \in [0,1]$ has ternary expansion $x = 0.a_1a_2\ldots$ with $a_n = 0, 1, 2$ define $N = \min\{n : a_n = 1\}$ and $N = \infty$ if none of the a_n are one (i.e. when $x \in C$). Now set $b_n = a_n/2$ for $n < N$ and $b_N = 1$ and set

$$F(x) = \sum_{n=1}^{N} \frac{b_n}{2^n}, \qquad \text{for } x \in [0,1].$$

Clearly $F(x)$ is a monotone non-decreasing function with $F(0) = 0$, $F(1) = 1$ and $F'(x) = 0$ for a.e. x.

Lemma 7.4. *The Cantor-Lebesgue function is* continuous *on $[0,1]$.*

Proof.

Fix $x \in (0,1)$ (and then $F(x) \in (0,1)$) and $\epsilon > 0$. We want to find $\delta > 0$ so that $x - \delta < y < x + \delta$ implies $F(x) - \epsilon < F(y) < F(x) + \epsilon$.

First choose $a = k/2^n$, $b = m/2^n$ so that

$$F(x) - \epsilon < a < F(x) < b < F(x) + \epsilon$$

(we can always do this for sufficiently large n). Now by construction a is the value taken by F on an interval with ternary endpoints which we denote by (a_1, a_2) say. Similarly, b is the value taken by F on an interval (b_1, b_2), again with ternary endpoints. Finally choose δ so that

$$a_1 < x - \delta < x + \delta < b_2$$

and by monotonicity of F we have $F(y) \in (F(x) - \epsilon, F(x) + \epsilon)$ whenever $y \in (x - \delta, x + \delta)$ as required. Similar arguments give continuity at 0 and at 1. □

Now for the point of Example 7.3. $F'(x) = 0$ except for $x \in C$, the Cantor set, which is null. So $F'(x)$ exists a.e. but

$$0 = \int_0^1 F'(x)dx \neq F(1) - F(0) = 1.$$

The Fundamental Theorem of Calculus fails!

In fact there is a more general result which tells us that if F is a monotone non-decreasing continuous function then F' exists a.e. and

$$\int_a^b F'(x)dx \leq F(b) - F(a).$$

But when do we get equality?

First let's think about why we *don't* get equality for the Cantor-Lebesgue function. The function increases by one as we pass through C, a null set. From the point of view of the Fundamental Theorem of Calculus, this is as bad as a jump.

So when do we have the Fundamental Theorem of Calculus?

Definition 7.5 (Absolutely continuous function)**.** *A function F defined on $[a,b]$ is* absolutely continuous *if for any $\epsilon > 0$ there exists $\delta > 0$ so that*

$$\sum_{k=1}^{N} |F(b_k) - F(a_k)| < \epsilon$$

whenever

$$\sum_{k=1}^{N} (b_k - a_k) < \delta$$

and the intervals (a_k, b_k) are disjoint.

(In this definition we are implicitly assuming that $a_k < b_k$.) Note that absolutely continuous functions are *uniformly* continuous. If

$$F(x) = \int_a^x f(y)dy$$

where f is integrable, then F is absolutely continuous (see Problem Sheet 4) so this looks like a good condition to impose on a function to ensure that it can be expressed as an integral and indeed we have the following result.

Theorem 7.6. *Suppose that F is absolutely continuous on $[a,b]$. Then F' exists a.e. and is integrable. Moreover,*

$$F(x) - F(a) = \int_a^x F'(y)dy, \qquad \text{for all } a \leq x \leq b.$$

Conversely, if f is integrable, then there exists an absolutely continuous function F such that $F' = f$ a.e. and in fact we may take

$$F(x) = \int_a^x f(y)dy.$$

But these results are solely for integrals against Lebesgue measure. They can be easily modified to cover integrals against measures that can be written as

$$\mu(A) = \int_A f(y)dy$$

for some (necessarily non-negative) integrable function f. (Such measures are called absolutely continuous with respect to Lebesgue measure and f is called the density of the measure.)

Suppose that instead I take a *discrete* measure. For example if I set

$$\mu(A) = \begin{cases} 1 & \text{if } 1/2 \in A, \\ 0 & \text{if } 1/2 \notin A \end{cases}$$

then the function F of Example 7.2 can be written as

$$F(x) = \int_0^x 1\mu(dy), \qquad x \in [0,1].$$

From a probability mass function such as we meet in discrete probability we can obtain an F with a countable number of jumps. Or we can take a mixture. By considering a broader class of measures, the class of functions obtained by integration becomes very rich indeed.

***Lebesgue-Stieltjes measures.**

Suppose that F is a non-decreasing function on $[a,b]$. Then we can define a measure by

$$m_F([c,d]) = F(d+) - F(c-) \equiv \lim_{x \downarrow d} F(x) - \lim_{x \uparrow c} F(x)$$

and follow our procedure for the Lebesgue integral to define the Lebesgue-Stieltjes integral

$$\int f(x) m_F(dx) = \int f(x) dF(x).$$

If F is absolutely continuous then we can write this as

$$\int f(x) F'(x) dx.$$

For a general non-decreasing function F we can decompose F as $F = G + H$ where G is non-decreasing and absolutely continuous and H is non-decreasing with $H' = 0$ a.e.. Then $F' = G'$ a.e. and so

$$G(x) = G(a) + \int_a^x F'(y)dy.$$

This gives a decomposition of the measure, $m_F = m_G + m_H$. The part m_G is said to be *absolutely continuous* with respect to Lebesgue measure and m_H is said to be *singular* with respect to Lebesgue measure. A measure m_G is absolutely continuous with respect to Lebesgue measure if and only if $m_G(E) = 0$ whenever the Lebesgue measure $m(E) = 0$.

8 The L^p spaces

Addition of two measurable functions and multiplication of a measurable function by a scalar both result in a measurable function. So the space of measurable functions is a *vector space*. But armed with our integral we can impose more structure because we also have a notion of 'size'.

Definition 8.1 (Normed space)**.** *Let V be a vector space over $\mathbb{R}$. A function $\|\cdot\| : V \to (0, \infty)$ is a* norm *if*

1. $\|v\| = 0$ *if and only if* $v = 0$.
2. $\|\alpha v\| = |\alpha|\|v\|$ *whenever* $\alpha \in \mathbb{R}$.
3. $\|u + v\| \leq \|u\| + \|v\|$ *whenever* $u, v \in V$.

We then call $(V, \|\cdot\|)$ *a* normed vector space*.*

In Euclidean space, our usual notion of size defines a norm.

Once you have a norm, you automatically have a notion of distance,

$$d(u, v) = \|u - v\|.$$

There are *lots* of ways to define norms on the spaces of measurable functions (with respect to any given measure). For concreteness we'll stick with Lebesgue measure. Then the most important norms are the so-called L^p norms.

Definition 8.2. *For $1 \leq p < \infty$ and E a measurable subset of $\mathbb{R}^n$, $L^p(E)$ denotes the set of* equivalence classes *of measurable functions f on E such that*

$$\int_E |f(x)|^p dx < \infty$$

where $f \sim g$ if and only if $f = g$ a.e.. We write

$$\|f\|_p = \left(\int_E |f(x)|^p dx\right)^{1/p}$$

for the L^p-norm*.*

The reason that we take equivalence classes here is that otherwise $\|\cdot\|_p$ would not define a *norm*- Condition 1 would not be satisfied since $\|f\|_p = 0$ only guarantees that $f = 0$ a.e. on E. That Condition 2 is satisfied is obvious. Condition 3 is known as *Minkowski's inequality* which we prove in Lemma 8.4.

When $p = \infty$ we take

$$\|f\|_\infty = \inf\{z : |f(x)| \le z \text{ a.e.}\} = \inf_{F:m(F)=0} \sup_{x\in E\setminus F} |f(x)|.$$

(Sometimes this is called the 'essential supremum' of $|f|$.)

But now we have a normed vector space we can use tools developed in that setting to say things about integrals. For example, recall the Cauchy-Schwarz inequality from Euclidean space (or more generally from an inner product space),

$$\langle u, v\rangle \le \|u\|\|v\|.$$

So for vectors $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ this says

$$\sum_{i=1}^n |u_i v_i| \le \left(\sum_{i=1}^n u_i^2\right)^{1/2} \left(\sum_{i=1}^n v_i^2\right)^{1/2}.$$

In $L^2(E)$ this becomes, for $f, g \in L^2(E)$,

$$\int_E |f(x)g(x)|dx \le \left(\int_E |f(x)|^2 dx\right)^{1/2} \left(\int_E |g(x)|^2 dx\right)^{1/2}.$$

Of course we didn't need to take Lebesgue measure and taking a suitable discrete measure we'd recover the result in Euclidean space.

In fact Cauchy-Schwartz is a special case of a more general result:

Lemma 8.3 (Hölder's inequality). *Suppose that $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(E)$, $g \in L^q(E)$ then*

$$\int_E |f(x)g(x)|dx \le \left(\int_E |f(x)|^p dx\right)^{1/p} \left(\int_E |g(x)|^q dx\right)^{1/q} = \|f\|_p \|g\|_q.$$

Proof.

The key to the proof is a simple result about real numbers. Suppose that $a, b > 0$. First note that either $b \le a^{p-1}$ or $b \ge a^{p-1}$ and in the second case $b^{q-1} \ge a^{(p-1)(q-1)} = a$ (since $(p-1)(q-1) = 1$). Suppose that $b \le a^{p-1}$. The curve in the picture denotes $y = x^{p-1}$ or, equivalently, $x = y^{q-1}$. Then with A and B denoting the areas of the regions marked in the picture we see that

$$A + B \ge ab.$$

Now (by integration) the area of A is a^p/p and that of B is b^q/q and so the inequality reads

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}. \qquad (12)$$

Similarly, if $b \ge a^{p-1}$ then $a \le b^{q-1}$ and so interchanging the rôles of (a, p) and (b, q) we once again obtain equation (12).

Now for $p > 1$ take

$$a = \frac{|f(x)|}{\|f\|_p}, \qquad b = \frac{|g(x)|}{\|g\|_q}.$$

Then using equation (12) and integrating over E we find

$$\int_E \frac{|f(x)g(x)|}{\|f\|_p\|g\|_q}dx \leq \frac{1}{p\|f\|_p^p}\int_E |f(x)|^p dx + \frac{1}{q\|g\|_q^q}\int_E |g(x)|^q dx = \frac{1}{p} + \frac{1}{q} = 1.$$

Rearranging gives the result. □

Now let's check the triangle inequality for L^p spaces.

Lemma 8.4 (Minkowski's inequality)**.** *For $p \geq 1$,*

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

Proof. The result is clear for $p = 1$, so suppose that $p > 1$. Then

$$\begin{aligned} \|f+g\|_p^p &= \int_E |f(x)+g(x)|^p dx \\ &\leq \int |f(x)| \cdot |f(x)+g(x)|^{p-1}dx + \int |g(x)| \cdot |f(x)+g(x)|^{p-1}dx \\ &\leq \left(\int |f(x)|^p dx\right)^{1/p} \cdot \left(\int |f(x)+g(x)|^{(p-1)q}dx\right)^{1/q} \\ &\qquad + \left(\int |g(x)|^p dx\right)^{1/p} \cdot \left(\int |f(x)+g(x)|^{(p-1)q}dx\right)^{1/q} \end{aligned}$$

where $\frac{1}{p} + \frac{1}{q} = 1$ and we have used Hölder's inequality to obtain the last line. Now $q(p-1) = p$ and so this becomes

$$\|f+g\|_p^p \leq \|f\|_p \cdot \|f+g\|_p^{p/q} + \|g\|_p \cdot \|f+g\|_p^{p/q}.$$

Finally, since $p(1-1/q) = 1$, dividing both sides by $\|f+g\|_p^{p/q}$ we obtain

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

as required. □

Much of the motivation for introducing the Lebesgue integral was its stability under limits and that is inherited by the L^p spaces. From the point of view of an analyst a key property of L^p spaces is that they are *complete*. What this means is that every Cauchy sequence of functions in L^p converges to a limit in L^p - just as every Cauchy sequence in $\mathbb{R}$ converges to a limit in $\mathbb{R}$. Often in analysis one defines a function through a limiting procedure and it cannot be written down explicitly. This result gives us the power to control the behaviour of the limit.

Proposition 8.5. *For $1 < p < \infty$, L^p is* complete*.*

The result is still true for $p = \infty$ but we omit the proof here.

Proof of Proposition 8.5.

Let $\{f_n\}_{n\geq 1}$ be a Cauchy sequence in L^p. That is, given $\epsilon > 0$ there exists N so that $n, m > N$ implies $\|f_n - f_m\|_p < \epsilon$. By passing to a subsequence if necessary we can assume that

$$\|f_n - f_{n+1}\| < \frac{1}{2^n}.$$

Let

$$g_k = \sum_{j=1}^{k} |f_{j+1} - f_j|, \quad g = \sum_{j=1}^{\infty} |f_{j+1} - f_j|.$$

By the iterated Minkowski inequality $\|g_k\|_p < 1$. By Fatou's Lemma,

$$\|g\|_p^p = \int \lim_{k\to\infty} g_k^p(x)dx \le \liminf_{k\to\infty} \int g_k^p(x)dx \le 1.$$

So g is finite a.e. and

$$f = f_1 + \sum_{n=1}^{\infty}(f_{n+1} - f_n) = \lim_{k\to\infty} f_k$$

is absolutely convergent a.e. (c.f Corollary 5.30). Define f arbitrarily where the series does not converge (recall that a function in L^p is only defined up to a null set).

If $\epsilon > 0$ there exists N so that $\|f_n - f_m\|_p < \epsilon$ for $n, m \ge N$ so for each $m > N$

$$\begin{aligned} \int |f(x) - f_m(x)|^p dx &= \int \lim_{k\to\infty} |f_k(x) - f_m(x)|^p dx \\ \text{(Fatou)} \qquad &\le \liminf_{k\to\infty} \int |f_k(x) - f_m(x)|^p dx \le \epsilon^p. \end{aligned}$$

So $f - f_m \in L^p$ with $\|f - f_m\|_p \le \epsilon$ and hence $f = (f - f_m) + f_m \in L^p$ as required. □

Example 8.6. *Show that*

$$(m+n)! \le \sqrt{(2m)!(2n)!}.$$

Solution.

$$\begin{aligned} \int_0^\infty x^{m+n}e^{-x}dx &= \int_0^\infty x^m e^{-x/2} x^n e^{-x/2} dx \\ &\le \left(\int_0^\infty x^{2m}e^{-x}dx\right)^{1/2} \left(\int_0^\infty x^{2n}e^{-x}dx\right)^{1/2} \\ &= \sqrt{(2m)!}\sqrt{(2n)!}. \end{aligned}$$

□

Useful Reading

[1] H. Lebesgue *Leçons sur l'intégration et la recherche des fonctions primitives.* Gauthier-Villars, Paris 1904.

[2] M. Capinski & P.E. Kopp *Measure, Integral and Probability.* Springer 2004.

[3] E.M. Stein & R. Shakarchi *Real Analysis: Measure Theory, Integration and Hilbert Spaces. Princeton Lectures in Analysis III.* Princeton University Press 2005.

[4] H.L. Royden *Real Analysis. 3rd Edition.* Macmillan 1988.

[5] R.G. Bartle *The Elements of Integration.* Wiley 1966.

[6] D.S. Kurtz & C.W.Swartz *Theories of Integration. Series in Real Analysis Vol.9.* World Scientific 2004.

[7] G. Temple *The structure of Lebesgue integration theory.* Oxford Clarendon Press 1971.