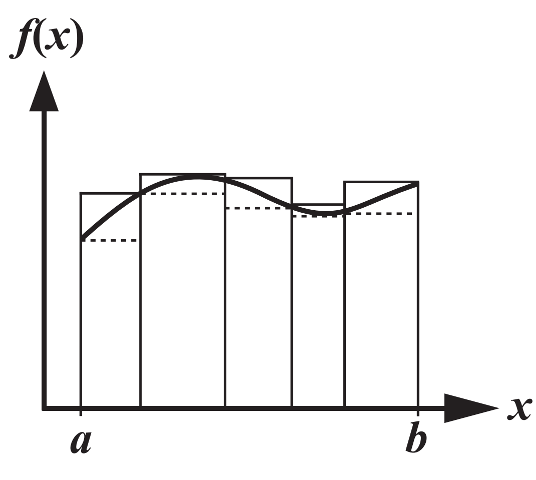
AMSI 2013: MEASURE THEORY
Handout 1
Introduction to Measure Theory

Marty Ross
martinirossi@gmail.com

January 7, 2013

NOTE: *This introduction is intended to be a quick, general and gentle overview, and necessarily the definitions and details are vague. We'll be much more careful when we begin the formal study of measures, beginning with Handout 2.*

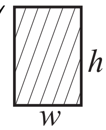
The standard form of integration one sees in high school, and in early undergraduate courses, is *Riemann Integration.* Here, the intuitive "area" under the graph of a function $f:[a,b]\to\mathbb{R}$ is approximated by the sums of areas of rectangles.



If we consider a "lower sum" and an "upper sum" then

$$L_\alpha = \sum \text{area(lower rectangles)} \leqslant \text{area under graph}(f) \leqslant \sum \text{area(upper rectangles)} = U_\alpha\,.$$

Here, α denotes some *partition* of the interval $[a, b]$, and then L_α and U_α are the corresponding lower and upper sums. Of course, we're also applying here the fundamental notion of the area of a rectangle:

Area () $= w \cdot h\,.$

The idea, then, is to consider finer and finer partitions of $[a, b]$, giving (hopefully) better and better approximations to the precise area. *If*, as we take the (suitable) limit of finer partitions, we find that

$$\lim_{\alpha} L_\alpha = \lim_{\alpha} U_\alpha\,,$$

then we define the *Riemann integral* of f on $[a, b]$ to be that common limit:

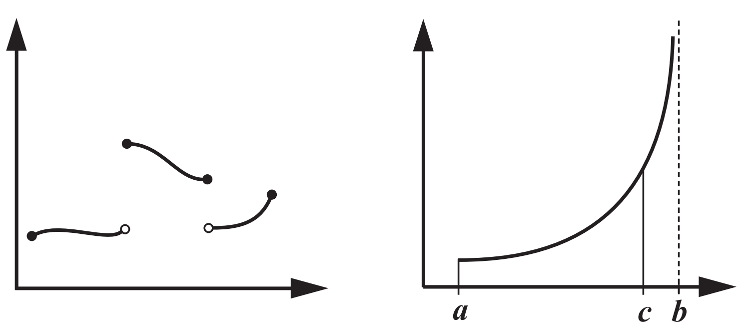
$$\lim_{\alpha} L_\alpha = \int_a^b f = \lim_{\alpha} U_\alpha\,.$$

We then say that f is *Riemann integrable*, meaning the integral – intuitively the area under the graph – makes sense. (Of course, this is an entirely different question to the practical calculation of $\int f$ for any specific f). We then have

Fundamental Riemann Theorem:

If f is continuous on $[a, b]$ then f is Riemann Integrable on $[a, b]$.

Of course, more than just continuous functions are Riemann integrable. The obvious generalization is to functions with a finite number of (suitably benign) discontinuities. Another important class is that of *improper integrals*, where the functions under consideration may have one or more vertical or horizontal asymptotes.



For example, for the function pictured on the right, we would define (as long as it exists)

$$\int_a^b f = \lim_{c \to b^-} \int_a^c f\,.$$

So, though such a function cannot technically be Riemann integrable,[1] if f is sufficiently well behaved then we can still use Riemann integration to make sense of the entire integral.

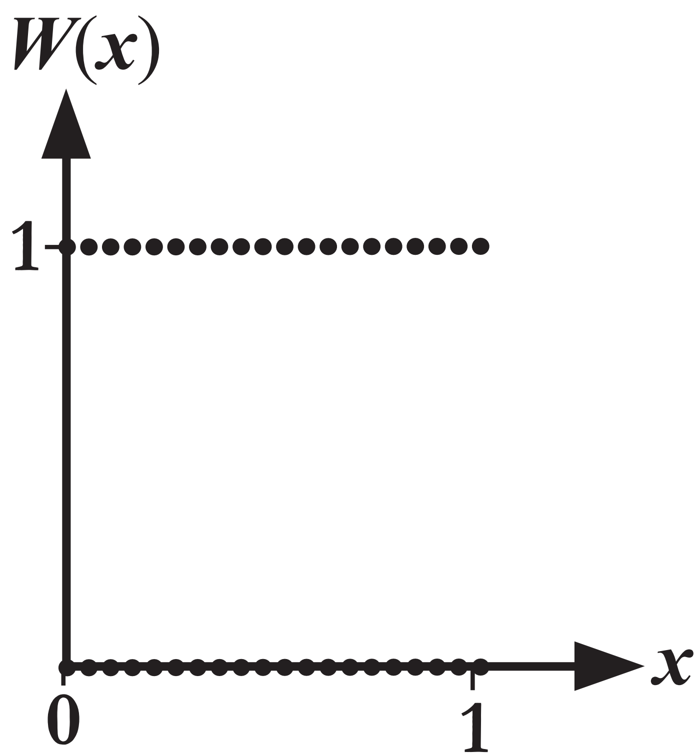
So, what are the shortcomings of Riemann integration? The first thing to realise is:

Problem with Riemann, Version 1: Not all functions are Riemann integrable.

For example, consider the Weird function $W : [0,1] \to \mathbb{R}$, defined by

$$W(x) = \begin{cases} 1 & x \in \mathbb{Q}\,, \\ 0 & x \in \sim\mathbb{Q}\,. \end{cases}$$

The function W is not Riemann integrable: since $\mathbb{Q}$ and $\sim\mathbb{Q}$ are both dense in $[0,1]$, any lower sum satisfies $L_\alpha \leqslant 0$, and any upper sum satisfies $U_\alpha \geqslant 1$.



This may not seem like much of an issue: the function W is seemingly cooked up, and a natural reaction is to remark that the area under the graph of such a Weird function *shouldn't* make sense. However, we also have:

[1]Clearly, a function f must be bounded in order for the upper and lower sums of f to make sense, and it is not hard to show that we also need the domain (i.e. the support) of f to be bounded. This upper-lower approach is due to Darboux, but also with Riemann's original approach, the same boundedness hypotheses are required on f. However, modern generalisations of the Riemann integral can apply directly to unbounded functions. See Handout 5.

Problem with Riemann, Version 2: There is a sequence $\{f_j\}_{j=1}^{\infty}$ of uniformly bounded functions on $[a,b]$,[2] and a function f, such that:

$$\begin{cases} \text{each } f_j \text{ is Riemann integrable;} \\ f_j \to f \text{ pointwise (that is, for any fixed } x \in [a,b] \text{, we have } f_j(x) \to f(x)); \\ f \text{ is } \textit{not} \text{ Riemann integrable.} \end{cases}$$

We'll give an example in a moment, but notice that this really is an issue. In both pure and applied mathematics, we are constantly taking limits of sequences of functions, and it is definitely of concern if the integral of the limiting function needn't make sense. To make it more concrete, consider the question

$$\lim_{j\to\infty} \int_a^b f_j \stackrel{?}{=} \int_a^b \lim_{j\to\infty} f_j\,.$$

This kind of manipulation, effectively the interchanging of limits, is something one does all the time. For the promised example, much less than the two sides being equal, the right hand side (i.e. $\int f$) is not even defined.

So what is an example of such a badly behaved limit? Since $\mathbb{Q}$ is a countable set, we can list the rationals in $[0,1]$ as a sequence:

$$\mathbb{Q} \cap [0,1] = \{q_1, q_2, \dots\}\,.$$

We then define

$$f_j(x) = \begin{cases} 1 & x = q_1, q_2, \dots, q_j, \\ 0 & \text{otherwise.} \end{cases}$$

That is, one by one, we raise the value of the rationals to 1. Now, each f_j has a finite number of harmless discontinuities, and thus is Riemann integrable. And one can (and should) check that $f_j \to W$ pointwise for each $x \in [0,1]$,[3] where W is the function defined previously. And, we have already noted that W is not Riemann integrable.

This may seem like a contrived example, but here is an alternative, less contrived way of defining W as double limit of very nice functions:

$$W(x) = \lim_{j\to\infty} \lim_{k\to\infty} \left(\cos(j!\pi x)\right)^{2k}\,.$$

The point is, Riemann integration is premised upon the functions integrated being relatively nice, and it is very easy for the limit of nice functions to be not-nice. True, such problems

[2] That is, there is an $M \in \mathbb{R}$ such that $|f_j(x)| \leqslant M$ for all $j \in \mathbb{N}$ and all $x \in [a,b]$.

[3] Consider separately the two cases: when x is irrational; and, when x is rational.

don't arise if the sequence of functions converges *uniformly*,[4] but uniform convergence is a very strong, and often unsatisfied, hypothesis.

There are in fact ways to deal directly with these problems of Riemann integration: firstly, one can give general hypotheses that guarantee the problems above do not arise; secondly, there are modern generalisations of the Riemann integral which do not suffer the same drawbacks. We'll briefly discuss both of these approaches in Handout 5. Nonetheless, as we shall see, the measure theory approach avoids the Riemann problems entirely, and also provide an excellent framework for integration in much more general contexts.

Lebesgue integration, which was discovered/invented early in the 20th Century is a largely successful attempt to avoid the problems of Riemann integration. The formal definitions will come later, but the idea is to assign, to any $A \subseteq \mathbb{R}$ the *Lebesgue measure* $\mathscr{L}(A)$ of A. $\mathscr{L}(A)$ is supposed to somehow measure the "size" or "length" of A. If A is a crazy set, it is not at all obvious what $\mathscr{L}(A)$ should be, but it is at least clear that we want

$$(*) \quad \mathscr{L}([a,b]) = b - a\,.$$

As well, $\mathscr{L}$ should behave in a way which reasonably reflects some notion of size. In particular, we want

$$(**) \quad \mathscr{L}(A \cup B) \leqslant \mathscr{L}(A) + \mathscr{L}(B) \quad \text{(and hopefully} = \mathscr{L}(A) + \mathscr{L}(B) \text{ if } A \cap B = \emptyset)\,.$$

We'll be precise later, but (i) and (ii) give us some sense of how we want Lebesgue measure to behave.[5]

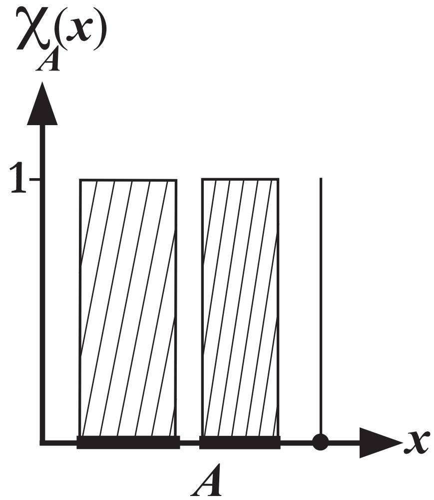
How do we then use Lebesgue measure to define integrals? The idea is to begin with what are called *characteristic functions*. Given $A \subseteq \mathbb{R}$, the characteristic function $\chi_A : \mathbb{R} \to \mathbb{R}$ is defined by

$$\chi_A(x) = \begin{cases} 1 & x \in A\,, \\ 0 & x \in \sim A\,. \end{cases}$$

So, for example, the Weird function W is the characteristic function $\chi_{\mathbb{Q} \cap [0,1]}$.

[4]Fix j and suppose $|f_j - f| < \epsilon$ everywhere on $[a,b]$. Then f is within a band around f_j of width $b-a$ and height ϵ. This means upper and lower sums for f are within $\epsilon \cdot (b-a)$ of the corresponding upper and lower sums for f_j. The uniform convergence of $\{f_j\}$ to f then makes it easy to to show f is Riemann integrable, with the desired limit integral.

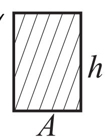
[5]The main further issue is that we'll want a countable version of $(**)$. This is motivated by the fact that an integral, however defined, amounts to a countable sum, or a limit.



Then we define the *Lebesgue integral* of χ_A to be

$$\int \chi_A \, \mathrm{d}\mathscr{L} = \mathscr{L}(A) \cdot 1 = \mathscr{L}(A)\,.$$

This is just the Lebesgue version of the area of a rectangle:

Area () $= \mathscr{L}(A) \cdot h\,.$

For a general function f, we attempt to approximate f by linear combinations of characteristic functions (so-called *simple functions*), and then f is *Lebesgue integrable* if a suitable limit of the integrals converges: this is the Lebesgue analogue of the Riemann process of approximating by lower and upper sums. Not surprisingly, the details are technical, but the underlying ideas are very natural.

Now, how does Lebesgue integration compare to Riemann? We have

Fundamental "Lebesgue is better than Riemann" Theorem:

(a) *Any Riemann integrable function is Lebesgue integrable, with the same integral value.*[6]

(b) *Suppose* $\{f_j\}_{j=1}^{\infty}$ *is a sequence of uniformly bounded Lebesgue integrable functions on* $[a,b]$*, and suppose* $f_j \to f$ *pointwise. Then* f *is Lebesgue integrable and*

$$\lim_{j\to\infty} \int f_j \, \mathrm{d}\mathscr{L} \;=\; \int \lim_{j\to\infty} f_j \, \mathrm{d}\mathscr{L} \;=\; \int f \, \mathrm{d}\mathscr{L}\,.$$

[6] There do exist functions which fail to be Lebesgue integrable but are improperly Riemann integrable. We'll discuss this later.

As a simple example, this theorem tells us that the Weird function W is Lebesgue integrable, and that $\int W \, \mathrm{d}\mathscr{L} = 0$. However, we are not suggesting that Lebesgue solves all our problems: there are still functions which are not Lebesgue integrable; and, we still need to be careful when taking limits (note the hypothesis of uniform boundedness). But there is a fundamental manner in which Lebesgue is more robust than Riemann in the taking of limits. In fact, in a certain sense we can consider the space of Lebesgue integrable functions as the *completion* of the space of continuous functions, in the same manner as the set $\mathbb{R}$ of real numbers is the completion $\overline{\mathbb{Q}}$ of the set of rationals. This leads to the very important theory of L^p *spaces*, which we shall consider in Handout 6.

Once we have the notion of Lebesgue measure, various generalizations are possible. The immediate thought is of *m-dimensional Lebesgue measure*, $\mathscr{L}^m$ on $\mathbb{R}^m$. Here, the ideas is to capture the notion of m-dimensional volume of subsets of $\mathbb{R}^m$. So, similar to considering the length of an interval in $\mathbb{R}$, the fundamental property we want is:

$$\begin{aligned}\mathscr{L}^m\left([a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_m, b_m]\right) &= (b_1 - a_1) \cdot (b_2 - a_2) \cdots (b_m - a_m) \\ &= \text{the } m\text{-volume of an } m\text{-box.}\end{aligned}$$

Then, for suitable $A \subseteq \mathbb{R}^m$ and $f : A \to \mathbb{R}^m$, we can define $\int_A f \, \mathrm{d}\mathscr{L}^m$, representing the $(m+1)$-volume under the graph of f. So, note that we have two methods of calculating $(m+1)$-volumes; by the $\mathscr{L}^m$-integral of functions over $\mathbb{R}^m$; and by the $\mathscr{L}^{m+1}$-measure of subsets of $\mathbb{R}^{m+1}$.

One can go on to consider a general measure μ on an arbitrary set X: so, for each $A \subseteq X$, $\mu(A)$ is some notion of the size of A. As for Lebesgue measure, we would want

$$(**) \quad \mu(A \cup B) \leqslant \mu(A) + \mu(B) \quad (\text{and hopefully} = \mu(A) + \mu(B) \text{ if } A \cap B = \emptyset)\,.$$

Then, by first considering characteristic functions, we can define the integral $\int_A f \, \mathrm{d}\mu$ for suitable $A \subseteq X$ and $f : A \to \mathbb{R}$.

There are three broad approaches to general measure theory (with lots of room for overlap):

Functional-analytic

One notes that a measure μ can be thought of as a *linear operator* T_μ on a suitable class of functions $f : X \to \mathbb{R}$:

$$T_\mu(f) = \int_X f \, d\mu \, .$$

Conversely, given a linear operator T on functions on X, we may hope to find a measure μ for which $T = T_\mu$. Thus, measure theory fits very naturally into the world of (Banach and Hilbert) spaces of functions.

Probabilistic

Here X is a space of possible *outcomes*, with $\mu(X) = 1$ (meaning the probability is certain that something will happen). Then, for $A \subseteq X$, $\mu(A)$ is the probability that the *event* A will occur. The machinery of probabilistic measure theory is much the same, but the language tends to be very different.

Geometric

Our approach in these notes will be largely geometric. Here, the concern is to focus upon specific measures that reflect the underlying geometry of Euclidean space $\mathbb{R}^m$ or, more generally, a metric space X. Lebesgue measure is a clear example of this. Another very important example is *m-dimensional Hausdorff measure*, $\mathscr{H}^m$, which gives the notion of m-dimensional volume of subsets of Euclidean space $\mathbb{R}^n$, or in fact of any metric space.

The focus upon specific and geometrically motivated measures allows us to correspondingly prove strongly geometric theorems. These theorems are not only beautiful in themselves, they are also central to the modern study of PDEs and differential geometry.