CHAPTER 2

The Lebesgue integral

In this second part of the course the basic theory of the Lebesgue integral is presented. Here I follow an idea of Jan Mikusiński, of completing the space of step functions on the line under the L^1 norm but in such a way that the limiting objects are seen directly as functions (defined almost everywhere). There are other places you can find this, for instance the book of Debnaith and Mikusiński [**1**]. Here I start from the Riemann integral, since this is a prerequisite of the course; this streamlines things a little. The objective is to arrive at a working knowledge of Lebesgue integration as quickly as seems acceptable, to pass on to the discussion of Hilbert space and then to more analytic questions.

So, the treatment of the Lebesgue integral here is intentionally compressed, while emphasizing the completeness of the spaces L^1 and L^2. In lectures everything is done for the real line but in such a way that the extension to higher dimensions – carried out partly in the text but mostly in the problems – is not much harder.

1. Integrable functions

Recall that the Riemann integral is defined for a certain class of bounded functions $u : [a, b] \longrightarrow \mathbb{C}$ (namely the Riemann integrable functions) which includes all continuous functions. It depends on the compactness of the interval and the boundedness of the function, but can be extended to an 'improper integral' on the whole real line for which however some of the good properties fail. This is NOT what we will do. Rather we consider the space of continuous functions 'with compact support':

(2.1)
$$\mathcal{C}_{\mathrm{c}}(\mathbb{R}) = \{u : \mathbb{R} \longrightarrow \mathbb{C}; u \text{ is continuous and } \exists\ R \text{ such that } u(x) = 0 \text{ if } |x| > R\}.$$

Thus each element $u \in \mathcal{C}_{\mathrm{c}}(\mathbb{R})$ vanishes outside an interval $[-R, R]$ where the R depends on the u. Note that the *support* of a continuous function is defined to be the complement of the largest open set on which it vanishes (or as the closure of the set of points at which it is non-zero – make sure you see why these are the same). Thus (2.1) says that the support, which is necessarily closed, is contained in some interval $[-R, R]$, which is equivalent to saying it is compact.

LEMMA 2.1. *The Riemann integral defines a continuous linear functional on $\mathcal{C}_c(\mathbb{R})$ equipped with the L^1 norm*

$$\begin{aligned}\int_{\mathbb{R}} u &= \lim_{R\to\infty}\int_{[-R,R]} u(x)dx,\\ \|u\|_{L^1} &= \lim_{R\to\infty}\int_{[-R,R]} |u(x)|dx, \qquad (2.2)\\ |\int_{\mathbb{R}} u| &\le \|u\|_{L^1}.\end{aligned}$$

The limits here are trivial in the sense that the functions involved are constant for large R.

PROOF. These are basic properties of the Riemann integral see Rudin [**4**]. □

Note that $\mathcal{C}_c(\mathbb{R})$ is a normed space with respect to $\|u\|_{L^1}$ as defined above; that it is not complete is one of the main reasons for passing to the Lebesgue integral. With this small preamble we can directly define the 'space' of Lebesgue integrable functions on $\mathbb{R}$.

DEFINITION 2.1. A function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is *Lebesgue integrable*, written $f \in \mathcal{L}^1(\mathbb{R})$, if there exists a series with partial sums $f_n = \sum\limits_{j=1}^{n} w_j$, $w_j \in \mathcal{C}_c(\mathbb{R})$ which is absolutely summable,

$$\sum_j \int |w_j| < \infty \tag{2.3}$$

and such that

$$\sum_j |w_j(x)| < \infty \Longrightarrow \lim_{n\to\infty} f_n(x) = \sum_j w_j(x) = f(x). \tag{2.4}$$

This is a somewhat convoluted definition which you should think about a bit. Its virtue is that it is all there. The problem is that it takes a bit of unravelling. Before we go any further note that the sequence w_j obviously determines the sequence of partial sums f_n, both in $\mathcal{C}_c(\mathbb{R})$ but the converse is also true since

$$\begin{gathered} w_1 = f_1,\ w_k = f_k - f_{k-1},\ k > 1,\\ \sum_j \int |w_j| < \infty \Longleftrightarrow \sum_{k>1} \int |f_k - f_{k-1}| < \infty. \end{gathered} \tag{2.5}$$

You might also notice that can we do some finite manipulation, for instance replace the sequence w_j by

$$W_1 = \sum_{j\le N} w_j,\ W_k = w_{N+k-1},\ k > 1 \tag{2.6}$$

and nothing much changes, since the convergence conditions in (2.3) and (2.4) are properties only of the *tail* of the sequences and the sum in (2.4) for $w_j(x)$ converges if and only if the corresponding sum for $W_k(x)$ converges and then converges to the same limit.

Before massaging the definition a little, let me give a simple example and check that this definition does include continuous functions defined on an interval and extended to be zero outside – the theory we develop will include the usual Riemann integral although I will not quite prove this in full, but only because it is not particularly interesting.

LEMMA 2.2. *If $f \in \mathcal{C}([a,b])$ then*

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in [a,b] \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

is an integrable function.

PROOF. Just 'add legs' to $\tilde{f}$ by considering the sequence

$$f_n(x) = \begin{cases} 0 & \text{if } x < a - 1/n \text{ or } x > b + 1/n, \\ (1 + n(x-a))f(a) & \text{if } a - 1/n \le x < a, \\ (1 - n(x-b))f(b) & \text{if } b < x \le b + 1/n, \\ f(x) & \text{if } x \in [a,b]. \end{cases} \tag{2.8}$$

This is a continuous function on each of the open subintervals in the description with common limits at the endpoints, so $f_n \in \mathcal{C}_c(\mathbb{R})$. By construction, $f_n(x) \to \tilde{f}(x)$ for each $x \in \mathbb{R}$. Define the sequence w_j which has partial sums the f_n, as in (2.5) above. Then $w_j = 0$ in $[a,b]$ for $j > 1$ and it can be written in terms of the 'legs'

$$l_n = \begin{cases} 0 & \text{if } x < a - 1/n,\ x \ge a \\ (1 + n(x-a)) & \text{if } a - 1/n \le x < a, \end{cases}$$

$$r_n = \begin{cases} 0 & \text{if } x \le b,\ x > b + 1/n \\ (1 - n(x-b)) & \text{if } b \le x \le b + 1/n, \end{cases}$$

as

$$|w_n(x)| = (l_n - l_{n-1})|f(a)| + (r_n - r_{n-1})|f(b)|,\ n > 1. \tag{2.9}$$

It follows that

$$\int |w_n(x)| = \frac{(|f(a)| + |f(b)|)}{n(n-1)}$$

so $\{w_n\}$ is an absolutely summable sequence showing that $\tilde{f} \in \mathcal{L}^1(\mathbb{R})$. □

Returning to the definition, notice that we only say 'there exists' an absolutely summable sequence and that it is required to converge to the function *only* at points at which the pointwise sequence is absolutely summable. At other points anything is permitted. So it is not immediately clear that there are any functions *not* satisfying this condition. Indeed if there was a sequence like w_j above with $\sum_j |w_j(x)| = \infty$ always, then (2.4) would represent no restriction at all. So the point of the definition is that absolute summability – a condition on the integrals in (2.3) – does imply something about (absolute) convergence of the pointwise series. Let us reenforce this idea with another definition:-

DEFINITION 2.2. A set $E \subset \mathbb{R}$ is said to be *of measure zero* in the sense of Lebesgue (which is pretty much always the meaning here) if there is a series $g_n = \sum\limits_{j=1}^{n} v_j$, $v_j \in \mathcal{C}_c(\mathbb{R})$ which is absolutely summable, $\sum\limits_j \int |v_j| < \infty$, and such that

$$\sum_j |v_j(x)| = \infty \ \forall\ x \in E. \tag{2.10}$$

Notice that we do not require E to be precisely the set of points at which the series in (2.10) diverges, only that it does so at all points of E, so E is just a subset of the set on which some absolutely summable series of functions in $\mathcal{C}_c(\mathbb{R})$ does not converge absolutely. So any subset of a set of measure zero is automatically of measure zero. To introduce the little trickery we use to unwind the definition above, consider first the following (important) result.

LEMMA 2.3. *Any finite union of sets of measure zero is a set of measure zero.*

PROOF. Since we can proceed in steps, it suffices to show that the union of two sets of measure zero has measure zero. So, let the two sets be E and F and two corresponding absolutely summable sequences, as in Definition 2.2, be v_j and w_j. Consider the alternating sequence

$$u_k = \begin{cases} v_j & \text{if } k = 2j-1 \text{ is odd} \\ w_j & \text{if } k = 2j \text{ is even.} \end{cases} \tag{2.11}$$

Thus $\{u_k\}$ simply interlaces the two sequences. It follows that u_k is absolutely summable, since

$$\sum_k \|u_k\|_{L^1} = \sum_j \|v_j\|_{L^1} + \sum_j \|w_j\|_{L^1}. \tag{2.12}$$

Moreover, the pointwise series $\sum\limits_k |u_k(x)|$ diverges precisely where one or other of the two series $\sum\limits_j |v_j(x)|$ or $\sum\limits_j |w_j(x)|$ diverges. In particular it must diverge on $E \cup F$ which is therefore, from the definition, a set of measure zero. □

The definition of $f \in \mathcal{L}^1(\mathbb{R})$ above certainly requires that the equality on the right in (2.4) should hold outside a set of measure zero, but in fact a specific one, the one on which the series on the left diverges. Using the same idea as in the lemma above we can get rid of this restriction.

PROPOSITION 2.1. *If $f : \mathbb{R} \longrightarrow \mathbb{C}$ and there exists a series $f_n = \sum\limits_{j=1}^{n} w_j$ with $w_j \in \mathcal{C}_c(\mathbb{R})$ which is absolutely summable, so $\sum\limits_j \|w_j\|_{L^1} < \infty$, and a set $E \subset \mathbb{R}$ of measure zero such that*

$$x \in \mathbb{R} \setminus E \Longrightarrow f(x) = \lim_{n\to\infty} f_n(x) = \sum_{j=1}^{\infty} w_j(x) \tag{2.13}$$

then $f \in \mathcal{L}^1(\mathbb{R})$.

Recall that when one writes down an equality such as on the right in (2.13) one is implicitly saying that $\sum\limits_{j=1}^{\infty} w_j(x)$ converges *and* the equality holds for the limit. We will call a sequence as the w_j above an 'approximating series' for $f \in \mathcal{L}^1(\mathbb{R})$.
This is indeed a refinement of the definition since all $f \in \mathcal{L}^1(\mathbb{R})$ arise this way, taking E to be the set where $\sum\limits_j |w_j(x)| = \infty$ for a series as in the defintion.

PROOF. By definition of a set of measure zero there is some series v_j as in (2.10). Now, consider the series obtained by alternating the terms between w_j, v_j

and $-v_j$. Explicitly, set

$$u_j = \begin{cases} w_k & \text{if } j = 3k-2 \\ v_k & \text{if } j = 3k-1 \\ -v_k(x) & \text{if } j = 3k. \end{cases} \tag{2.14}$$

This defines a series in $\mathcal{C}_c(\mathbb{R})$ which is absolutely summable, with

$$\sum_j \|u_j(x)\|_{L^1} = \sum_k \|w_k\|_{L^1} + 2\sum_k \|v_k\|_{L^1}. \tag{2.15}$$

The same sort of identity is true for the pointwise series which shows that

$$\sum_j |u_j(x)| < \infty \text{ iff } \sum_k |w_k(x)| < \infty \text{ and } \sum_k |v_k(x)| < \infty. \tag{2.16}$$

So if the pointwise series on the left converges absolutely, then $x \notin E$, by definition and hence, using (2.13), we find that

$$\sum_j |u_j(x)| < \infty \Longrightarrow f(x) = \sum_j u_j(x) \tag{2.17}$$

since the sequence of partial sums of the u_j cycles through f_n, $f_n(x) + v_n(x)$, then $f_n(x)$ and then to $f_{n+1}(x)$. Since $\sum_k |v_k(x)| < \infty$ the sequence $|v_n(x)| \to 0$ so (2.17) indeed follows from (2.13). □

This is the trick at the heart of the definition of integrability above. Namely we can manipulate the series involved in this sort of way to prove things about the elements of $\mathcal{L}^1(\mathbb{R})$. One point to note is that if w_j is an absolutely summable series in $\mathcal{C}_c(\mathbb{R})$ then

$$F(x) = \begin{cases} \sum_j |w_j(x)| & \text{when this is finite} \\ 0 & \text{otherwise} \end{cases} \Longrightarrow F \in \mathcal{L}^1(\mathbb{R}). \tag{2.18}$$

The sort of property (2.13), where some condition holds on the complement of a set of measure zero is so commonly encountered in integration theory that we give it a simpler name.

DEFINITION 2.3. A condition that holds on $\mathbb{R} \setminus E$ for some set of measure zero, E, is said to hold *almost everywhere*. In particular we write

$$f = g \text{ a.e. if } f(x) = g(x) \ \forall\ x \in \mathbb{R} \setminus E,\ E \text{ of measure zero.} \tag{2.19}$$

Of course as yet we are living dangerously because we have done nothing to show that sets of measure zero are 'small' let alone 'ignorable' as this definition seems to imply. Beware of the trap of 'proof by declaration'!

Now Proposition 2.1 can be paraphrased as 'A function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is Lebesgue integrable if and only if it is the pointwise sum a.e. of an absolutely summable series in $\mathcal{C}_c(\mathbb{R})$.'

2. Linearity of $\mathcal{L}^1$

The word 'space' is quoted in the definition of $\mathcal{L}^1(\mathbb{R})$ above, because it is not immediately obvious that $\mathcal{L}^1(\mathbb{R})$ is a linear space, even more importantly it is far from obvious that the integral of a function in $\mathcal{L}^1(\mathbb{R})$ is well defined (which is the point of the exercise after all). In fact we wish to define the integral to be

$$\int_{\mathbb{R}} f = \sum_n \int w_n \tag{2.20}$$

where $w_n \in \mathcal{C}_{\mathrm{c}}(\mathbb{R})$ is any 'approximating series' meaning now as the w_j in Proposition 2.1. This is fine in so far as the series on the right (of complex numbers) does converge – since we demanded that $\sum\limits_n \int |w_n| < \infty$ so this series converges absolutely – but not fine in so far as the answer might well depend on *which* series we choose which 'approximates f' in the sense of the definition or Proposition 2.1.

So, the immediate aim is to prove these two things. First we will do a little more than prove the linearity of $\mathcal{L}^1(\mathbb{R})$. Recall that a function is 'positive' if it takes only non-negative values.

Proposition 2.2. *The space $\mathcal{L}^1(\mathbb{R})$ is linear (over $\mathbb{C}$) and if $f \in \mathcal{L}^1(\mathbb{R})$ the real and imaginary parts,* $\operatorname{Re} f$, $\operatorname{Im} f$ *are Lebesgue integrable as are their positive parts and as is also the absolute value, $|f|$. For a real Lebesgue integrable function there is an approximating sequence as in Proposition 2.1 which is real and if $f \geq 0$ the sequence of partial sums can be arranged to be non-negative.*

Proof. We first consider the real part of a function $f \in \mathcal{L}^1(\mathbb{R})$. Suppose $w_n \in \mathcal{C}_{\mathrm{c}}(\mathbb{R})$ is an approximating series as in Proposition 2.1. Then consider $v_n = \operatorname{Re} w_n$. This is absolutely summable, since $\int |v_n| \leq \int |w_n|$ and

$$\sum_n w_n(x) = f(x) \Longrightarrow \sum_n v_n(x) = \operatorname{Re} f(x). \tag{2.21}$$

Since the left identity holds a.e., so does the right and hence $\operatorname{Re} f \in \mathcal{L}^1(\mathbb{R})$ by Proposition 2.1. The same argument with the imaginary parts shows that $\operatorname{Im} f \in \mathcal{L}^1(\mathbb{R})$. This also shows that a real element has a real approximating sequence.

The fact that the sum of two integrable functions is integrable really is a simple consequence of Proposition 2.1 and Lemma 2.3. Indeed, if $f,\ g \in \mathcal{L}^1(\mathbb{R})$ have approximating series w_n and v_n as in Proposition 2.1 then $u_n = w_n + v_n$ is absolutely summable,

$$\sum_n \int |u_n| \leq \sum_n \int |w_n| + \sum_n \int |v_n| \tag{2.22}$$

and

$$\sum_n w_n(x) = f(x),\ \sum_n v_n(x) = g(x) \Longrightarrow \sum_n u_n(x) = f(x) + g(x).$$

The first two conditions hold outside (probably different) sets of measure zero, E and F, so the conclusion holds outside $E \cup F$ which is of measure zero. Thus $f + g \in \mathcal{L}^1(\mathbb{R})$. The case of cf for $c \in \mathbb{C}$ is more obvious.

The proof that $|f| \in \mathcal{L}^1(\mathbb{R})$ if $f \in \mathcal{L}^1(\mathbb{R})$ is similar but perhaps a little trickier. Again, let $\{w_n\}$ be an approximating series as in the definition showing that $f \in$

$\mathcal{L}^1(\mathbb{R})$. To make a series for $|f|$ we can try the 'obvious' thing. Namely we know that

$$\sum_{j=1}^{n} w_j(x) \to f(x) \text{ if } \sum_j |w_j(x)| < \infty \tag{2.23}$$

so certainly it follows that

$$|\sum_{j=1}^{n} w_j(x)| \to |f(x)| \text{ if } \sum_j |w_j(x)| < \infty.$$

So, set

$$v_1(x) = |w_1(x)|,\ v_k(x) = |\sum_{j=1}^{k} w_j(x)| - |\sum_{j=1}^{k-1} w_j(x)|\ \forall\ x \in \mathbb{R}. \tag{2.24}$$

Then, for sure,

$$\sum_{k=1}^{N} v_k(x) = |\sum_{j=1}^{N} w_j(x)| \to |f(x)| \text{ if } \sum_j |w_j(x)| < \infty. \tag{2.25}$$

So equality holds off a set of measure zero and we only need to check that $\{v_j\}$ is an absolutely summable series.

The triangle inequality in the 'reverse' form $||v| - |w|| \le |v - w|$ shows that, for $k > 1$,

$$|v_k(x)| = ||\sum_{j=1}^{k} w_j(x)| - |\sum_{j=1}^{k-1} w_j(x)|| \le |w_k(x)|. \tag{2.26}$$

Thus

$$\sum_k \int |v_k| \le \sum_k \int |w_k| < \infty \tag{2.27}$$

so the v_k's do indeed form an absolutely summable series and (2.25) holds almost everywhere, so $|f| \in \mathcal{L}^1(\mathbb{R})$.

For a positive function this last argument yields a real approximating sequence with positive partial sums. $\square$

By combining these results we can see again that if $f,\ g \in \mathcal{L}^1(\mathbb{R})$ are both real valued then

$$f_+ = \max(f, 0),\ \max(f, g),\ \min(f, g) \in \mathcal{L}^1(\mathbb{R}). \tag{2.28}$$

Indeed, the positive part, $f_+ = \frac{1}{2}(|f| + f)$, $\max(f, g) = g + (f - g)_+$, $\min(f, g) = -\max(-f, -g)$.

3. The integral on $\mathcal{L}^1$

Next we want to show that the integral is well defined via (2.20) for any approximating series. From Propostion 2.2 it is enough to consider only real functions. For this, recall a result concerning a case where uniform convergence of continuous functions follows from pointwise convergence, namely when the convergence is monotone, the limit is continuous, and the space is compact. It works on a general compact metric space but we can concentrate on the case at hand.

LEMMA 2.4. *If $u_n \in \mathcal{C}_c(\mathbb{R})$ is a decreasing sequence of non-negative functions such that $\lim_{n\to\infty} u_n(x) = 0$ for each $x \in \mathbb{R}$ then $u_n \to 0$ uniformly on $\mathbb{R}$ and*

$$\lim_{n\to\infty} \int u_n = 0. \tag{2.29}$$

PROOF. Since all the $u_n(x) \geq 0$ and they are decreasing (which really means not increasing of course) if $u_1(x)$ vanishes at x then all the other $u_n(x)$ vanish there too. Thus there is one $R > 0$ such that $u_n(x) = 0$ if $|x| > R$ for all n, namely one that works for u_1. So we only need consider what happens on $[-R, R]$ which is compact. For any $\epsilon > 0$ look at the sets

$$S_n = \{x \in [-R, R]; u_n(x) \geq \epsilon\}.$$

This can also be written $S_n = u_n^{-1}([\epsilon, \infty)) \cap [-R, R]$ and since u_n is continuous it follows that S_n is closed and hence compact. Moreover the fact that the $u_n(x)$ are decreasing means that $S_{n+1} \subset S_n$ for all n. Finally,

$$\bigcap_n S_n = \emptyset$$

since, by assumption, $u_n(x) \to 0$ for each x. Now the property of compact sets in a metric space that we use is that if such a sequence of decreasing compact sets has empty intersection then the sets themselves are empty from some n onwards. This means that there exists N such that $\sup_x u_n(x) < \epsilon$ for all $n > N$. Since $\epsilon > 0$ was arbitrary, $u_n \to 0$ uniformly.

One of the basic properties of the Riemann integral is that the integral of the limit of a uniformly convergent sequence (even of Riemann integrable functions but here continuous) is the limit of the sequence of integrals, which is (2.29) in this case. □

We can easily extend this in a useful way – the direction of monotonicity is reversed really just to mentally distinquish this from the preceding lemma.

LEMMA 2.5. *If $v_n \in \mathcal{C}_c(\mathbb{R})$ is any increasing sequence such that $\lim_{n\to\infty} v_n(x) \geq 0$ for each $x \in \mathbb{R}$ (where the possibility $v_n(x) \to \infty$ is included) then*

$$\lim_{n\to\infty} \int v_n dx \geq 0 \textit{ including possibly } +\infty. \tag{2.30}$$

PROOF. This is really a corollary of the preceding lemma. Consider the sequence of functions

$$w_n(x) = \begin{cases} 0 & \text{if } v_n(x) \geq 0 \\ -v_n(x) & \text{if } v_n(x) < 0. \end{cases} \tag{2.31}$$

Since this is the maximum of two continuous functions, namely $-v_n$ and 0, it is continuous and it vanishes for large x, so $w_n \in \mathcal{C}_c(\mathbb{R})$. Since $v_n(x)$ is increasing, w_n is decreasing and it follows that $\lim w_n(x) = 0$ for all x – either it gets there for some finite n and then stays 0 or the limit of $v_n(x)$ is zero. Thus Lemma 2.4 applies to w_n so

$$\lim_{n\to\infty} \int_{\mathbb{R}} w_n(x) dx = 0.$$

Now, $v_n(x) \geq -w_n(x)$ for all x, so for each n, $\int v_n \geq -\int w_n$. From properties of the Riemann integral, $v_{n+1} \geq v_n$ implies that $\int v_n dx$ is an increasing sequence and it is bounded below by one that converges to 0, so (2.30) is the only possibility. □

From this result applied carefully we see that the integral behaves sensibly for absolutely summable series.

LEMMA 2.6. *Suppose* $u_n \in \mathcal{C}_c(\mathbb{R})$ *is an absolutely summable series of real-valued functions, so* $\sum\limits_n \int |u_n|dx < \infty$, *and also suppose that*

$$\sum_n u_n(x) = 0 \ a.e. \tag{2.32}$$

then

$$\sum_n \int u_n dx = 0. \tag{2.33}$$

PROOF. As already noted, the series (2.33) does converge, since the inequality $|\int u_n dx| \le \int |u_n| dx$ shows that it is absolutely convergent (hence Cauchy, hence convergent).

If E is a set of measure zero such that (2.32) holds on the complement then we can modify u_n as in (2.14) by adding and subtracting a non-negative absolutely summable sequence v_k which diverges absolutely on E. For the new sequence u_n (2.32) is strengthened to

$$\sum_n |u_n(x)| < \infty \Longrightarrow \sum_n u_n(x) = 0 \tag{2.34}$$

and the conclusion (2.33) holds for the new sequence if and only if it holds for the old one.

Now, we need to get ourselves into a position to apply Lemma 2.5. To do this, just choose some integer N (large but it doesn't matter yet) and consider the sequence of functions – it depends on N but I will suppress this dependence –

$$U_1(x) = \sum_{n=1}^{N+1} u_n(x),\ U_j(x) = |u_{N+j}(x)|,\ j \ge 2. \tag{2.35}$$

This is a sequence in $\mathcal{C}_c(\mathbb{R})$ and it is absolutely summable – the convergence of $\sum\limits_j \int |U_j| dx$ only depends on the 'tail' which is the same as for u_n. For the same reason,

$$\sum_j |U_j(x)| < \infty \Longleftrightarrow \sum_n |u_n(x)| < \infty. \tag{2.36}$$

Now the sequence of partial sums

$$g_p(x) = \sum_{j=1}^{p} U_j(x) = \sum_{n=1}^{N+1} u_n(x) + \sum_{j=2}^{p} |u_{N+j}| \tag{2.37}$$

is increasing with p – since we are adding non-negative functions. If the two equivalent conditions in (2.36) hold then

$$\sum_n u_n(x) = 0 \Longrightarrow \sum_{n=1}^{N+1} u_n(x) + \sum_{j=2}^{\infty} |u_{N+j}(x)| \ge 0 \Longrightarrow \lim_{p\to\infty} g_p(x) \ge 0, \tag{2.38}$$

since we are only increasing each term. On the other hand if these conditions do not hold then the tail, any tail, sums to infinity so

$$\lim_{p\to\infty} g_p(x) = \infty. \tag{2.39}$$

Thus the conditions of Lemma 2.5 hold for g_p and hence

$$\sum_{n=1}^{N+1}\int u_n + \sum_{j\ge N+2}\int |u_j(x)|dx \ge 0. \tag{2.40}$$

Using the same inequality as before this implies that

$$\sum_{n=1}^{\infty}\int u_n \ge -2\sum_{j\ge N+2}\int |u_j(x)|dx. \tag{2.41}$$

This is true for any N and as $N\to\infty$, $\lim_{N\to\infty}\sum\limits_{j\ge N+2}\int |u_j(x)|dx = 0$. So the fixed number on the left in (2.41), which is what we are interested in, must be non-negative.

In fact the signs in the argument can be reversed, considering instead

$$h_1(x) = -\sum_{n=1}^{N+1} u_n(x),\ h_p(x) = |u_{N+p}(x)|,\ p\ge 2 \tag{2.42}$$

and the final conclusion is the opposite inequality in (2.41). That is, we conclude what we wanted to show, that

$$\sum_{n=1}^{\infty}\int u_n = 0. \tag{2.43}$$

□

Finally then we are in a position to show that the integral of an element of $\mathcal{L}^1(\mathbb{R})$ is well-defined.

PROPOSITION 2.3. *If $f\in\mathcal{L}^1(\mathbb{R})$ then*

$$\int f = \lim_{n\to\infty}\sum_n\int u_n \tag{2.44}$$

is independent of the approximating sequence, u_n, used to define it. Moreover,

$$\begin{gathered}\int |f| = \lim_{N\to\infty}\int |\sum_{k=1}^{N} u_k|,\\ |\int f| \le \int |f| \text{ and}\\ \lim_{n\to\infty}\int |f-\sum_{j=1}^{n} u_j| = 0.\end{gathered} \tag{2.45}$$

So in some sense the definition of the Lebesgue integral 'involves no cancellations'. There are various extensions of the integral which do exploit cancellations – I invite you to look into the definition of the Henstock integral (and its relatives).

PROOF. The uniqueness of $\int f$ follows from Lemma 2.6. Namely, if u_n and u_n' are two series approximating f as in Proposition 2.1 then the real and imaginary parts of the difference $u_n'-u_n$ satisfy the hypothesis of Lemma 2.6 so it follows that

$$\sum_n\int u_n = \sum_n\int u_n'.$$

Then the first part of (2.45) follows from this definition of the integral applied to $|f|$ and the approximating series for $|f|$ devised in the proof of Proposition 2.2. The inequality

$$|\sum_n \int u_n| \le \sum_n \int |u_n|, \tag{2.46}$$

which follows from the finite inequalities for the Riemann integrals

$$|\sum_{n\le N} \int u_n| \le \sum_{n\le N} \int |u_n| \le \sum_n \int |u_n|$$

gives the second part.

The final part follows by applying the same arguments to the series $\{u_k\}_{k>n}$, as an absolutely summable series approximating $f - \sum\limits_{j=1}^{n} u_j$ and observing that the integral is bounded by

$$\int |f - \sum_{k=1}^{n} u_k| \le \sum_{k=n+1}^{\infty} \int |u_k| \to 0 \text{ as } n \to \infty. \tag{2.47}$$

□

4. Summable series in $\mathcal{L}^1(\mathbb{R})$

The next thing we want to know is when the 'norm', which is in fact only a seminorm, on $\mathcal{L}^1(\mathbb{R})$, vanishes. That is, when does $\int |f| = 0$? One way is fairly easy. The full result we are after is:-

Proposition 2.4. *For an integrable function $f \in \mathcal{L}^1(\mathbb{R})$, the vanishing of $\int |f|$ implies that f is a null function in the sense that*

$$f(x) = 0 \ \forall \ x \in \mathbb{R} \setminus E \text{ where } E \text{ is of measure zero.} \tag{2.48}$$

Conversely, if (2.48) holds then $f \in \mathcal{L}^1(\mathbb{R})$ and $\int |f| = 0$.

Proof. The main part of this is the first part, that the vanishing of $\int |f|$ implies that f is null. The converse is the easier direction in the sense that we have already done it.

Namely, if f is null in the sense of (2.48) then $|f|$ is the limit a.e. of the absolutely summable series with all terms 0. It follows from the definition of the integral above that $|f| \in \mathcal{L}^1(\mathbb{R})$ and $\int |f| = 0$. □

For the forward argument we will use the following more technical result, which is also closely related to the completeness of $L^1(\mathbb{R})$ (note the small notational difference, L^1 is the Banach space which is the quotient by the null functions, see below).

Proposition 2.5. *If $f_n \in \mathcal{L}^1(\mathbb{R})$ is an absolutely summable series, meaning that $\sum\limits_n \int |f_n| < \infty$, then*

$$E = \{x \in \mathbb{R}; \sum_n |f_n(x)| = \infty\} \text{ has measure zero.} \tag{2.49}$$

If $f : \mathbb{R} \longrightarrow \mathbb{C}$ satisfies

$$f(x) = \sum_n f_n(x) \text{ a.e.} \tag{2.50}$$

then $f \in \mathcal{L}^1(\mathbb{R})$,

$$
\begin{gathered}
\int f = \sum_n \int f_n, \\
|\int f| \le \int |f| = \lim_{n\to\infty} \int |\sum_{j=1}^n f_j| \le \sum_j \int |f_j| \text{ and} \qquad (2.51)\\
\lim_{n\to\infty} \int |f - \sum_{j=1}^n f_j| = 0.
\end{gathered}
$$

This basically says we can replace 'continuous function of compact support' by 'Lebesgue integrable function' in the definition and get the same result. Of course this makes no sense without the original definition, so what we are showing is that iterating it makes no difference – we do not get a bigger space.

Proof. The proof is very like the proof of completeness via absolutely summable series for a normed space outlined in the preceding chapter.

By assumption each $f_n \in \mathcal{L}^1(\mathbb{R})$, so there exists a sequence $u_{n,j} \in \mathcal{C}_c(\mathbb{R})$ with $\sum_j \int |u_{n,j}| < \infty$ and

$$
\sum_j |u_{n,j}(x)| < \infty \Longrightarrow f_n(x) = \sum_j u_{n,j}(x). \qquad (2.52)
$$

We might hope that $f(x)$ is given by the sum of the $u_{n,j}(x)$ over both n and j, but in general, this double series is not absolutely summable. However we can replace it by one that is. For each n choose N_n so that

$$
\sum_{j>N_n} \int |u_{n,j}| < 2^{-n}. \qquad (2.53)
$$

This is possible by the assumed absolute summability – the tail of the series therefore being small. Having done this, we replace the series $u_{n,j}$ by

$$
u'_{n,1} = \sum_{j\le N_n} u_{n,j}(x),\ u'_{n,j}(x) = u_{n,N_n+j-1}(x)\ \forall\ j \ge 2, \qquad (2.54)
$$

summing the first N_n terms. This still sums to f_n on the same set as in (2.52). So in fact we can simply replace $u_{n,j}$ by $u'_{n,j}$ and we have in addition the estimate

$$
\sum_j \int |u'_{n,j}| \le \int |f_n| + 2^{-n+1}\ \forall\ n. \qquad (2.55)
$$

This follows from the triangle inequality since, using (2.53),

$$
\int |u'_{n,1} + \sum_{j=2}^N u'_{n,j}| \ge \int |u'_{n,1}| - \sum_{j\ge 2} \int |u'_{n,j}| \ge \int |u'_{n,1}| - 2^{-n} \qquad (2.56)
$$

and the left side converges to $\int |f_n|$ by (2.45) as $N \to \infty$. Using (2.53) again gives (2.55).

Dropping the primes from the notation and denoting the new series again as $u_{n,j}$ we can let v_k be some enumeration of the $u_{n,j}$ – using the standard diagonalization procedure for instance. This gives a new series of continuous functions of compact support which is absolutely summable since

$$\sum_{k=1}^{N}\int|v_k| \le \sum_{n,j}\int|u_{n,j}| \le \sum_n(\int|f_n| + 2^{-n+1}) < \infty. \tag{2.57}$$

Using the freedom to rearrange absolutely convergent series we see that

$$\sum_{n,j}|u_{n,j}(x)| < \infty \Longrightarrow f(x) = \sum_k v_k(x) = \sum_n\sum_j u_{n,j}(x) = \sum_n f_n(x). \tag{2.58}$$

The set where (2.58) fails is a set of measure zero, by definition. Thus $f \in \mathcal{L}^1(\mathbb{R})$ and (2.49) also follows. To get the final result (2.51), rearrange the double series for the integral (which is also absolutely convergent). □

For the moment we only need the weakest part, (2.49), of this. To paraphrase this, for any absolutely summable series of integrable functions the absolute point-wise series converges off a set of measure zero – it can only diverge on a set of measure zero. It is rather shocking but this allows us to prove the rest of Proposition 2.4! Namely, suppose $f \in \mathcal{L}^1(\mathbb{R})$ and $\int|f| = 0$. Then Proposition 2.5 applies to the series with each term being $|f|$. This is absolutely summable since all the integrals are zero. So it must converge pointwise except on a set of measure zero. Clearly it diverges whenever $f(x) \neq 0$,

$$\int|f| = 0 \Longrightarrow \{x; f(x) \neq 0\} \text{ has measure zero} \tag{2.59}$$

which is what we wanted to show to finally complete the proof of Proposition 2.4.

5. The space $L^1(\mathbb{R})$

At this point we are able to define the standard Lebesgue space

$$L^1(\mathbb{R}) = \mathcal{L}^1(\mathbb{R})/\mathcal{N},\ \mathcal{N} = \{\text{null functions}\} \tag{2.60}$$

and to check that it is a Banach space with the norm (arising from, to be pedantic) $\int|f|$.

Theorem 2.1. *The quotient space $L^1(\mathbb{R})$ defined by* (2.60) *is a Banach space in which the continuous functions of compact support form a dense subspace.*

The elements of $L^1(\mathbb{R})$ are *equivalence classes of functions*

$$[f] = f + \mathcal{N},\ f \in \mathcal{L}^1(\mathbb{R}). \tag{2.61}$$

That is, we 'identify' two elements of $\mathcal{L}^1(\mathbb{R})$ if (and only if) their difference is null, which is to say they are equal off a set of measure zero. Note that the set which is ignored here is not fixed, but can depend on the functions.

Proof. For an element of $L^1(\mathbb{R})$ the integral of the absolute value is well-defined by Propositions 2.2 and 2.4

$$\|[f]\|_{L^1} = \int|f|,\ f \in [f] \tag{2.62}$$

and gives a *semi-norm* on $\mathcal{L}^1(\mathbb{R})$. It follows from Proposition 1.5 that on the quotient, $\|[f]\|$ is indeed a norm.

The completeness of $L^1(\mathbb{R})$ is a direct consequence of Proposition 2.5. Namely, to show a normed space is complete it is enough to check that any absolutely summable series converges. If $[f_j]$ is an absolutely summable series in $L^1(\mathbb{R})$ then f_j is absolutely summable in $\mathcal{L}^1(\mathbb{R})$ and by Proposition 2.5 the sum of the series exists so we can use (2.50) to define f off the set E and take it to be zero on E. Then, $f \in \mathcal{L}^1(\mathbb{R})$ and the last part of (2.51) means precisely that

$$\lim_{n\to\infty} \|[f] - \sum_{j<n}[f_j]\|_{L^1} = \lim_{n\to\infty} \int |f - \sum_{j<n} f_j| = 0 \tag{2.63}$$

showing the desired completeness. □

Note that despite the fact that it is technically incorrect, everyone says '$L^1(\mathbb{R})$ is the space of Lebesgue integrable functions' even though it is really the space of equivalence classes of these functions modulo equality almost everywhere. Not much harm can come from this mild abuse of language.

Another consequence of Proposition 2.5 and the proof above is an extension of Lemma 2.3.

Proposition 2.6. *Any countable union of sets of measure zero is a set of measure zero.*

Proof. If E is a set of measure zero then any function f which is defined on $\mathbb{R}$ and vanishes outside E is a null function – is in $\mathcal{L}^1(\mathbb{R})$ and has $\int |f| = 0$. Conversely if the characteristic function of E, the function equal to 1 on E and zero in $\mathbb{R} \setminus E$ is integrable and has integral zero then E has measure zero. This is the characterization of null functions above. Now, if E_j is a sequence of sets of measure zero and χ_k is the characteristic function of

$$\bigcup_{j\le k} E_j \tag{2.64}$$

then $\int |\chi_k| = 0$ so this is an absolutely summable series with sum, the characteristic function of the union, integrable and of integral zero. □

6. The three integration theorems

Even though we now 'know' which functions are Lebesgue integrable, it is often quite tricky to use the definitions to actually show that a particular function has this property. There are three standard results on convergence of sequences of integrable functions which are powerful enough to cover most situations that arise in practice – a Monotonicity Lemma, Fatou's Lemma and Lebesgue's Dominated Convergence theorem.

Lemma 2.7 (Montonicity). *If $f_j \in \mathcal{L}^1(\mathbb{R})$ is a monotone sequence, either $f_j(x) \ge f_{j+1}(x)$ for all $x \in \mathbb{R}$ and all j or $f_j(x) \le f_{j+1}(x)$ for all $x \in \mathbb{R}$ and all j, and $\int f_j$ is bounded then*

$$\{x \in \mathbb{R}; \lim_{j\to\infty} f_j(x) \text{ is finite}\} = \mathbb{R} \setminus E \tag{2.65}$$

where E has measure zero and

$$\begin{gathered} f = \lim_{j\to\infty} f_j(x) \text{ a.e. is an element of } \mathcal{L}^1(\mathbb{R}) \\ \text{with } \int f = \lim_{j\to\infty} \int f_j \text{ and } \lim_{j\to\infty} \int |f - f_j| = 0. \end{gathered} \tag{2.66}$$

In the usual approach through measure one has the concept of a measureable, non-negative, function for which the integral 'exists but is infinite' – we do not have this (but we could easily do it, or rather you could). Using this one can drop the assumption about the finiteness of the integral but the result is not significantly stronger.

PROOF. Since we can change the sign of the f_i it suffices to assume that the f_i are monotonically increasing. The sequence of integrals is therefore also montonic increasing and, being bounded, converges. Turning the sequence into a series, by setting $g_1 = f_1$ and $g_j = f_j - f_{j-1}$ for $j \ge 2$ the g_j are non-negative for $j \ge 1$ and

$$\sum_{j\ge 2} \int |g_j| = \sum_{j\ge 2} \int g_j = \lim_{n\to\infty} \int f_n - \int f_1 \tag{2.67}$$

converges. So this is indeed an absolutely summable series. We therefore know from Proposition 2.5 that it converges absolutely a.e., that the limit, f, is integrable and that

$$\int f = \sum_j \int g_j = \lim_{n\to\infty} \int f_j. \tag{2.68}$$

The second part, corresponding to convergence for the equivalence classes in $L^1(\mathbb{R})$ follows from the fact established earlier about $|f|$ but here it also follows from the monotonicity since $f(x) \ge f_j(x)$ a.e. so

$$\int |f - f_j| = \int f - \int f_j \to 0 \text{ as } j \to \infty. \tag{2.69}$$

□

Now, to Fatou's Lemma. This really just takes the monotonicity result and applies it to a sequence of integrable functions with bounded integral. You should recall that the max and min of two real-valued integrable functions is integrable and that

$$\int \min(f,g) \le \min(\int f, \int g). \tag{2.70}$$

This follows from the identities

$$2\max(f,g) = |f-g| + f + g,\ 2\min(f,g) = -|f-g| + f + g. \tag{2.71}$$

LEMMA 2.8 (Fatou). *Let $f_j \in \mathcal{L}^1(\mathbb{R})$ be a sequence of real-valued non-negative integrable functions such that $\int f_j$ is bounded then*

$$\begin{gathered} f(x) = \liminf_{n\to\infty} f_n(x) \text{ exists a.e., } f \in \mathcal{L}^1(\mathbb{R}) \text{ and} \\ \int \liminf f_n = \int f \le \liminf \int f_n. \end{gathered} \tag{2.72}$$

PROOF. You should remind yourself of the properties of lim inf as necessary! Fix k and consider

$$F_{k,n} = \min_{k\le p\le k+n} f_p(x) \in \mathcal{L}^1(\mathbb{R}). \tag{2.73}$$

As discussed above this is integrable. Moreover, this is a decreasing sequence, as n increases, because the minimum is over an increasing set of functions. The $F_{k,n}$ are non-negative so Lemma 2.7 applies and shows that

$$g_k(x) = \inf_{p\ge k} f_p(x) \in \mathcal{L}^1(\mathbb{R}),\ \int g_k \le \int f_n \ \forall\ n \ge k. \tag{2.74}$$

Note that for a decreasing sequence of non-negative numbers the limit exists and is indeed the infimum. Thus in fact,

$$\int g_k \le \liminf \int f_n \ \forall\ k. \tag{2.75}$$

Now, let k vary. Then, the infimum in (2.74) is over a set which decreases as k increases. Thus the $g_k(x)$ are increasing. The integrals of this sequence are bounded above in view of (2.75) since we assumed a bound on the $\int f_n$'s. So, we can apply the monotonicity result again to see that

$$\begin{gathered} f(x) = \lim_{k\to\infty} g_k(x) \text{ exists a.e and } f \in \mathcal{L}^1(\mathbb{R}) \text{ has} \\ \int f \le \liminf \int f_n. \end{gathered} \tag{2.76}$$

Since $f(x) = \liminf f_n(x)$, by definition of the latter, we have proved the Lemma. □

Now, we apply Fatou's Lemma to prove what we are really after:-

Theorem 2.2 (Dominated convergence). *Suppose $f_j \in \mathcal{L}^1(\mathbb{R})$ is a sequence of integrable functions such that*

$$\begin{gathered} \exists\ h \in \mathcal{L}^1(\mathbb{R}) \text{ with } |f_j(x)| \le h(x) \text{ a.e. and} \\ f(x) = \lim_{j\to\infty} f_j(x) \text{ exists a.e.} \end{gathered} \tag{2.77}$$

then $f \in \mathcal{L}^1(\mathbb{R})$ and $[f_j] \to [f]$ in $L^1(\mathbb{R})$, so $\int f = \lim_{n\to\infty} \int f_n$ (including the assertion that this limit exists).

Proof. First, we can assume that the f_j are real since the hypotheses hold for the real and imaginary parts of the sequence and together give the desired result. Moreover, we can change all the f_j's to make them zero on the set on which the initial estimate in (2.77) does not hold. Then this bound on the f_j's becomes

$$-h(x) \le f_j(x) \le h(x)\ \forall\ x \in \mathbb{R}. \tag{2.78}$$

In particular this means that $g_j = h - f_j$ is a non-negative sequence of integrable functions and the sequence of integrals is also bounded, since (2.77) also implies that $\int |f_j| \le \int h$, so $\int g_j \le 2 \int h$. Thus Fatou's Lemma applies to the g_j. Since we have *assumed* that the sequence $g_j(x)$ converges a.e. to $h - f$ we know that

$$\begin{gathered} h - f(x) = \liminf g_j(x) \text{ a.e. and} \\ \int h - \int f \le \liminf \int (h - f_j) = \int h - \limsup \int f_j. \end{gathered} \tag{2.79}$$

Notice the change on the right from liminf to limsup because of the sign.

Now we can apply the same argument to $g_j'(x) = h(x) + f_j(x)$ since this is also non-negative and has integrals bounded above. This converges a.e. to $h(x) + f(x)$ so this time we conclude that

$$\int h + \int f \leq \liminf \int (h + f_j) = \int h + \liminf \int f_j. \tag{2.80}$$

In both inequalities (2.79) and (2.80) we can cancel an $\int h$ and combining them we find

$$\limsup \int f_j \leq \int f \leq \liminf \int f_j. \tag{2.81}$$

In particular the limsup on the left is smaller than, or equal to, the liminf on the right, for the same real sequence. This however implies that they are equal and that the sequence $\int f_j$ converges. Thus indeed

$$\int f = \lim_{n\to\infty} \int f_n. \tag{2.82}$$

Convergence of f_n to f in $L^1(\mathbb{R})$ follows by applying the results proved so far to $|f - f_n|$, converging almost everywhere to 0. In this case (2.82) becomes

$$\lim_{n\to\infty} \int |f - f_n| = 0.$$

□

Generally in applications it is Lebesgue's dominated convergence which is used to prove that some function is integrable. Of course, since we deduced it from Fatou's lemma, and the latter from the Monotonicity lemma, you might say that Lebesgue's theorem is the weakest of the three! However, it is very handy and often a combination does the trick. For instance

Lemma 2.9. *A continuous function $u \in \mathcal{C}(\mathbb{R})$ is Lebesgue integrable if and only if the 'improper Riemann integral'*

$$\lim_{R\to\infty} \int_{-R}^{R} |u(x)| dx < \infty. \tag{2.83}$$

Note that the 'improper integral' without the absolute value can converge without u being Lebesgue integrable.

Proof. If (2.83) holds then consider the sequence of functions $v_N = \chi_{[-N,N]}|u|$, which we know to be in $L^1(\mathbb{R})$ by Lemma 2.2. This is monotonic increasing with limit $|u|$, so the Monotonicity Lemma shows that $|u| \in L^1(\mathbb{R})$. Then consider $w_N = \chi_{[-N,N]}u$ which we also know to be in $L^1(\mathbb{R})$. Since it is bounded by $|u|$ and converges pointwise to u, it follows from Dominated Convergence that $u \in L^1(\mathbb{R})$. Conversely, if $u \in L^1(\mathbb{R})$ then $|u| \in L^1(\mathbb{R})$ and $\chi_{[-N,N]}|u| \in L^1(\mathbb{R})$ converges to $|u|$ so by Dominated Convergence (2.83) must hold. □

So (2.83) holds for any $u \in L^1(\mathbb{R})$.

7. Notions of convergence

We have been dealing with two basic notions of convergence, but really there are more. Let us pause to clarify the relationships between these different concepts.

(1) Convergence of a sequence in $L^1(\mathbb{R})$ (or by slight abuse of language in $\mathcal{L}^1(\mathbb{R})$) – f and $f_n \in L^1(\mathbb{R})$ and

$$\|f - f_n\|_{L^1} \to 0 \text{ as } n \to \infty. \tag{2.84}$$

(2) Convergence almost everywhere:- For some sequence of functions f_n and function f,

$$f_n(x) \to f(x) \text{ as } n \to \infty \text{ for } x \in \mathbb{R} \setminus E \tag{2.85}$$

where $E \subset \mathbb{R}$ is of measure zero.

(3) Dominated convergence:- For $f_j \in L^1(\mathbb{R})$ (or representatives in $\mathcal{L}^1(\mathbb{R})$) such that $|f_j| \le F$ (a.e.) for some $F \in L^1(\mathbb{R})$ and (2.85) holds.

(4) What we might call 'absolutely summable convergence'. Thus $f_n \in L^1(\mathbb{R})$ are such that $f_n = \sum\limits_{j=1}^{n} g_j$ where $g_j \in L^1(\mathbb{R})$ and $\sum\limits_j \int |g_j| < \infty$. Then (2.85) holds for some f.

(5) Monotone convergence. For $f_j \in \mathcal{L}^1(\mathbb{R})$, real valued and montonic, we require that $\int f_j$ is bounded and it then follows that $f_j \to f$ almost everywhere, with $f \in \mathcal{L}^1(\mathbb{R})$ and that the convergence is $\mathcal{L}^1$ and also that $\int f = \lim_j \int f_j$.

So, one important point to know is that 1 does not imply 2. Nor conversely does 2 imply 1 even if we assume that all the f_j and f are in $L^1(\mathbb{R})$.

However, montone convergence implies dominated convergence. Namely if f is the limit then $|f_j| \le |f| + 2|f_1|$ and $f_j \to f$ almost everywhere. Also, monotone convergence implies convergence with absolute summability simply by taking the sequence to have first term f_1 and subsequence terms $f_j - f_{j-1}$ (assuming that f_j is monotonically increasing) one gets an absolutely summable series with sequence of finite sums converging to f. Similarly absolutely summable convergence implies dominated convergence for the sequence of partial sums; by montone convergence the series $\sum\limits_n |f_n(x)|$ converges a.e. and in L^1 to some function F which dominates the partial sums which in turn converge pointwise. I suggest that you make a diagram with these implications in it so that you are clear about the relationships between them.

8. The space $L^2(\mathbb{R})$

So far we have discussed the Banach space $L^1(\mathbb{R})$. The real aim is to get a good hold on the (Hilbert) space $L^2(\mathbb{R})$. This can be approached in several ways. We could start off as for $L^1(\mathbb{R})$ and define $L^2(\mathbb{R})$ as the completion of $\mathcal{C}_{\mathrm{c}}(\mathbb{R})$ with respect to the norm

$$\|f\|_{L^2} = \left(\int |f|^2 \right)^{\frac{1}{2}}. \tag{2.86}$$

This would be rather repetitious; instead we adopt an approach based on Dominated Convergence. You might think, by the way, that it is enough just to ask that $|f|^2 \in \mathcal{L}^1(\mathbb{R})$. This does not work, since even if real the sign of f could jump around and make it non-integrable (provided you believe in the axiom of choice).

Nor would this approach work for $L^1(\mathbb{R})$ since $|f| \in L^1(\mathbb{R})$ does not imply that $f \in L^1(\mathbb{R})$.

Definition 2.4. A function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is said to be 'Lebesgue square integrable', written $f \in \mathcal{L}^2(\mathbb{R})$, if there exists a sequence $u_n \in \mathcal{C}_c(\mathbb{R})$ such that

$$u_n(x) \to f(x) \text{ a.e. and } |u_n(x)|^2 \le F(x) \text{ a.e. for some } F \in \mathcal{L}^1(\mathbb{R}). \tag{2.87}$$

Proposition 2.7. *The space $\mathcal{L}^2(\mathbb{R})$ is linear, $f \in \mathcal{L}^2(\mathbb{R})$ implies $|f|^2 \in \mathcal{L}^1(\mathbb{R})$ and* (2.86) *defines a seminorm on $\mathcal{L}^2(\mathbb{R})$ which vanishes precisely on the null functions $\mathcal{N} \subset \mathcal{L}^2(\mathbb{R})$.*

Definition 2.5. We define $L^2(\mathbb{R}) = \mathcal{L}(\mathbb{R})/\mathcal{N}$.

So we know that $L^2(\mathbb{R})$ is a normed space. It is in fact complete and much more!

Proof. First to see the linearity of $\mathcal{L}^2(\mathbb{R})$ note that if $f \in \mathcal{L}^2(\mathbb{R})$ and $c \in \mathbb{C}$ then $cf \in \mathcal{L}^2(\mathbb{R})$ since if u_n is a sequence as in the definition for f then cu_n is such a sequence for cf.

Similarly if f, $g \in \mathcal{L}^2(\mathbb{R})$ with sequences u_n and v_n then $w_n = u_n + v_n$ has the first property – since we know that the union of two sets of measure zero is a set of measure zero and the second follows from the estimate

$$|w_n(x)|^2 = |u_n(x) + v_n(x)|^2 \le 2|u_n(x)|^2 + 2|v_n(x)|^2 \le 2(F + G)(x) \tag{2.88}$$

where $|u_n(x)|^2 \le F(x)$ and $|v_n(x)|^2 \le G(x)$ with F, $G \in \mathcal{L}^1(\mathbb{R})$.

Moreover, if $f \in \mathcal{L}^2(\mathbb{R})$ then the sequence $|u_n(x)|^2$ converges pointwise almost everywhere to $|f(x)|^2$ so by Lebesgue's Dominated Convergence, $|f|^2 \in \mathcal{L}^1(\mathbb{R})$. Thus $\|f\|_{L^2}$ is well-defined. It vanishes if and only if $|f|^2 \in \mathcal{N}$ but this is equivalent to $f \in \mathcal{N}$ – conversely $\mathcal{N} \subset \mathcal{L}^2(\mathbb{R})$ since the zero sequence works in the definition above.

So we only need to check the triangle inquality, absolute homogeneity being clear, to deduce that $L^2 = \mathcal{L}^2/\mathcal{N}$ is at least a normed space. In fact we checked this earlier on $\mathcal{C}_c(\mathbb{R})$ and the general case follows by continuity:-

$$\|u_n + v_n\|_{L^2} \le \|u_n\|_{L^2} + \|v_n\|_{L^2} \ \forall\ n \Longrightarrow \\ \|f + g\|_{L^2} = \lim_{n\to\infty} \|u_n + v_n\|_{L^2} \le \|f\|_{L^2} + \|g\|_{L^2}. \tag{2.89}$$

□

We will get a direct proof of the triangle inequality as soon as we start talking about (pre-Hilbert) spaces.

So it only remains to check the completeness of $L^2(\mathbb{R})$, which is really the whole point of the discussion of Lebesgue integration.

Theorem 2.3. *The space $L^2(\mathbb{R})$ is complete with respect to $\|\cdot\|_{L^2}$ and is a completion of $\mathcal{C}_c(\mathbb{R})$ with respect to this norm.*

Proof. That $\mathcal{C}_c(\mathbb{R}) \subset L^2(\mathbb{R})$ follows directly from the definition and the fact that a continuous null function must vanish. This is a dense subset since, if $f \in \mathcal{L}^2(\mathbb{R})$ a sequence $u_n \in \mathcal{C}_c(\mathbb{R})$ as in Definition 2.4 satisfies

$$|u_n(x) - u_m(x)|^2 \le 4F(x) \ \forall\ n,\ m, \tag{2.90}$$

and converges almost everwhere to $|f(x)-u_m(x)|^2$ as $n\to\infty$. Thus, by Dominated Convergence, $|f(x)-u_m(x)|^2\in\mathcal{L}^1(\mathbb{R})$. As $m\to\infty$, $|f(x)-u_m(x)|^2\to 0$ almost everywhere and $|f(x)-u_m(x)|^2\le 4F(x)$ so again by dominated convergence

$$\|f-u_m\|_{L^2}=\left(\|(|f-u_m|^2)\|_{L^1})\right)^{\frac12}\to 0. \tag{2.91}$$

This shows the density of $\mathcal{C}_{\mathrm{c}}(\mathbb{R})$ in $L^2(\mathbb{R})$, the quotient by the null functions.

To prove completeness, we only need show that any absolutely L^2-summable sequence in $\mathcal{C}_{\mathrm{c}}(\mathbb{R})$ converges in L^2 and the general case follows by density. So, suppose $\phi_n\in\mathcal{C}_{\mathrm{c}}(\mathbb{R})$ is such a sequence:

$$\sum_n\|\phi_n\|_{L^2}<\infty.$$

Consider $F_k(x)=\left(\sum\limits_{n\le k}|\phi_k(x)|\right)^2$. This is an increasing sequence in $\mathcal{C}_{\mathrm{c}}(\mathbb{R})$ and its L^1 norm is bounded:

$$\|F_k\|_{L^1}=\|\sum_{n\le k}|\phi_n|\|^2_{L^2}\le\left(\sum_{n\le k}\|\phi_n\|_{L^2}\right)^2\le C^2<\infty \tag{2.92}$$

using the triangle inequality and absolutely L^2 summability. Thus, by Monotone Convergence, $F_k(x)\to F(x)$ a.e., $F_k\to F\in\mathcal{L}^1(\mathbb{R})$ and $F_k(x)\le F(x)$ a.e., where we define $F(x)$ to be the limit when this exists and zero otherwise.

Thus the sequence of partial sums $u_k(x)=\sum\limits_{n\le k}\phi_n(x)$ converges almost everywhere – since it converges (absoliutely) on the set where F_k is bounded. Let $f(x)$ be the limit. We want to show that $f\in\mathcal{L}^2(\mathbb{R})$ but this follows from the definition since

$$|u_k(x)|^2\le(\sum_{n\le k}|\phi_n(x)|)^2=F_k(x)\le F(x)\text{ a.e.} \tag{2.93}$$

As in (2.91) it follows that

$$\int|u_k(x)-f(x)|^2\to 0. \tag{2.94}$$

As for the case of $L^1(\mathbb{R})$ it now follows that $L^2(\mathbb{R})$ is complete. □

We want to check that $L^2(\mathbb{R})$ is a Hilbert space (which I will define very soon, even though it is in the next Chapter); to do so observe that if $f,\ g\in\mathcal{L}^2(\mathbb{R})$ have approximating sequences u_n, v_n as in Definition 2.4, so $|u_n(x)|^2\le F(x)$ and $|v_n(x)|^2\le G(x)$ with $F,\ G\in\mathcal{L}^1(\mathbb{R})$ then

$$u_n(x)v_n(x)\to f(x)g(x)\text{ a.e. and }|u_n(x)v_n(x)|\le\frac12(F(x)+G(x)) \tag{2.95}$$

shows that $fg\in\mathcal{L}^1(\mathbb{R})$ by Dominated Convergence. This leads to the basic property of the norm on a (pre)-Hilbert space – that it comes from an inner product. In this case

$$\langle f,g\rangle_{L^2}=\int f(x)\overline{g(x)},\ \|f\|_{L^2}=\langle f,f\rangle^{\frac12}. \tag{2.96}$$

At this point I normally move on to the next chapter on Hilbert spaces with $L^2(\mathbb{R})$ as one motivating example.

9. Measurable and non-measurable sets

The σ-algebra of Lebesgue measurable sets on the line is discussed below but we can directly consider the notion of a set of finite Lebesgue measure. Namely such a set $A \subset \mathbb{R}$ is defined by the condition that the chactacteristic function

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \tag{2.97}$$

is Lebesgue integrable, $\chi_A \in \mathcal{L}^1(\mathbb{R})$. The measure of the set (think 'length') is then $\mu(A) = \int_{\mathbb{R}} \chi_A$ the properties of which are discussed below. Certainly if $A \subset [-R, R]$ has finite measure then $\mu(A) \le 2R$ from the properties of the integral. Similalry if $A_i \subset [-R, R]$ are a sequence of sets of finite measure which are disjoint, $A_i \cap A_j = \emptyset$, $i \ne j$, then

$$A = \bigsqcup_i A_i \text{ has finite measure and } \mu(A) = \sum_i \mu(A_i) \tag{2.98}$$

using Monotone Convergence.

Now the question arises, enquiring minds want to know after all:- Are there bounded sets which are not of finite measure? Similarly, are there functions of bounded support which are not integrable? It turns out this question gets us into somewhat deep water, but it is important to understand some of the limitiations that the insistence on precision in Mathematics places on its practitioners!

Let me present a standard construction of a non-(Lebesgue-)measurable subset of $[0, 1]$ and then comment on the issues that it raises. We start with the quotient space and quotient map

$$q : \mathbb{R} \longrightarrow \mathbb{R}/\mathbb{Q},\ q(x) = \{y \in \mathbb{R}; y = x + r,\ r \in \mathbb{Q}\}. \tag{2.99}$$

This partitions $\mathbb{R}$ into disjoint subsets

$$\mathbb{R} = \bigsqcup_{\tau \in \mathbb{R}/\mathbb{Q}} q^{-1}(\tau). \tag{2.100}$$

Two of these sets intersect if and only if they have elements differing by a rational, and then they are the same.

Now, each of these sets $q^{-1}(\tau)$ intersects $[0, 1]$. This follows from the density of the rationals in the reals, since if $x \in q^{-1}(\tau)$ there exists $r \in \mathbb{Q}$ such that $|x - r| < \frac{1}{2}$ and then $x' = x + (-r + \frac{1}{2}) \in q^{-1}(\tau) \cap [0, 1]$. So we can 'localize' (2.100) to

$$[0, 1] = \bigsqcup_{\tau \in \mathbb{R}/\mathbb{Q}} L(\tau),\ L(\tau) = q^{-1}(\tau) \cap [0, 1] \tag{2.101}$$

where all the sets $L(\tau)$ are non-empty.

DEFINITION 2.6. A *Vitali set*, $V \subset [0, 1]$, is a set which contains precisely one element from each of the $L(\tau)$.

Take such a set V and consider the translates of it by rationals in $[-1, 1]$,

$$V_r = \{y \in [-1, 2]; y = x + r,\ x \in V\},\ r \in \mathbb{Q},\ |r| \le 1. \tag{2.102}$$

For different r these are disjoint – since by construction no two distinct elements of V differ by a rational. The union of these sets however satisfies

$$[0, 1] \subset \bigsqcup_{r \in \mathbb{Q}, |r| \le 1} V_r \subset [-1, 2]. \tag{2.103}$$

Now, we can simply order the sets V_r into a sequence A_i by ordering the rationals in $[-1,1]$.

Suppose V is of finite Lebesgue measure. Then we know that all the V_r are of finite measure and $\mu(V_r) = \mu(V) = \mu(A_i)$ for all i, from the properties of the Lebesgue integral. This means that (2.98) applies, so we have the inequalities

$$\mu([0,1]) = 1 \le \sum_{i=1}^{\infty} \mu(V) \le 3 = \mu([-1,2]). \tag{2.104}$$

Clearly we have a problem! The only way the right-hand inequality can hold is if $\mu(V) = 0$, but then the left-hand inequality fails.

Our conclusion then is that V cannot be Lebesgue measurable! Or is it? Since we are careful people we trace back through the discussion and see (it took people a long, long, time to recognize this) more precisely:-

Proposition 2.8. *If a Vitali set, $V \subset [0,1]$ exists, containing precisely one element of each of the sets $L(\tau)$, then it is bounded and not of finite Lebesgue measure; its characteristic function is a non-negative function of bounded support which is not Lebesgue integrable.*

Okay, so what is the 'issue' here. It is that the existence of such a Vitali set requires the *Axiom of Choice.* There are lots of sets $L(\tau)$ so from the standard (Zermelo-Fraenkel) axions of set theory it does not follow that you can 'choose an element from each' to form a new set. That is a (slightly informal) version of the additional axiom. Now, it has been shown (namely by Gödel and Cohen) that the Axiom of Choice is independent of the Zermelo-Fraenkel Axioms. This does not mean consistency, it means conditional consistency. The Zermelo-Fraenkel axioms together with the Axiom of Choice are inconsistent if and only if the Zermelo-Fraenkel axioms on their own are inconsistent.

Conclusion: As a working Mathematician you are free to choose to believe in the Axiom of Choice or not. It will make your life easier if you do, but it is up to you. Note that if you do not admit the Axiom of Choice, it does not mean that all bounded real sets are measurable, in the sense that you can prove it. Rather it means that it is consistent to believe this (as shown by Solovay).

See also the discussion of the Hahn-Banach Theorem in Section 1.12.

10. Measurable functions

From our original definition of $\mathcal{L}^1(\mathbb{R})$ we know that $\mathcal{C}_c(\mathbb{R})$ is dense in $L^1(\mathbb{R})$. We also know that elements of $\mathcal{C}_c(\mathbb{R})$ can be approximated uniformly, and hence in $L^1(\mathbb{R})$ by step functions – finite linear combinations of the characteristic functions of intervals. It is usual in measure theory to consider a somewhat larger class of functions which contains the step functions:

Definition 2.7. A *simple* function on $\mathbb{R}$ is a finite linear combination (generally with complex coefficients) of characteristic functions of subsets of finite measure:

$$f = \sum_{j=1}^{N} c_j \chi(B_j),\ \chi(B_j) \in \mathcal{L}^1(\mathbb{R}),\ c_j \in \mathbb{C}. \tag{2.105}$$

The real and imaginary parts of a simple function are simple and the positive and negative parts of a real simple function are simple. Since step functions are simple, we know that simple functions are dense in $\mathcal{L}^1(\mathbb{R})$ and that if $0 \le F \in \mathcal{L}^1(\mathbb{R})$ then there exists a sequence of simple functions (take them to be a summable sequence of step functions) $f_n \ge 0$ such that $f_n \to F$ almost everywhere and $f_n \le G$ for some other $G \in \mathcal{L}^1(\mathbb{R})$.

We elevate a special case of the second notion of convergence above to a definition.

Definition 2.8. A function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is *(Lebesgue) measurable* if it is the pointwise limit almost everywhere of a sequence of simple functions.

Lemma 2.10. *A function is Lebesgue measurable if and only if it is the pointwise limit, almost everywhere, of a sequence of continuous functions of compact support.*

Proof. Continuous functions of compact support are the uniform limits of step functions, so this condition certainly implies measurability in the sense of Definition 2.8. Conversely, suppose a function f is the limit almost everywhere of a squence u_n of simple functions. Each of these functions is integrable, so we can find $\phi_n \in \mathcal{C}_c(\mathbb{R})$ such that $\|u_n - \phi_n\|_{L^1} < 2^{-n}$. Then the telescoped sequence $v_1 = u_1 - \phi_1$, $v_k = (u_k - \phi_k) - (u_{k-1} - \phi_{k-1})$, $k > 1$, is absolutely summable so $u_n - \phi_n \to 0$ almost everywhere, and hence $\phi_n \to f$ off a set of measure zero. □

So replacing 'simple functions' by continuous functions in Definition 2.8 makes no difference – and the same for approximation by elements of $\mathcal{L}^1(\mathbb{R})$.

The measurable functions form a linear space since if f and g are measurable and f_n, g_n are sequences of simple functions as required by the definition then $c_1 f_n(x) + c_2 f_2(x) \to c_1 f(x) + c_2 g(x)$ on the intersection of the sets where $f_n(x) \to f(x)$ and $g_n(x) \to g(x)$ which is the complement of a set of measure zero.

Now, from the discussion above, we know that each element of $\mathcal{L}^1(\mathbb{R})$ is measurable. Conversely:

Lemma 2.11. *A function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is an element of $\mathcal{L}^1(\mathbb{R})$ if and only if it is measurable and there exists $F \in \mathcal{L}^1(\mathbb{R})$ such that $|f| \le F$ almost everywhere.*

Proof. If f is measurable there exists a sequence of simple functions f_n such that $f_n \to f$ almost everywhere. The real part, $\operatorname{Re} f$, is also measurable as the limit almost everywhere of $\operatorname{Re} f_n$ and from the hypothesis $|\operatorname{Re} f| \le F$. We know that there exists a sequence of simple functions g_n, $g_n \to F$ almost everywhere and $0 \le g_n \le G$ for another element $G \in \mathcal{L}^1(\mathbb{R})$. Then set

$$u_n(x) = \begin{cases} g_n(x) & \text{if } \operatorname{Re} f_n(x) > g_n(x) \\ \operatorname{Re} f_n(x) & \text{if } -g_n(x) \le \operatorname{Re} f_n(x) \le g_n(x) \\ -g_n(x) & \text{if } \operatorname{Re} f_n(x) < -g_n(x). \end{cases} \tag{2.106}$$

Thus $u_n = \max(v_n, -g_n)$ where $v_n = \min(\operatorname{Re} f_n, g_n)$ so u_n is simple and $u_n \to f$ almost everywhere. Since $|u_n| \le G$ it follows from Lebesgue Dominated Convergence that $\operatorname{Re} f \in \mathcal{L}^1(\mathbb{R})$. The same argument shows $\operatorname{Im} f = -\operatorname{Re}(if) \in \mathcal{L}^1(\mathbb{R})$ so $f \in \mathcal{L}^1(\mathbb{R})$ as claimed. □

11. The spaces $L^p(\mathbb{R})$

We use Lemma 2.11 as a model:

Definition 2.9. For $1 \le p < \infty$ we set

$$\mathcal{L}^p(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{C}; f \text{ is measurable and } |f|^p \in \mathcal{L}^1(\mathbb{R})\}. \tag{2.107}$$

For $p = \infty$ we set

$$\mathcal{L}^\infty(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{C}; f \text{ measurable and } \exists\ C \text{ s.t. } |f(x)| \le C \text{ a.e}\} \tag{2.108}$$

Observe that, in view of Lemma 2.10, the case $p = 2$ gives the same space as Definition 2.4.

Proposition 2.9. *For each* $1 \le p < \infty$,

$$\|u\|_{L^p} = \left(\int |u|^p\right)^{\frac{1}{p}} \tag{2.109}$$

is a seminorm on the linear space $\mathcal{L}^p(\mathbb{R})$ *vanishing only on the null functions and making the quotient* $L^p(\mathbb{R}) = \mathcal{L}^p(\mathbb{R})/\mathcal{N}$ *into a Banach space.*

Proof. The real part of an element of $\mathcal{L}^p(\mathbb{R})$ is in $\mathcal{L}^p(\mathbb{R})$ since it is measurable and $|\operatorname{Re} f|^p \le |f|^p$ so $|\operatorname{Re} f|^p \in \mathcal{L}^1(\mathbb{R})$. Similarly, $\mathcal{L}^p(\mathbb{R})$ is linear; it is clear that $cf \in \mathcal{L}^p(\mathbb{R})$ if $f \in \mathcal{L}^p(\mathbb{R})$ and $c \in \mathbb{C}$ and the sum of two elements, f, g, is measurable and satisfies $|f + g|^p \le 2^p(|f|^p + |g|^p)$ so $|f + g|^p \in \mathcal{L}^1(\mathbb{R})$.

We next strengthen (2.107) to the approximation condition that there exists a sequence of simple functions v_n such that

$$v_n \to f \text{ a.e. and } |v_n|^p \le F \in \mathcal{L}^1(\mathbb{R}) \text{ a.e.} \tag{2.110}$$

which certainly implies (2.107). As in the proof of Lemma 2.11, suppose $f \in \mathcal{L}^p(\mathbb{R})$ is real and choose f_n real-valued simple functions and converging to f almost everywhere. Since $|f|^p \in \mathcal{L}^1(\mathbb{R})$ there is a sequence of simple functions $0 \le h_n$ such that $|h_n| \le F$ for some $F \in \mathcal{L}^1(\mathbb{R})$ and $h_n \to |f|^p$ almost everywhere. Then set $g_n = h_n^{\frac{1}{p}}$ which is also a sequence of simple functions and define v_n by (2.106). It follows that (2.110) holds for the real part of f but combining sequences for real and imaginary parts such a sequence exists in general.

The advantage of the approximation condition (2.110) is that it allows us to conclude that the triangle inequality holds for $\|u\|_{L^p}$ defined by (2.109) since we know it for simple functions and from (2.110) it follows that $|v_n|^p \to |f|^p$ in $\mathcal{L}^1(\mathbb{R})$ so $\|v_n\|_{L^p} \to \|f\|_{L^p}$. Then if w_n is a similar sequence for $g \in \mathcal{L}^p(\mathbb{R})$

$$\|f+g\|_{L^p} \le \limsup_n \|v_n+w_n\|_{L^p} \le \limsup_n \|v_n\|_{L^p}+\limsup_n \|w_n\|_{L^p} = \|f\|_{L^p}+\|g\|_{L^p}. \tag{2.111}$$

The other two conditions being clear it follows that $\|u\|_{L^p}$ is a seminorm on $\mathcal{L}^p(\mathbb{R})$.

The vanishing of $\|u\|_{L^p}$ implies that $|u|^p$ and hence $u \in \mathcal{N}$ and the converse follows immediately. Thus $L^p(\mathbb{R}) = \mathcal{L}^p(\mathbb{R})/\mathcal{N}$ is a normed space and it only remains to check completeness.

We know that completeness is equivalent to the convergence of any absolutely summable series. So, we can suppose $f_n \in \mathcal{L}^p(\mathbb{R})$ have

$$\sum_n \left(\int |f_n|^p\right)^{\frac{1}{p}} < \infty. \tag{2.112}$$

Consider the sequence $g_n = f_n\chi_{[-R,R]}$ for some fixed $R > 0$. This is in $\mathcal{L}^1(\mathbb{R})$ and

$$\|g_n\|_{L^1} \le (2R)^{\frac{1}{q}}\|f_n\|_{L^p} \tag{2.113}$$

by the integral form of Hölder's inequality
(2.114)

$$f \in \mathcal{L}^p(\mathbb{R}),\ g \in \mathcal{L}^q(\mathbb{R}),\ \frac{1}{p} + \frac{1}{q} = 1 \Longrightarrow fg \in \mathcal{L}^1(\mathbb{R}) \text{ and } |\int fg| \le \|f\|_{L^p} \||g\|_{L^q}$$

which can be proved by the same approximation argument as above, see Problem **??**. Thus the series g_n is absolutely summable in L^1 and so converges absolutely almost everywhere. It follows that the series $\sum_n f_n(x)$ converges absolutely almost everywhere – since it is just $\sum_n g_n(x)$ on $[-R, R]$, to a function, f.

So, we only need show that $f \in \mathcal{L}^p(\mathbb{R})$ and that $\int |f - F_n|^p \to 0$ as $n \to \infty$ where $F_n = \sum_{k=1}^n f_k$. By Minkowski's inequality we know that $h_n = (\sum_{k=1}^n |f_k|)^p$ has bounded L^1 norm, since

$$\||h_n|\|_{L^1}^{\frac{1}{p}} = \| \sum_{k=1}^n |f_k|\|_{L^p}. \le \sum_k \|f_k\|_{L^p}. \tag{2.115}$$

Thus, h_n is an increasing sequence of functions in $\mathcal{L}^1(\mathbb{R})$ with bounded integral, so by the Monotonicity Lemma it converges a.e. to a function $h \in \mathcal{L}^1(\mathbb{R})$. Since $|F_n|^p \le h$ and $|F_n|^p \to |f|^p$ a.e. it follows by Dominated convergence that

$$|f|^p \in \mathcal{L}^1(\mathbb{R}),\ \||f|^p\|_{L^1}^{\frac{1}{p}} \le \sum_n \|f_n\|_{L^p} \tag{2.116}$$

and hence $f \in \mathcal{L}^p(\mathbb{R})$. Applying the same reasoning to $f - F_n$ which is the sum of the series starting at term $n + 1$ gives the norm convergence:

$$\|f - F_n\|_{L^p} \le \sum_{k>n} \|f_k\|_{L^p} \to 0 \text{ as } \ n \to \infty. \tag{2.117}$$

□

A function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is *locally integrable* if

$$F_{[-N,N]} = \begin{cases} f(x) & x \in [-N, N] \\ 0 & x \text{ if } |x| > N \end{cases} \Longrightarrow F_{[-N,N]} \in \mathcal{L}^1(\mathbb{R})\ \forall\ N. \tag{2.118}$$

So any continuous function on $\mathbb{R}$ is locally integrable as is any element of $\mathcal{L}^1(\mathbb{R})$.

LEMMA 2.12. *The locally integrable functions form a linear space, $\mathcal{L}^1_{\text{loc}}(\mathbb{R})$ and*

$$\begin{gathered} \mathcal{L}^p(\mathbb{R}) = \{f \in \mathcal{L}^1_{\text{loc}}(\mathbb{R}); |f|^p \in \mathcal{L}^1(\mathbb{R})\}\ 1 \le p < \infty \\ \mathcal{L}^\infty(\mathbb{R}) = \{f \in \mathcal{L}^1_{\text{loc}}(\mathbb{R}); \sup_{\mathbb{R}\setminus E} |f(x)| < \infty \textit{ for some } E \textit{ of measure zero.}\} \end{gathered} \tag{2.119}$$

The proof is left as an exercise.

12. Lebesgue measure

In case anyone is interested in how to define Lebesgue measure from where we are now we can just use the integral.

DEFINITION 2.10. A set $A \subset \mathbb{R}$ is *measurable* if its characteristic function χ_A is locally integrable. A measurable set A has finite measure if $\chi_A \in \mathcal{L}^1(\mathbb{R})$ and then

$$\mu(A) = \int \chi_A \tag{2.120}$$

is the Lebesgue measure of A. If A is measurable but not of finite measure then $\mu(A) = \infty$ by definition.

We know immediately that any interval (a, b) is measurable (indeed whether open, semi-open or closed) and has finite measure if and only if it is bounded – then the measure is $b - a$.

PROPOSITION 2.10. *The complement of a measurable set is measurable and any* countable *union or countable intersection of measurable sets is measurable.*

PROOF. The first part follows from the fact that the constant function 1 is locally integrable and hence $\chi_{\mathbb{R}\setminus A} = 1 - \chi_A$ is locally integrable if and only if χ_A is locally integrable.

Notice the relationship between a characteristic function and the set it defines:-

$$\chi_{A\cup B} = \max(\chi_A, \chi_B),\ \chi_{A\cap B} = \min(\chi_A, \chi_B). \tag{2.121}$$

If we have a sequence of sets A_n then $B_n = \bigcup_{k\le n} A_k$ is clearly an increasing sequence of sets and

$$\chi_{B_n} \to \chi_B,\ B = \sum_n A_n \tag{2.122}$$

is an increasing sequence which converges pointwise (at each point it jumps to 1 somewhere and then stays or else stays at 0.) Now, if we multiply by $\chi_{[-N,N]}$ then

$$f_n = \chi_{[-N,N]}\chi_{B_n} \to \chi_{B\cap[-N,N]} \tag{2.123}$$

is an increasing sequence of integrable functions – assuming that is that the A_k's are measurable – with integral bounded above, by $2N$. Thus by the monotonicity lemma the limit is integrable so χ_B is locally integrable and hence $\bigcup_n A_n$ is measurable.

For countable intersections the argument is similar, with the sequence of characteristic functions decreasing. □

COROLLARY 2.1. *The (Lebesgue) measurable subsets of* $\mathbb{R}$ *form a collection,* $\mathcal{M}$, *of the power set of* $\mathbb{R}$, *including* $\emptyset$ *and* $\mathbb{R}$ *which is closed under complements, countable unions and countable intersections.*

Such a collection of subsets of a set X is called a 'σ-algebra' – so a σ-algebra Σ in a set X is a collection of subsets of X containing X, $\emptyset$, the complement of any element and countable unions and intersections of any element. A (positive) measure is usually defined as a map $\mu : \Sigma \longrightarrow [0, \infty]$ with $\mu(\emptyset) = 0$ and such that

$$\mu(\bigcup_n E_n) = \sum_n \mu(E_n) \tag{2.124}$$

for any sequence $\{E_m\}$ of sets in Σ which are disjoint (in pairs).

As for Lebesgue measure a set $A \in \Sigma$ is 'measurable' and if $\mu(A)$ is not of finite measure it is said to have infinite measure – for instance $\mathbb{R}$ is of infinite measure in this sense. Since the measure of a set is always non-negative (or undefined if it isn't measurable) this does not cause any problems and in fact Lebesgue measure is countably additive as in (2.124) provided we allow ∞ as a value of the measure. It is a good exercise to prove this!

13. Higher dimensions

I have never actually covered this in lectures – there is simply not enough time. Still it is worth knowing that the Lebesgue integral in higher dimensional Euclidean spaces can be obtained following the same line of reasoning. So, we want – with the advantage of a little more experience – to go back to the beginning and define $\mathcal{L}^1(\mathbb{R}^n)$, $L^1(\mathbb{R}^n)$, $\mathcal{L}^2(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n)$. In fact relatively little changes but there are some things that one needs to check a little carefully.

The first hurdle is that I am not assuming that you have covered the Riemann integral in higher dimensions; it is in my view a rather pointless thing to do anyway. Fortunately we do not really need that since we can just iterate the one-dimensional Riemann integral for continuous functions. So, define

$$\mathcal{C}_{\mathrm{c}}(\mathbb{R}^n) = \{u : \mathbb{R}^n \longrightarrow \mathbb{C};\ \text{continuous and such that } u(x) = 0 \text{ for } |x| > R\} \tag{2.125}$$

where of course the R can depend on the element. Now, if we hold say the last $n-1$ variables fixed, we get a continuous function of one variable which vanishes when $|x| > R$:

$$u(\cdot, x_2, \dots, x_n) \in \mathcal{C}_{\mathrm{c}}(\mathbb{R}) \text{ for each } (x_2, \dots, x_n) \in \mathbb{R}^{n-1}. \tag{2.126}$$

So we can integrate it and get a function

$$I_1(x_2, \dots, x_n) = \int_{\mathbb{R}} u(x, x_1, \dots, x_n),\ I_1 : \mathbb{R}^{n-1} \longrightarrow \mathbb{C}. \tag{2.127}$$

Lemma 2.13. *For each $u \in \mathcal{C}_c(\mathbb{R}^n)$, $I_1 \in \mathcal{C}_c(\mathbb{R}^{n-1})$.*

Proof. Certainly if $|(x_2, \dots, x_n)| > R$ then $u(\cdot, x_2, \dots, x_n) \equiv 0$ as a function of the first variable and hence $I_1 = 0$ in $|(x_2, \dots, x_n)| > R$. The continuity follows from the uniform continuity of a function on the compact set $|x| \le R$, $|(x_2, \dots, x_n) \le R$ of $\mathbb{R}^n$. Thus given $\epsilon > 0$ there exists $\delta > 0$ such that

$$|x - x'| < \delta,\ |y - y'|_{\mathbb{R}^{n-1}} < \delta \Longrightarrow |u(x, y) - u(x', y')| < \epsilon. \tag{2.128}$$

From the standard estimate for the Riemann integral,

$$|I_1(y) - I_1(y')| \le \int_{-R}^{R} |u(x, y) - u(x, y')| dx \le 2R\epsilon \tag{2.129}$$

if $|y - y'| < \delta$. This implies the (uniform) continuity of I_1. Thus $I_1 \in \mathcal{C}_{\mathrm{c}}(\mathbb{R}^{n-2})$ □

The upshot of this lemma is that we can integrate again, and hence a total of n times and so define the (iterated) Riemann integral as

$$\int_{\mathbb{R}^n} u(z) dz = \int_{-R}^{R} \int_{-R}^{R} \cdots \int_{-R}^{R} u(x_1, x_2, x_3, \dots, x_n) dx_1 dx_2 \dots dx_n \in \mathbb{C}. \tag{2.130}$$

Lemma 2.14. *The interated Riemann integral is a well-defined linear map*

$$\mathcal{C}_c(\mathbb{R}^n) \longrightarrow \mathbb{C} \tag{2.131}$$

which satisfies

$$|\int u| \le \int |u| \le (2R)^n \sup |u| \text{ if } u \in \mathcal{C}_c(\mathbb{R}^n) \text{ and } u(z) = 0 \text{ in } |z| > R. \tag{2.132}$$

Proof. This follows from the standard estimate in one dimension. □

Now, one slightly annoying thing is that we would really want to know that the integral is independent of the order of integration. In fact it is not hard – see Problem XX. Again using properties of the one-dimensional Riemann integral we find:-

Lemma 2.15. *The iterated integral*

$$\|u\|_{L^1} = \int_{\mathbb{R}^n} |u| \tag{2.133}$$

is a norm on $\mathcal{C}_c(\mathbb{R}^n)$.

Definition 2.11. The space $\mathcal{L}^1(\mathbb{R}^n)$ is defined to consist of those functions $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ such that there exists a sequence $\{f_n\}$ which is absolutely summable with respect to the L^1 norm and such that

$$\sum_n |f_n(x)| < \infty \Longrightarrow \sum_n f_n(x) = f(x). \tag{2.134}$$

Now you can go through the whole discusion above in this higher dimensional case, and the only changes are really notational!

Things get a littlem more complicated in the discussion of change of variable. This is covered in the problems. There are also a few other theorems it is good to know!