AMSI 2013: MEASURE THEORY
Handout 2
General Measures and Lebesgue Measure

Marty Ross
martinirossi@gmail.com

January 9, 2013

THE EXTENDED REAL NUMBERS

In what is to follow, we'll definitely want to consider the possibility of sets having infinite measure: for example, $\mathbb{R}$ will have infinite Lebesgue measure. Thus, in order to give a precise definition of a measure, we first need to consider the concept of ∞ in a clear and explicit manner. To do this, we define the *extended real number system*, $\mathbb{R}^* = \mathbb{R} \cup \{\infty, -\infty\}$, consisting of $\mathbb{R}$ together with two new elements. To begin, ∞ and $-\infty$ are just two abstract objects, with no designated properties: we now have to decide how to extend the operations on $\mathbb{R}$ to operations on $\mathbb{R}^*$, so that the properties of $\pm\infty$ reflect our intuition.

As a first step, we can obviously extend the linear ordering of $\mathbb{R}$ to $\mathbb{R}^*$, by declaring that $-\infty < \infty$ and that

$$-\infty < a < \infty \quad \text{for all } a \in \mathbb{R}.$$

Having done so, it now makes sense to write $\mathbb{R}^*$ as a "closed" interval:

$$\mathbb{R}^* = [-\infty, \infty]\,.$$

We can now define bounds for subsets of $\mathbb{R}^*$, just as we do for $\mathbb{R}$. And, of course any subset of $\mathbb{R}^*$ is bounded above, by ∞ if nothing else. Then, by the least upper bound property of $\mathbb{R}$, we have:

Sup-Inf Property of $\mathbb{R}^*$.

Every $A \subseteq \mathbb{R}^*$ has a least upper bound and a greatest lower bound.

Of course, if $A \subseteq \mathbb{R}$ is non-empty and bounded above (below) by a real number, then $\sup A \in \mathbb{R}$ ($\inf A \in \mathbb{R}$) is the same, whether A is considered a subset of $\mathbb{R}$ or $\mathbb{R}^*$.

In a similar manner, we can consider sequences, in particular monotonic sequences, in $\mathbb{R}^*$. However, we have to be clear about what it means for a sequence $\{a_j\}$ in $\mathbb{R}^*$ to converge to $a \in \mathbb{R}^*$. If $a \in \mathbb{R}$ then there is no real issue: all but finitely many of the a_j must be real as well, and then $a_j \to a$ has just the same meaning as it does for real sequences. If $a = \pm\infty$, the simplest approach is to reinterpret the intuitive notion of $a_j \to \pm\infty$ as a formal definition. So, we can define $a_j \to \infty$ if, for every $N \in \mathbb{R}$, there is an $M \in \mathbb{N}$ such that

$$j \geqslant M \implies a_j > N\,.$$

Similarly, we can define $a_j \to -\infty$. With these definitions, we can then use the monotonic sequence property of $\mathbb{R}$ to prove:

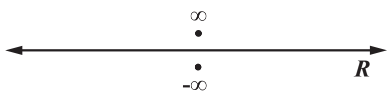
Monotonic Sequence Property of $\mathbb{R}^*$.

Every monotonic sequence in $\mathbb{R}^*$ converges.

It follows that the definitions and results of Handout 0 hold quite generally, without concern for whether the relevant limits might be infinite or not. In particular, we have that $\limsup a_j$ and $\liminf a_j$ are well-defined for any sequence $\{a_j\}$ in $\mathbb{R}^*$.

This ad hoc approach to sequences will suffice for our purposes. However, it does leave open the natural question of whether the convergence of sequences in $\mathbb{R}^*$ can be treated in a more systematic manner. That is, can we regard $\mathbb{R}^*$ as a *metric space*?

Recall that $\mathbb{R}$ is a metric space with the distance $d(a,b)$ from a to b defined as $d(a,b) = |a-b|$. Is it then possible to define a corresponding metric d^* on in $\mathbb{R}^*$? Stated as such, the question is too vague. As a silly response, we could just put the discrete metric on $\mathbb{R}^*$. But of course, the discrete metric has nothing to do with the way we want to think about $\mathbb{R}^*$. Alternatively, we could try to impose the condition that $d^*(a,b) = d(a,b)$ whenever $a, b \in \mathbb{R}$, so that d^* in some sense extends d. However, even this won't suffice. For example, we could simply interpret $\pm\infty = (0, \pm 1)$ as points in $\mathbb{R}^2$, and set $d^* = d$ to be Euclidean distance in $\mathbb{R}^2$. But of course, in this case $\pm\infty$ are not where we want them to be: in particular, it would then be impossible for a sequence $\{a_j\}$ of real numbers to converge to $\pm\infty$, which is the very notion we are trying to capture.



The one property we definitely want from d^* is that a sequence $\{a_j\}$ converges to ∞ with respect to d^* (i.e. $d^*(a_j, \infty) \to 0$) if and only if $a_j \to \infty$ in the special-case way that we defined above. If we then also try to demand that $d^*(a,b) = d(a,b)$ for $a, b \in \mathbb{R}$, it quickly becomes clear that *no* such metric d^* on $\mathbb{R}^*$ exists: we are forced to define $d^*(a, \infty) = \infty$, which is illegal. (Note that, even if we are making sense of ∞ in $\mathbb{R}^*$, a metric d^* on $\mathbb{R}^*$ – as for any metric on any metric space – must still be real-valued: we do not permit an infinite distance between two given points.)[1]

[1]This suggests that we might be able to generalise the concept of a metric space, to include the possibility

If the metric situation is unclear, at least we can naturally extend the *topology* of $\mathbb{R}$ to $\mathbb{R}^*$. Recall that on $\mathbb{R}$ we begin with the open intervals (a, b); then a set $A \subseteq \mathbb{R}$ is said to be open if it is a (possibly infinite) union of such open intervals. We can then define a topology on $\mathbb{R}^*$ in an identical manner: we declare that for any $a \in \mathbb{R}$, the intervals $[-\infty, a)$ and $(a, \infty]$ are also open, and then a set $A \subseteq \mathbb{R}^*$ is open if it is a union of open intervals. This is easily shown to be a topology on R^*, and we have

Exercise. $\mathbb{R}^*$ is compact with the topology just defined.

We can now note that, though $\mathbb{R}^*$ is not naturally a metric space, it is *metrizable*, and in a way which captures the correct notion of convergence to $\pm\infty$. To be precise:

Exercise. There is a metric d^* on $\mathbb{R}^*$ on such that $A \subseteq \mathbb{R}^*$ is open (in the topology defined above) iff A is open with respect to d^*. Furthermore,

$$a_j \to a \in \mathbb{R}^* \text{ (as defined in the special-case way above) iff } d^*(a_j, a) \to 0\,.$$

Of course, since $\mathbb{R}^*$ can now be considered a compact metric space, the metric characterizations of compactness now apply. In particular, any sequence in $\mathbb{R}^*$ must have a convergent subsequence.

Finally, we want to consider algebraic operations on $\mathbb{R}^*$. We shall declare

$$\begin{cases} \infty + \infty = \infty \cdot \infty = \infty & \\ \infty \pm x = \infty & x \in \mathbb{R}, \\ \infty \times x = \infty \text{ (or } -\infty) & x > 0 \text{ (or } x < 0), \\ \dfrac{1}{\infty} = 0\,. & \end{cases}$$

We make similar definitions for $-\infty$. With these definitions, the usual algebraic (field) properties of $\mathbb{R}$ (distributivity, associativity, etc) continue to hold, as long as we don't run into undefined quantities such as $\infty - \infty$ or $\dfrac{\infty}{\infty}$. Also, with more caution, we shall define

$$\infty \times 0 = 0\,.$$

This last equation is not intuitively true, and and so is sometimes referred to suspiciously as a "convention". However, in the context of measure theory it turns out to be a useful shorthand. We will spell out carefully when it is applied.

Note that, since the convergence of infinite series is defined in terms of infinite sequences of partial sums, the meaning of the convergence (or divergence) of such series in $\mathbb{R}^*$ is also resolved.

---

of infinite distances. That indeed can be done. However, so defined, *no* sequence $\{a_j\}$ of real numbers can converge to ∞, since all of the a_j will be infinitely far away. So, once again, this would fail to capture the intuitive notion of sequences converging to infinity.

GENERAL MEASURES

Historically, much work went into coming up with just the right definitions in measure theory, and proving exactly what one could and could not expect to gain from such concepts. We shall avoid almost all of this groundwork, and simply begin with the definition which works best for us.[2]

A *measure* μ on a set X is a function $\mu : \wp(X) \to [0, \infty]$ such that

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

(ii) $A \subseteq B \implies \mu(A) \leqslant \mu(B)$ for any $A, B \subseteq X$ (monotonicity),

(iii) $\mu\left(\bigcup_{j=1}^{\infty} A_j\right) \leqslant \sum_{j=1}^{\infty} \mu(A_j)$ for any $A_1, A_2, \dots \subseteq X$ (countable subadditivity).

What we refer to as a measure, many texts will refer to as an *outer measure*. The critical point is that, for us, a measure on a set X must be defined on *all* subsets of X. What we do not demand is *countable additivity*, i.e. equality in (iii) in the case that the sets A_j are pairwise disjoint; we'll address this in the next Handout, when we consider the so-called *measurable sets*. Notice also that we do *not* ever need to apply the equation $0 \cdot \infty = 0$ to evaluate the right hand side of (iii): if each $\mu(A_j) = 0$ then $\sum \mu(A_j) = 0$ directly by the definition of convergence of infinite series, as a sequence of partial sums.

We are keen to define Lebesgue measure, but we first consider some simpler examples.

Dirac Measure: For any set X and any fixed $a \in X$ we define

$$\mu_a(A) = \begin{cases} 1 & a \in A, \\ 0 & a \notin A. \end{cases}$$

The fact that μ_a is a measure is quite trivial, but let's quickly give the proof of (iii). If the LHS of (iii) is 0 then (iii) is obvious. On the other hand, if the LHS of (iii) is 1 (the only other possibility), that means $a \in \cup A_j$. But then $a \in A_j$ for some specific j, implying $\mu_a(A_j) = 1$. Thus the RHS of (iii) is at least 1, and we have LHS $\leqslant$ RHS, as desired.

Anything-Will-Do Measure: For any set X, we define

$$\mu(A) = \begin{cases} 1 & A \neq \emptyset, \\ 0 & A = \emptyset. \end{cases}$$

[2]For a thorough treatment of this groundwork, see *Measure Theory* by Paul Halmos (Springer, 1978).

Anything-Is-Wonderful Measure: For any set X, we define

$$\mu(A) = \begin{cases} \infty & A \neq \emptyset, \\ 0 & A = \emptyset. \end{cases}$$

Infinite-Is-Better Measure: For an infinite set X, we define

$$\mu(A) = \begin{cases} \infty & A \text{ is infinite}, \\ 0 & A = \emptyset, \\ 1 & \text{otherwise}. \end{cases}$$

Here, to prove (iii), note that if $\bigcup A_j$ is infinite then either one of the A_j is infinite, or infinitely many of the A_j are non-empty.

Eeverything-Is-Better Measure: For any set X with at least 2 elements, we define

$$\mu(A) = \begin{cases} 3 & A = X, \\ 0 & A = \emptyset, \\ 2 & \text{otherwise}. \end{cases}$$

Tons-Will-Do Measure: For an uncountable set X, we define

$$\mu(A) = \begin{cases} 1 & A \text{ is uncountable}, \\ 0 & A \text{ is countable}. \end{cases}$$

For this last measure, note that (iii) follows from the fact that a countable union of countable sets is countable.

A more interesting and useful measure is

Counting Measure: For any set X, we define

$$\mu^0(A) = \begin{cases} \text{the number of elements in } A & A \text{ is finite}, \\ \infty & A \text{ is infinite}. \end{cases}$$

The fact that μ^0 is a measure is intuitive but a bit fiddly to prove. If X is countable then μ^0 is just the sum of all the delta measures on X:

$$\mu^0(A) = \sum_{a \in X} \mu_a(A) \, .$$

If X is uncountable (e.g. $X = \mathbb{R}$), we can still do something similar. We then interpret the sum as the supremum over the possible finite sums corresponding to finite $F \subseteq X$. The work then required to show counting measure is always a measure is then nicely encapsulated in the following exercises:

③ **Exercise**: Suppose μ and ν are measures on a set X, and $a, b \in [0, \infty]$. Then $a\mu + b\nu$ is also a measure on X.

Of course, by induction, it then follows that any *finite* nonnegative linear combination of measures is a measure. The extension to the (even uncountably) infinite is then given by

④ **Exercise**: Suppose $\{\mu_\alpha\}_{\alpha \in I}$ is a collection of measures on a set X, and define $\mu : \wp(X) \to [0, \infty]$ by

$$\mu(A) = \sup_{\alpha \in I} \mu_\alpha(A) \, .$$

Then μ is a measure on X.

Of course ③ and ④ can be applied more generally, but it's not of much interest unless we have interesting measures to sum or sup. For that, we need to do some work.

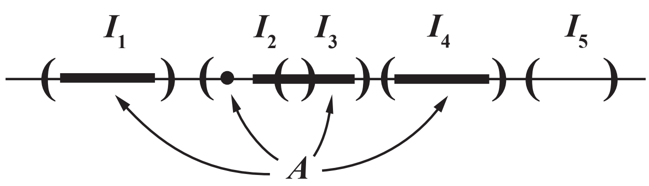
LEBESGUE MEASURE

Recall that Lebesgue measure $\mathscr{L}$ on $\mathbb{R}$ is designed to extend to the notion of the length of an interval I to arbitrary sets. So, we begin by defining $l(I)$ to be the length of the interval I. In particular

$$l((a, b)) = b - a \, ,$$

with $l(I)$ defined similarly for closed and half-open intervals: also, the interval I can be infinite in extent, in which case, of course $l(I) = \infty$. Now, given $A \subseteq \mathbb{R}$, we consider covering A by a countable collection $\{I_j\}_{j=1}^{\infty}$ of open intervals:[3]

$$A \subseteq \bigcup_{j=1}^{\infty} I_j \, .$$



[3]The collection $\{I_j\}$ is also permitted to be finite. For notation pedants, one can include this situation by setting all but finitely many $I_j = \emptyset$.

Then, the "length" of A should be no greater than $\sum l(I_j)$. On the other hand, it may be strictly less, because of the overlap of the I_j, or because of poorly placed intervals. So, we want to consider all such coverings, and to consider $\sum l(I_j)$ for those coverings which are most efficient. This leads us to the precise definition of *Lebesgue measure* on $\mathbb{R}$:

$$\boxed{\mathscr{L}(A) = \inf\left\{\sum_{j=1}^{\infty} l(I_j) : A \subseteq \bigcup_{j=1}^{\infty} I_j,\ \text{each } I_j \text{ an open interval}\right\}} \qquad A \subseteq \mathbb{R}$$

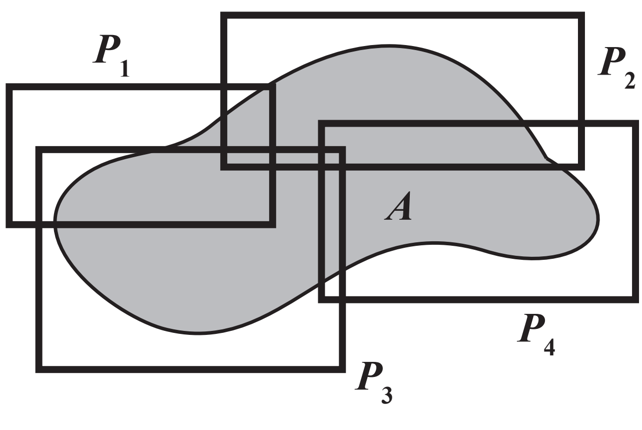
We shall prove that Lebesgue measure is inded a measure, but we first consider the generalization to $\mathbb{R}^m$. Here the fundamental notion is the m-dimensional volume v of an m-box P:

$$P = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_m, b_m) \implies v(P) = (b_1 - a_1) \cdot (b_2 - a_2) \cdots (b_m - a_m)\,.$$

As for intervals, we also allow the possibility of an m-box P being open, or closed, or somewhere inbetween (some parts of some edges included but not others): in all cases, the volume $v(P)$ is defined in the obvious manner. Note, however, that the m-box must be oriented as indicated above, with its edges parallel to the coordinate axes.

In exact analogy to the 1-dimensional situation, we now consider coverings of a set $A \subseteq \mathbb{R}^m$ by open m-boxes.This leads to the definition of *m-dimensional Lebesgue measure*:[4]

$$\boxed{\mathscr{L}^m(A) = \inf\left\{\sum_{j=1}^{\infty} v(P_j) : A \subseteq \bigcup_{j=1}^{\infty} P_j,\ \text{each } P_j \text{ an open } m\text{-box}\right\}} \qquad A \subseteq \mathbb{R}^m$$



[4]Though the approach here is natural, most texts define m-dimensional Lebesgue measure by the use of product measures. See Handout 8.

Notice that it is *not* obvious that $\mathscr{L}^m(P) = v(P)$ for P an m-box, even when $m = 1$. We shall indeed prove that, but first things first. We shall first prove that $\mathscr{L}^m$, including $\mathscr{L}^1 = \mathscr{L}$ as a special case, is in fact a measure. For this purpose, and in general for proving properties of $\mathscr{L}^m$, it is worth spelling out the exact meaning of the "inf" in the definition:

(♡) For *any* covering $\{P_j\}$ of A, we have

$$\mathscr{L}^m(A) \leqslant \sum v(P_j)\,.$$

(♠) For any $\epsilon > 0$ there *exists* a covering $\{P_j\}$ of A with

$$\sum v(P_j) \leqslant \mathscr{L}^m(A) + \epsilon\,.$$

In particular, any covering of A gives an upper bound for $\mathscr{L}^m$, which usually makes the establishment of the desired upper bound quite easy. For lower bounds, on the other hand, we have to consider all coverings, which can sometimes be tricky.

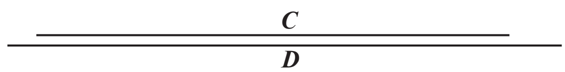
We now prove that $\mathscr{L}^m$ is a measure, including $\mathscr{L} = \mathscr{L}^1$ as a special case.

PROOF OF (i) *FOR* $\mathscr{L}^m$*:* Covering $\emptyset$ by $\emptyset$ ($a_1 = b_1$, etc.) we obviously have $\mathscr{L}^m(\emptyset) \leqslant 0$. But any covering of a set gives a nonnegative estimate, and so clearly $\mathscr{L}^m(\emptyset) \geqslant 0$. Thus $\mathscr{L}^m(\emptyset) = 0$.

PROOF OF (ii) *FOR* $\mathscr{L}^m$*:* This is also pretty easy: if $A \subseteq B$ then it is easier to cover A, and so the *inf* should be smaller. For general purposes, we'll isolate an explicit argument that we use:

LEMMA 1: Thrilling Sup-Inf Lemma

$$C \subseteq D \subseteq R^* \implies \begin{cases} \inf(D) \leqslant \inf(C)\,, \\ \sup(D) \geqslant \sup(C)\,. \end{cases}$$



Now, to prove (ii), we note

$A \subseteq B$

$\implies$ any covering of B is also a covering of A

$\implies$ the collection $\mathcal{D}$ of A-coverings is *larger* than the collection $\mathcal{C}$ of B-coverings

$\implies$ the *inf* of $D = \{\sum v(P_j) : A \subseteq \cup P_j\}$ is *smaller* than the *inf* of $C = \{\sum v(P_j) : B \subseteq \cup P_j\}$

$\implies$ $\mathscr{L}^m(A) \leqslant \mathscr{L}^m(B)$

PROOF OF (iii) *FOR* $\mathscr{L}^m$*:* Again, this is easy in principle, as any coverings for A_j can be combined to give a covering for $\cup A_j$. Again, we'll isolate a useful lemma.

5b LEMMA 2: Thrilling $\boldsymbol{\epsilon}$-Lemma

Suppose $a, b \in \mathbb{R}^*$ and suppose that for every $\epsilon > 0$, we have

$$a \leqslant b + \epsilon\,.$$

Then $a \leqslant b$.

Now, to prove (iii), fix $\epsilon > 0$ and for each j let $\{P_{jk}\}_{k=1}^{\infty}$ be a covering of A_j such that

$$\sum_{k=1}^{\infty} v(P_{jk}) \leqslant \mathscr{L}^m(A_j) + \frac{\epsilon}{2^j} \qquad \text{(possible, by } \spadesuit\text{)}\,.$$

Combining the coverings, $\{P_{jk}\}_{j,k=1}^{\infty}$ is a covering of $\displaystyle\bigcup_{j=1}^{\infty} A_j$, and so (by $\heartsuit$)

$$\mathscr{L}^m\left(\bigcup_{j=1}^{\infty} A_j\right) \leqslant \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} v(P_{jk}) \leqslant \sum_{j=1}^{\infty}\left(\mathscr{L}^m(A_j) + \frac{\epsilon}{2^j}\right) = \sum_{j=1}^{\infty}\mathscr{L}^m(A_j) \,+\, \epsilon\,.$$

By the Thrilling ϵ-lemma, we're done.

It's worth noting an alternative proof of (ii) above, using instead the Thrilling ϵ-lemma. Given $B \supseteq A$, and given $\epsilon > 0$, we can cover B by $\{P_j\}_{j=1}^{\infty}$ with

$$\sum v(P_j) \leqslant \mathscr{L}^m(B) + \epsilon \qquad (\text{by } \spadesuit)\,.$$

But $\{P_j\}_{j=1}^{\infty}$ also covers A, and so

$$\begin{aligned} & \mathscr{L}^m(A) \leqslant \sum v(P_j) \qquad\qquad (\text{by } \heartsuit) \\ \Longrightarrow \quad & \mathscr{L}^m(A) \leqslant \mathscr{L}^m(B) + \epsilon\,. \end{aligned}$$

Since this is true for every $\epsilon > 0$, the Thrilling ϵ-lemma implies $\mathscr{L}^m(A) \leqslant \mathscr{L}^m(B)$, as desired.

The proof that Lebesgue measure is a measure used nothing about the fact that the covering sets were boxes, or that we assigned the classical volume $v(P)$ to each box P: the exact same proof would have worked with the boxes replaced by any collection $\{P_\alpha\}$ of designated covering sets, and with any assigned nonnegative "volume" $v(P_\alpha)$ for each P_α. Historically, measures constructed in this way were called *Type 1 measures*.[5] What is not automatically clear is whether the constructed measure is interesting.[6] We now address that question for Lebesgue measure.

We want to show that for an m-box P we have $\mathscr{L}^m(P) = v(P)$. The $\leqslant$ is trivial (since P covers P), but for the $\geqslant$ we have to consider all covers, and we have to rule out the possibility that some tricky cover gives a strictly lower estimate. We'll deal with this in a moment, but we first set up some general results, and consider some easier sets.

When estimating Lebesgue measure, the following simple result can reduce the technicalities.

LEMMA 3: $\mathscr{L}^m$ doesn't change if we use closed boxes in the definition. That is, defining

$$\mathscr{L}_c^m(A) = \inf\left\{\sum_{j=1}^{\infty} v(P_j) : A \subseteq \bigcup_{j=1}^{\infty} P_j,\ \text{each } P_j \text{ a } \textbf{closed } m\text{-box}\right\},$$

we have $\mathscr{L}_c^m = \mathscr{L}^m$.

Since open boxes are as small as possible (they include none of the boundary), and closed boxes are as large as possible (they include all of the boundary), it follows that Lebesgue

[5] Yes, there are Type 2 measures as well. Hausdorff measure, which we'll come to later, is such a measure.

[6] In fact such measures are often interesting, and/or result in the same measure. For example, if we use balls B_α in $\mathbb{R}^n$ as covering sets, with $v(B_\alpha)$ the classical volume of the ball, then the resulting measure is exactly Lebesgue measure. We'll give the (very non-trivial) proof of this later.

measure is the same, whether we consider all or any types of boxes, perhaps including only part of the boundary.

The following proposition can also be helpful, and is also of interest in itself.

PROPOSITION 4: Suppose $A \subseteq \mathbb{R}^m$, $x \in \mathbb{R}^m$, $t > 0$. Then

(a) $\mathscr{L}^m(A + x) = \mathscr{L}^m(A)$ where $A + x = \{y + x : y \in A\}$,

(b) $\mathscr{L}^m(tA) = t^m \mathscr{L}^m(A)$ where $tA = \{ty : y \in A\}$.

Given this simple (and expected) behaviour with respect to translations and dilations, it is natural to ask how $\mathscr{L}$ behaves under rotations; if $\mathscr{L}$ does truly calculate m-dimensional volume, then it should indeed also be invariant under rotations. This is true, and we shall prove it later, but it turns out to be significantly harder to prove than Proposition 4: the difficulty is that Lebesgue measure is defined in terms of m-boxes oriented with the coordinate axes, and so a box-covering of a set cannot be rotated to give a box-covering of the rotated set.

After all that, we now calculate the Lebesgue measure of some specific sets. First of all, if $A = \{q\} \subseteq \mathbb{R}$ then $[q, q]$ covers A, and so

$$\mathscr{L}(\{q\}) = 0\,.$$

Then since $\mathbb{Q} = \{q_j\}$ is countable, the countable subadditivity of $\mathscr{L}$ implies

$$\mathscr{L}(\mathbb{Q}) = \mathscr{L}\left(\bigcup_{j=1}^{\infty}\{q_j\}\right) \leqslant \sum_{j=1}^{\infty}\mathscr{L}(\{q_j\}) = \sum_{j=1}^{\infty} 0 = 0\,.$$

The same is true for any countable subset of $\mathbb{R}$, and similarly for $\mathscr{L}^m$ and countable subsets of $\mathbb{R}^m$. In particular

$$\mathscr{L}^m(\mathbb{Q}^m) = 0\,,$$

where $\mathbb{Q}^m$ is the set of points in $\mathbb{R}^m$ with rational coordinates.[7]

[7] As an alternative proof, fix $\epsilon > 0$ and let $\{r_j\}$ be a listing of the points in $\mathbb{Q}^m$. For each r_j, let P_j be an open box containing r_j of volume $\frac{\epsilon}{2^j}$. then $P = \cap P_j$ contains $\mathbb{Q}^m$. By monotonicity and subadditivity $\mathscr{L}^m(\mathbb{Q}^m) \leqslant \mathscr{L}^m \leqslant \epsilon$. So, by the Thrilling ϵ-lemma, $\mathscr{L}^m(\mathbb{Q}^m) = 0$. Note that P has tiny measure, but is open and is a dense subset of $\mathbb{R}^n$.

A set with zero μ-measure (with respect to whatever measure μ) is called a *null set* (or *μ-null set* if there is some possibility of confusion). It is plausible that any $\mathscr{L}^m$-null set would have to be countable, but this is not the case. We now give a famous example, the *Cantor set*.

To construct the Cantor set, we begin with the closed unit interval

$$C_0 = [0,1] = I_{01}\,.$$

Removing the open middle third, we obtain

$$C_1 = \left[0, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, 1\right] = I_{11} \cup I_{12}\,.$$

Removing the middle third of each of these intervals,

$$C_2 = \left[0, \tfrac{1}{9}\right] \cup \left[\tfrac{2}{9}, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, \tfrac{7}{9}\right] \cup \left[\tfrac{8}{9}, 1\right] = \bigcup_{k=1}^{4} I_{2k}\,.$$

We continue the process. At each stage, C_j is the union of 2^j pairwise disjoint closed intervals of length $\frac{1}{3^j}$:

$$C_j = \bigcup_{k=1}^{2^j} I_{jk}\,.$$

Note that

$$I_{jk} \supseteq I_{j+1,2k-1} \cup I_{j+1,2k}\,.$$

Finally, we define

$$C = \bigcap_{j=1}^{\infty} C_j\,.$$

8 **Exercise**. C is uncountable.

As well, C is a null set. To see this, we just note that every C_j has the obvious covering. Together with monotonicity, this gives the upper bound

$$\mathscr{L}(C) \leqslant \mathscr{L}(C_j) \leqslant \sum_{k=1}^{2^j} l(I_{jk}) = \frac{2^j}{3^j}\,.$$

Since this is true for all j, the Thrilling ϵ-lemma implies $\mathscr{L}(C) = 0$, as desired.

Next, we have the key result:

PROPOSITION 5: If $P \subseteq \mathbb{R}^m$ is an m-box then $\mathscr{L}^m(P) = v(P)$. In particular, if $I \subseteq \mathbb{R}$ is an interval then $\mathscr{L}(I) = l(I)$.

We'll prove the $m = 1$ case for closed and bounded - i.e. compact - intervals; the cases of infinite and open intervals then follow easily. The proof for $m > 2$ is similar but much messier: we'll give a different proof later, using the theory of *product measures*.[8] As an (unpleasant) Exercise, 9, one can attempt the $m = 2$ case for $P = [a, b] \times [c, d]$ a closed and bounded box.

PROOF FOR $m = 1$:

The idea of the proof is clear: using more than one interval to cover I will include overlaps, which should only increase the sum of the lengths. It's only a matter of nailing down the argument with as little fussiness as possible. To do this, we assume throughout that I is a compact interval (noting the remarks above).

Since I covers I, we immediately have the upper bound

$$\mathscr{L}(I) \leqslant l(I)\,.$$

To prove the reverse inequality, consider any covering of any compact I by a collection of open intervals $\{I_j\}_{j=1}^{\infty}$. Then we want to show

$$\sum_{j=1}^{\infty} (I_j) \geqslant l(I)\,.$$

Now, since I is compact, in fact some *finite* subcollection $\{I_j\}_{j=1}^{n}$ of the covering intervals also covers I. It is then enough to show that the finite sum of $l(I_j)$ is at least $l(I)$, (since any excluded intervals only make the sum larger). We do this by induction on n. To be precise, the inductive claim is:

P(n): Any covering $\{I_j\}_{j=1}^{n}$ of any compact I by n open intervals satisfies $\displaystyle\sum_{j=1}^{n} l(I_j) > l(I)\,.$

Note a bit of sneakiness in the inductive claim: we're considering *all* compact intervals at the same time. Note also the strict inequality, a byproduct of covering closed intervals by open ones; this strictness does not contradict the reverse inequality, since the $>$ is lost when taking the *inf*.

Base Case: $n = 1$. In this case $I_1 \supseteq I$, and so obviously $l(I) < l(I_1)$. Done.

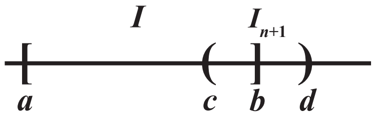
[8] Exactly because of this messiness, higher dimensional Lebesgue is usually introduced later, defined in terms of product measures.

Inductive Step: So we now assume P(n) is true, and we consider a covering $\{I_j\}_{j=1}^{n+1}$ of some compact I by $n+1$ open intervals. Write $I = [a,b]$. Then some I_j contains the point b, and after relabeling we can assume

$$b \in I_{n+1} = (c,d)\,.$$

If $I_{n+1} \supseteq I$ then we're clearly done (just as for the base case). So, let's assume otherwise. Then $c \in I$, and so

$$a \leqslant c < b < d\,.$$



But then the n intervals $I_1, \ldots, I_n$ must cover the interval $[a,c]$ remaining. So, by the inductive hypothesis applied to $[a,c]$,

$$\sum_{j=1}^{n} l(I_j) > c - a\,.$$

Including I_{n+1} back in the sum, we have

$$\sum_{j=1}^{n+1} l(I_j) > (c-a) + (d-c) = d - a > b - a = l(I)\,.$$

10 As a final remark, we note that knowing the Lebesgue measure of intervals allows us to show the existence of non-trivial Cantorlike sets.[9] To be precise, for any ϵ with $0 < \epsilon < 1$, there is a closed set $D \subseteq [0,1]$, with interior $D^\circ = \emptyset$ (i.e. D contains no open intervals),[10] and with $\mathscr{L}(D) = \epsilon$.

---

[9]Some results in the next Handout make things a little easier, but the point is, knowing the measures of intervals suffices to analyse Cantorlike sets.

[10]A set is called *nowhere dense* if its closure has no interior. The Cantor set (which is already closed) is such an example, but $\mathbb{Q}$ for example is not: even though the set of rationals contains no interval, its closure $\overline{\mathbb{Q}} = \mathbb{R}$ includes everything.

SOLUTIONS

2 We want to show that $\mathbb{R}^*$ is metrizable. One method is to simply shoot a very big gun: since $\mathbb{R}^*$ is compact, Hausdorff and second countable, it must be metrizable by Urysohn's metrization theorem (see Handout 0). Of course, that's a fair bit of overkill.

The easiest direct way to prove $\mathbb{R}^*$ is metrizable is to define a bijection between $\mathbb{R}^*$ and a closed interval in $\mathbb{R}$. For example, we can define $f : [-1,1] \to \mathbb{R}^*$ as

$$f(x) = \begin{cases} \dfrac{x}{1-x^2} & -1 < x < 1\,, \\[2ex] \pm\infty & x = \pm 1\,. \end{cases}$$

Clearly f is a bijection, and f is *order-preserving*:

$$a < b \iff f(a) < f(b)\,.$$

Thus the open intervals of $[-1,1]$ and $\mathbb{R}^*$ correspond, and thus automatically the open sets also correspond:

$$U \text{ is open in } [-1,1] \iff f(U) \text{ is open in } \mathbb{R}^*\,.$$

Since (by definition) the metric $d(a,b) = |a-b|$ on $\mathbb{R}$ gives the open sets on $[-1,1]$, it is immediate that $\mathbb{R}^*$ is metrizable by defining $d^*(x,y) = |f^{-1}(x) - f^{-1}(y)|$.

Next, we want to show $a_n \to a$ by our cases definition iff $d^*(a_n, a) \to 0$. First suppose $a \in \mathbb{R}$. Then, for either notion of convergence we have $a_n \in \mathbb{R}$ for n beyond some N. Then

$$\begin{aligned} a_n \to a \quad &\iff \quad f^{-1}(a_n) \to f^{-1}(a) \qquad \text{(continuity of } f \text{ and } f^{-1} \text{ on } \mathbb{R}) \\ &\iff \quad d^*(a_n, a) \to 0\,. \end{aligned}$$

Now suppose $a = \infty$ (with the case $a = -\infty$ handled similarly). Note that for $M \in \mathbb{R}$,

$$a_n > M \implies f^{-1}(a_n) > f^{-1}(M) = \frac{-1 + \sqrt{1 + 4M^2}}{2M}\,.$$

So,

$$\begin{aligned} a_n \to \infty \quad &\iff \quad f^{-1}(a_n) \to 1 \\ &\iff \quad d\left(f^{-1}(a_n), 1\right) \to 0 \\ &\iff \quad d^*(a_n, \infty) \to 0\,. \end{aligned}$$

6 Write

$$\left\{\begin{array}{l} \mathscr{L}^m_{\mathrm{O}}(A) = \inf\left\{\sum_{j=1}^{\infty} v(P_j) : A \subseteq \bigcup_{j=1}^{\infty} P_j,\ \text{each } P_j \text{ an open } m\text{-box}\right\} = \mathscr{L}^m(A) \\ \\ \mathscr{L}^m_{\mathrm{C}}(A) = \inf\left\{\sum_{j=1}^{\infty} v(Q_j) : A \subseteq \bigcup_{j=1}^{\infty} Q_j,\ \text{each } Q_j \text{ a closed } m\text{-box}\right\} \end{array}\right. \qquad A \subseteq \mathbb{R}^m\,.$$

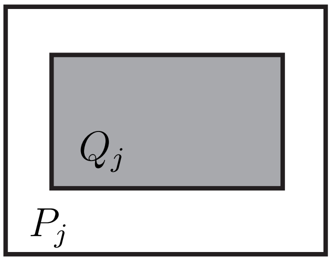
Then we want to show $\mathscr{L}^m_{\mathrm{C}} = \mathscr{L}^m_{\mathrm{O}}$.

Fix A, and consider a covering $\{P_j\}$ of A by open m-boxes. Then $\{\overline{P}_j\}$ is a covering of A by closed m-boxes. Thus

$$\mathscr{L}^m_{\mathrm{C}}(A) \leqslant \sum_{j=1}^{\infty} v(\overline{P}_j) = \sum_{j=1}^{\infty} v(P_j)\,.$$

Taking the *inf* over all coverings of A by open m-boxes, we see $\mathscr{L}^m_{\mathrm{C}}(A) \leqslant \mathscr{L}^m_{\mathrm{O}}(A)$.

For the reverse inequality, fix $\epsilon > 0$ and consider a covering $\{Q_j\}$ of A by closed m-boxes. For each Q_j we can easily find an open m-box $P_j \supseteq Q_j$ and with $v(P_j) \leqslant v(Q_j) + \frac{\epsilon}{2^j}$.



Then $\{P_j\}$ is a covering of A by open m-boxes, and so

$$\mathscr{L}^m_{\mathrm{O}}(A) \leqslant \sum_{j=1}^{\infty} v(P_j) \leqslant \sum_{j=1}^{\infty} v(Q_j) + \epsilon\,.$$

Taking the *inf* over all coverings of A by closed m-boxes, we see $\mathscr{L}^m_{\mathrm{O}}(A) \leqslant \mathscr{L}^m_{\mathrm{C}}(A) + \epsilon$. Thus, $\mathscr{L}^m_{\mathrm{O}}(A) \leqslant \mathscr{L}^m_{\mathrm{C}}(A)$, by the Thrilling ϵ-lemma.

7

(a) We want to show that $\mathscr{L}^m(A+x) = \mathscr{L}^m(A)$ for $A \subseteq \mathbb{R}^m, x \in \mathbb{R}^m$. If $\{P_j\}$ is a covering of A then $\{P_j + x\}$ is a covering of $A + x$, and obviously $v(P_j + x) = v(P_j)$. So

$$\mathscr{L}^m(A) \leqslant \sum_{j=1}^{\infty} v(P_j + x) = \sum_{j=1}^{\infty} v(P_j)\,.$$

Taking the *inf* over all such coverings, we have $\mathscr{L}^m(A+x) \leqslant \mathscr{L}^m(A)$. But replacing x by $-x$ and interchanging A and $A+x$, we also have $\mathscr{L}^m(A) = \mathscr{L}^m((A+x) - x) \leqslant \mathscr{L}^m(A+x)$.

(b) We want to show that $\mathscr{L}^m(tA) = t^m \mathscr{L}^m(A)$ for $A \subseteq \mathbb{R}^m$ and $t > 0$. For an m-box P, we obviously have $v(tP) = t^m v(P)$, and thus the result follows from exactly the same type of argument as in (a).

8 We want to prove that the Cantor set C is uncountable. Supposing not, we can write $C = \{c_n\}_{j=1}^{\infty}$ as a sequence. We now show that there is a $d \in C$ with $d \neq c_j$ for any j. To do this, we first inductively choose $\{d_n\}_{j=1}^{\infty} \subseteq D$ as follows (using the notation from Handout 2):

$$\begin{cases} c_1 \in I_{11} (\text{resp. } I_{12}) \quad \Longrightarrow \quad d_1 = \text{right endpoint of } I_{12} \text{ (resp. } I_{11}) \\ \left.\begin{cases} d_n \in I_{nk} \\ c_{n+1} \in I_{n+1,j} \ j \text{ odd (resp. } j \text{ even)} \end{cases}\right\} \Longrightarrow d_{n+1} = \text{right endpoint of } I_{n+1,2k} \text{ (resp } I_{n+1,2k-1}) \end{cases}$$

The point of this is that if $d_n \in I_{nk}$ then also $d_m \in I_{nk}$ for $m \geqslant n$, and $c_n \notin I_{nk}$. Since I_{nk} has length $\frac{1}{3^n} \to 0$, and since C is closed, this shows $d_n \to d$ for some $d \in C$. And, from the above observation for each c_n, clearly $c_n \neq d$.

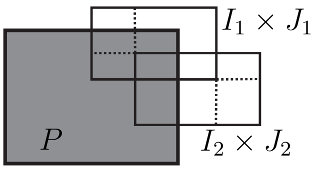
9 Given $P = [a,b] \times [c,d]$, we want to show $\mathscr{L}^2(P) = (b-a)(d-c) = v(P)$. Consider a covering $\{P_j\}$ of P by a sequence of open rectangles $P_j = I_j \times J_j$. By compactness of P, some finite subcollection $\{P_j\}_{j=1}^N$ covers P, and then we want to show

$$\sum_{j=1}^{N} v(P_j) \geqslant (b-a) \cdot (d-c)\,. \tag{$*$}$$

We can also assume each $P \cap P_j \neq \emptyset$.

Consider the corresponding closed covering $\{\overline{P}_j\}$, where $\overline{P}_j = \bar{I}_j \times \bar{J}_j$. Chopping the $\bar{I}_j$ and $\bar{J}_j$ into comparable pieces, we can assume

$$\text{For each } i \text{ and } j, \text{ either } \bar{I}_i = \bar{I}_j \text{ or } I_i \cap I_j = \emptyset, \text{ and similarly for } J_i, J_j. \tag{$\dagger$}$$



Throwing away any unnecessary intervals, we can also clearly assume each $P \cap P_k \neq \emptyset$. But then if $\bar{I}_i$ and $\bar{J}_j$ are any two such intervals, then $\bar{I}_i \times \bar{J}_j = \overline{P}_k$ for some k (since ($\dagger$) implies this is the only way to cover $P \cap (I_i \times J_j)$). Relabelling, we now have intervals $\{\bar{I}_i\}_{i=1}^L$ and $\{\bar{J}_j\}_{i=1}^M$ such that $\{\overline{P}_k\}_{k=1}^N \supseteq \{\bar{I}_i \times \bar{J}_j\}_{i,j}$, and where:

$$\begin{cases} (a,b) \subseteq \displaystyle\bigcup_{i=1}^{L} \bar{I}_i \\[2ex] (c,d) \subseteq \displaystyle\bigcup_{j=1}^{M} \bar{J}_j\,. \end{cases}$$

Then, by Lemma 3 and the $m = 1$ case of Proposition 5,

$$\begin{aligned} \sum_{k=1}^{N} v(P_j) &\geqslant \sum_{i=1}^{L} \sum_{j=1}^{M} v(\bar{I}_i \times \bar{J}_j) \\ &= \sum_{i=1}^{L} \sum_{j=1}^{M} l(\bar{I}_i) \cdot l(\bar{J}_j) = \left(\sum_{i=1}^{L} l(\bar{I}_i) \right) \cdot \left(\sum_{j=1}^{M} l(\bar{J}_j) \right) \geqslant (b-a) \cdot (d-c)\,. \end{aligned}$$