A Measure theory and integration

In these notes we give an alternative approach to measure theory and integration, emphasising the duality between a measure (a way of measuring the size of subsets of a space) and an integral (a rule for integrating some class of functions on the space).

A.1 Motivation from the Lebesgue integral

Let's think about Lebesgue measure for a moment, and try to abstract out what "integration" means.

Certainly, we can integrate any continuous compactly supported function $f \in C_c(\mathbb{R}^n)$, and this is a linear functional, which we write as

$$\int : C_c(\mathbb{R}^n) \to \mathbb{R}$$
$$f \mapsto \int_{\mathbb{R}^n} f(x)dx.$$

We can also integrate many other functions, but for now let's just concentrate on these — it will eventually turn out that everything else is determined by how the integral behaves here.

Remark. Is this a *bounded linear functional*?

Unfortunately no, as $\mathbb{R}^n$ is too big. It's easy to see that if we restrict to some finite volume $A \subset \mathbb{R}^n$, then $\int : C_c(A) \to \mathbb{R}$ is a bounded linear functional, as

$$\int_A f(x)dx \le (\operatorname{vol} A)(\sup_{x \in A} f(x))$$
$$= (\operatorname{vol} A)\|f\|_\infty.$$

In fact, the operator norm $\int$ is $\left\|\int\right\| = \operatorname{vol} A$.

What other properties does integration have? The obvious one is that it is positive: if $f(x) \ge 0$ everywhere (or even almost everywhere) then $\int_A f(x)dx \ge 0$. It turns out that these properties are all that 'essentially' matters about Lebesgue integration, so let's pull them out as a definition.

Definition A.1. A *Radon integral* on a 'nice' space X is a positive linear functional $C_c(X) \to \mathbb{R}$.

What should 'nice' mean? We have a few options. One can set up the theory at the most general level, allowing any *locally compact Hausdorff space X*. To make life easier in what follows, we'll also freely assume second countability (so that we don't need to futz about with nets, and

1

can just use sequences). In fact, in places we'll even just restrict to the case of a metric space (or even a separable metric space). If you need a reminder about these adjectives, see the appendix on point set topology below.

Our goal for the next while will be to prove a theorem that looks something like this:

Theorem A.2. *Let X be a nice space. There is a bijective correspondence between*

1. *Radon integrals on X, and*
2. *Radon measures on the Borel σ-algebra of X.*

This is the theorem that says that we can take either of two points of view, taking as the fundamental mathematical objects:

1. the integrals of compactly supported continuous functions, or
2. the measures of Borel sets.

(Of course measures of Borel sets are the same thing as integrals of characteristic functions of Borel sets.)

A.2 Basic measure theory

First we need to develop some of the basics of measure theory. I think Terry Tao's book "An introduction to measure theory" is an excellent source for this; this section is essentially a selection from §1.4 of that book.

A.2.1 Algebras of sets

Let X be some set (typically a topological space or metric space, but for now we don't need any structure at all).

Definition A.3. A Boolean algebra (of subsets of X) is a collection $\mathcal{M}$ of subsets of X such that

- $\emptyset \in \mathcal{M}$,
- if $E \in \mathcal{M}$, then $X \setminus E \in \mathcal{M}$, and
- if $E, F \in \mathcal{M}$, then $E \cup F \in \mathcal{M}$.

When it is completely clear that we are talking about an algebra of sets (and not an algebra in the sense of an associative algebra or a Lie algebra, for example), we may drop the word 'Boolean'.

One easily sees that $E \cap F$, $E \setminus F$, $F \setminus E$, and the symmetric difference $E \triangle F$ are all in $\mathcal{M}$ as long as E and F are, by combining the last two axioms. Moreover, by induction, any finite union (or finite intersection) of sets in $\mathcal{M}$ is in $\mathcal{M}$.

We can talk about a Boolean algebra $\mathcal{M}$ being *finer* than a Boolean algebra $\mathcal{M}'$ — this just means $\mathcal{M}' \subset \mathcal{M}$. One should think about this as the finer algebra 'making finer distinctions'. In the other direction, if $\mathcal{M}$ is finer that $\mathcal{M}'$ we say $\mathcal{M}'$ is coarser than $\mathcal{M}$.

In probability, one often thinks of the elements of the set X as 'states' of an underlying system (either at a moment, or representing an entire history), and an algebra of subsets of X describes the possible ways to partition up these states according to some observations we might make. When we consider further observations (perhaps just by 'time passing') we should expect to pass to a finer algebra.

Certainly on any set there is a finest algebra (the discrete algebra $\mathcal{M} = 2^X$, also known as the power set of X) and a coarsest algebra (just $\mathcal{M} = \{\emptyset, X\}$).

The intersection of two Boolean algebras is always a Boolean algebra (coarser than either of them), and in fact the intersection of an arbitrary family of Boolean algebras is again a Boolean algebra.

Definition A.4. Given an arbitrary collection $\mathcal{F}$ of subsets of X, we can form the Boolean algebra generated $\mathcal{F}$, which we denote $\langle \mathcal{F} \rangle_{\text{bool}}$, as the intersection of all Boolean algebras containing $\mathcal{F}$.

(This is an intersection of a nonempty collection, since the power set contains any collection of subsets.) One can see that the Boolean algebra generated by $\mathcal{F}$ is the unique coarsest Boolean algebra containing $\mathcal{F}$, and often this characterisation is useful.

Definition A.5. A σ-algebra is a Boolean algebra which moreover is closed under taking countable unions of sets in the algebra.

While we really want to be able to talk about σ-algebras, unfortunately they are much harder to deal with that Boolean algebras. While all the discussion above (finer, coarser, generation) passes over immediately to the corresponding notions for σ-algebras, we find that σ-algebras are somewhat slippery:

Exercise. Given a collection of subsets $\mathcal{F}$ of X, define a family $\{\mathcal{F}_n\}_{n \geq 0}$ inductively by:
- $\mathcal{F}_0 = \mathcal{F}$
- $\mathcal{F}_{n+1} = \{$finite unions of set in $\mathcal{F}_n$, or the complement of such a finite union$\}$

and similarly define $\{\mathcal{F}_n^\sigma\}$ replacing 'finite' everywhere with 'countable'.

1. Show that the Boolean algebra generated by $\mathcal{F}$ is exactly $\bigcup_{n \geq 0} \mathcal{F}_n$. (Hint: it is clear that every set in $\bigcup_{n \geq 0} \mathcal{F}_n$ lies in $\langle \mathcal{F} \rangle_{\text{bool}}$. After that, you just need to show that it really is a Boolean algebra. 2nd hint: given $E, F \in \bigcup_{n \geq 0} \mathcal{F}_n$, we may say $E \in \mathcal{F}_{n_1}$ and $F \in \mathcal{F}_{n_2}$ for some $n_1, n_2 \geq 0$.)

2. Explain why this argument fails when we try to show that the σ-algebra generated by $\mathcal{F}$ is exactly $\bigcup_{n\geq 0} \mathcal{F}_n^\sigma$. (Hint: what happens if we try to take a countable union of sets E_n, where $E_n \in \mathcal{F}_n^\sigma$?)

3. Can you produce a family $\mathcal{F}$ where in fact $\langle \mathcal{F} \rangle_\sigma \neq \bigcup_{n\geq 0} \mathcal{F}_n^\sigma$?

4. Learn about transfinite induction, and see if you're satisfied by the resulting 'explicit' description of $\langle \mathcal{F} \rangle_\sigma$.

How then do we work with σ-algebras? The following lemma provides an 'induction principle' for σ-algebras.

Lemma A.6. *Suppose $P(E)$ is a property of subsets $E \subset X$ such that*
- *$P(\emptyset)$ is true,*
- *if $P(E)$ is true, then $P(X \setminus E)$ is true, and*
- *if $\{E_i\}_{i\geq 0}$ are subsets of X, and $P(E_i)$ is true for all i, then $P\left(\bigcup_{i\geq 0} E_i\right)$ is true also.*

Now suppose that $P(E)$ is true for all $E \in \mathcal{F}$, where $\mathcal{F}$ is some family of subsets of X. Then $P(E)$ is true for all $E \in \langle \mathcal{F} \rangle_\sigma$.

Proof: The collection $\{E \mid P(E)\}$ is a σ-algebra containing $\mathcal{F}$, so $\langle \mathcal{F} \rangle_\sigma \subset \{E \mid P(E)\}$. □

Before we continue on to measures, we have to discuss the most important examples of σ-algebras.

Definition A.7. If X is a topological space, the *Borel σ-algebra* is the σ-algebra generated by the open sets of X.

We'll write this as $\mathcal{B}[X]$ — the square brackets rather than parentheses are just a local notation to make sure we don't mistake the Borel σ-algebra for the set of bounded linear transformations on something.

To make sure you understand what we've been doing, make sure you try these:

Exercise.
1. Equivalently, the Borel σ-algebra is generated by the closed sets of X.

2. If X is a separable metric space, the Borel σ-algebra is generated by the open balls of X. (Is there any hope of dropping the hypothesis 'separable' here?)

3. If you enjoy point set topology, repeat the previous exercise merely assuming X is a second countable, locally compact, Hausdorff space. (Can you drop any of those assumptions?)

4. When $X = \mathbb{R}^n$, the Borel σ-algebra is generated by boxes. (Or by boxes with rational coordinates, or by boxes with dyadic coordinates.)

5. What condition on a topological space X is needed so that the σ-algebra generated by compact sets agrees with the Borel σ-algebra?

Note that to prove $\langle \mathcal{F} \rangle_\sigma = \langle \mathcal{G} \rangle_\sigma$, it suffices to show you can write every set in $\mathcal{F}$ using a finite sequence of the operations of 'take a set from $\mathcal{G}$', complementation, and countable union of sets previously constructed, and vice versa switching $\mathcal{F}$ and $\mathcal{G}$.

Exercise. If E and F are Borel in $\mathbb{R}^a$ and $\mathbb{R}^b$ respectively, show that $E \times F$ is Borel in $\mathbb{R}^{a+b}$. (Hint: this is not trivial; assume F is a box, first.)

Exercise. Suppose E is Borel in $\mathbb{R}^{a+b}$. Show that the slice $E^{x_1} = \{x_2 \in \mathbb{R}^b \mid (x_1, x_2) \in E\}$ is Borel for each $x_1 \in \mathbb{R}^a$.

Now show that this statement is not true if we replace both instances of 'Borel' with 'Lebesgue'. (Hint, take a set C that isn't Lebesgue in $\mathbb{R}$ (ha!), and consider $C \times \{0\}$. Why is this Lebesgue?)

(You might interpret this exercise as telling you that Borel sets are better than Lebesgue sets.)

A.2.2 The definition of a measure

Definition A.8. If $\mathcal{M}$ is a Boolean algebra of subsets of X, a finitely additive measure for $\mathcal{M}$ is a function $\mu : \mathcal{M} \to [0, \infty]$ such that

- $\mu(\emptyset) = 0$
- if $E, F \in \mathcal{M}$ are disjoint, then $\mu(E \cup F) = \mu(E) + \mu(F)$.

Lemma A.9.

- $E \subset F$ implies $\mu(E) \leq \mu(F)$
- if $E_1, \ldots, E_n$ are in $\mathcal{M}$, and are disjoint, $\mu\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n \mu(E_i)$.
- without the disjointness assumption, $\mu\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n \mu(E_i)$.
- $\mu(E \cup F) + \mu(E \cap F) = \mu(E) + \mu(F)$

Exercise. Explain why one shouldn't write the last fact as $\mu(E \cup F) = \mu(E) + \mu(F) - \mu(E \cap F)$.

Definition A.10. If $\mathcal{M}$ is a σ-algebra on X, a measure for $\mathcal{M}$ is a finitely additive measure for $\mathcal{M}$ which further satisfies the condition

- If $\{E_i\}_{i \geq 0}$ is a countable family of disjoint sets, each in $\mathcal{M}$, then

$$\mu\left(\bigcup_{i \geq 0} E_i\right) = \sum_{i \geq 0} \mu(E_i).$$

The triple $(X, \mathcal{M}, \mu)$ is called a *measure space*. Sometimes the pair $(X, \mathcal{M})$ is called a *measurable space*.

Exercise.

- $\mu\left(\bigcup_{i\geq 0} E_i\right) \leq \sum_{i\geq 0}\mu(E_i)$
- If $E_1 \subset E_2 \subset \cdots$ are all in $\mathcal{M}$,

$$\mu\left(\bigcup_{i\geq 0} E_i\right) = \lim_i \mu(E_i) = \sup_i \mu(E_i).$$

- If $E_1 \supset E_2 \supset \cdots$ are all in $\mathcal{M}$, and eventually $\mu(E_i)$ is finite

$$\mu\left(\bigcap_{i\geq 0} E_i\right) = \lim_i \mu(E_i) = \inf_i \mu(E_i).$$

- Explain why the finiteness condition is required in the previous exercise.

Definition A.11. A *Borel measure* is just a measure defined on the σ-algebra of Borel sets. (Or sometimes, on a bigger σ-algebra containing all the Borel sets.)

A.2.3 Completeness

Definition A.12. A measure is *complete* if every subset of a measure zero set is measurable (and hence also has measure zero).

Sometimes completeness is a desirable property.

Definition A.13. Given a measure μ on a σ-algebra $\mathcal{M}$, we can *complete* the measure by defining a new σ-algebra $\mathcal{M}'$ whose sets are of the form $E \subset X$, such that there exists $E' \in \mathcal{M}$ so that $E\Delta E'$ is contained in a null set with respect to μ. We define the measure $\mu' : \mathcal{M}' \to [0,\infty]$ simply by $\mu'(E) = \mu(E')$.

Exercise.
- Check that $\mathcal{M}'$ as defined really is a σ-algebra.
- Check that μ' is well-defined (that is, it does not depend on the choice of E' for each E — clearly there are many different choices).
- Check that μ' is still countably additive.

Exercise. Observe that in the construction of Lebesgue measure in Analysis 2, we effectively built Lebesgue measure on the Borel σ-algebra, and then completed it.

A.2.4 Radon measures

Definition A.14. A Borel measure is *inner regular* if

$$\mu(A) = \sup\{\mu(K) \mid K \text{ is a compact subset of } A\}.$$

Definition A.15. A *Radon measure* is a Borel measure which
  - is inner regular, and
  - $\mu(K) < \infty$ for every compact set K.

Exercise. Lebesgue measure on $\mathbb{R}^n$ is inner regular, and so easily a Radon measure.

A.3 Extending integrals to all integrable functions

Throughout this section X is a locally compact Hausdorff space. We will assume X is second countable, although with some extra work this condition can be omitted.

This section closely follows the presentation in §6.1 of Pedersen's *Analysis Now*; I've tried to unpack it a bit and make it a little friendlier.

We now set out to extend a Radon integral, from being merely defined on compactly supported continuous functions to a much bigger class of functions. This has two purposes:
  - to include the characteristic functions of 'many' sets in the class of integrals we can integrate, to extract a measure from the integral
  - to provide a setting where we can prove the monotone and dominated convergence theorems.

A.3.1 Monotone limits of $C_c(X)$

To begin, we define two new classes of functions,

$$C_c(X)^m = \{f : X \to \mathbb{R} \cup \{+\infty\} \mid \exists f_n \in C_c(X) \text{ so } f_n \nearrow f\}$$
$$C_c(X)_m = \{f : X \to \mathbb{R} \cup \{-\infty\} \mid \exists f_n \in C_c(X) \text{ so } f_n \searrow f\}$$

where $f_n \nearrow f$ expresses that f_n is an increasing sequence, converging pointwise to f, and similarly for $f_n \searrow f$.

We're going to spend some time understanding these classes.

Example A.16. When $X = \mathbb{R}$, $\chi_{(a,b)}$ is in $C_c(X)^m$, but not in $C_c(X)_m$. Conversely, $\chi_{[a,b]}$ is in $C_c(X)_m$, but not in $C_c(X)^m$.

Some notation: if $\mathcal{F}$ is some collection of functions $X \to \mathbb{R}$, we write $\mathcal{F}_+$ for the subset of non-negative functions.

Lemma A.17.
- *Both $C_c(X)^m$ and $C_c(X)_m$ are closed under addition, and multiplication by positive scalars.*
- *When $f \in C_c(X)^m$, we have $-f \in C_c(X)_m$ and vice versa.*

Lemma A.18.
- *The function -1 is in $C_c(X)^m$ iff X is compact.*
- *The function 1 is in $C_c(X)^m$ iff X is a countable union of compact sets.*

Proof: The first fact is easy: we have some compactly supported function less than -1, so outside of a compact set $0 \leq -1$, so the whole space must be compact.

Taking $K_n = \operatorname{supp} f_n$, for some sequence $\{f_n\} \subset C_c(X)$ such that $f_n \nearrow f$, we see every point if eventually in some K_n, i.e. $X = \bigcup K_n$. In the other direction, if $X = \bigcup K_n$, for each n choose some $f_n \in C_c(X)$ which is identically 1 on K_n, and then take maximums as $g_n = \max(f_1, \ldots, f_n)$. We see that $g_n \in C_c(X)$ still, and $g_n \nearrow 1$. $\qquad\square$

Lemma A.19. *If X is a locally compact metric space, $\chi_K \in C_c(X)_m$ for every compact set K.*

Proof: Define
$$
f_n(x) = \begin{cases} 1 & \text{if } x \in K \\ 0 & \text{if } d(x, K) \geq \frac{1}{n} \\ 1 - nd(x, K) & \text{otherwise.} \end{cases}
$$
Observe that $d(x, K) = \inf\{d(x, y) \mid y \in K\} = \max\{d(x, y) \mid y \in K\}$ since K is compact, and use this to show $d(x, K)$ is continuous, and hence that f_n is continuous.

One has to be quite careful checking that eventually f_n is compactly supported! The problem is that even though K is compact, the set $\{x \mid d(x, K) \leq \frac{1}{n}\}$ could be noncompact. We need to use the hypothesis that X is locally compact, and Lemma A.51 from the appendix, which ensures this set is eventually compact. $\qquad\square$

One can also prove the following facts, although we won't need them here:

Lemma A.20.
- *If $f, g \in C_c(X)^m_+$, then $fg \in C_c(X)^m_+$.*
- *Every non-negative lower semicontinous function is in $C_c(X)^m$.*
- *$C_c(X)^m \cap C_c(X)_m = C_c(X)$.*

A.3.2 Upper and lower integrals

We next define linear functionals

$$\int^* : C_c(X)^m \to \mathbb{R} \cup \{+\infty\}$$

$$: f \mapsto \sup \left\{ \int g \mid f \geq g \in C_c(X) \right\}$$

$$\int_* : C_c(X)_m \to \mathbb{R} \cup \{-\infty\}$$

$$: f \mapsto \inf \left\{ \int g \mid f \leq g \in C_c(X) \right\}$$

Remember that $C_c(X)^m$ and $C_c(X)_m$ are cones, not vector spaces, so when we say linear functional we mean functions that commute with multiplication by non-negative scalars and with addition. It's not actually obvious that these *are* linear; we address this in Lemma A.24.

Observe that of course if $f \in C_c(X)$, then $\int^* f = \int_* f = \int f$. Moreover if $f \in C_c(X)^m$

$$\int^* f = -\int_* (-f). \tag{A.1}$$

We then define the class of *integrable functions* $L^1(X)$ (of course relative to the particular Radon measure, which we consider fixed throughout this section) to be those functions $f : \mathbb{R} \to \mathbb{R} \cup \pm\infty$ such that for every $\epsilon > 0$, we can find a larger function $g \in C_c(X)^m$ and a smaller function $h \in C_c(X)_m$, such that $\int^* g - \int_* h < \epsilon$.

Another way to express this same idea is to define upper and lower integrals for arbitrary functions $\mathbb{R} \to \mathbb{R} \cup \pm\infty$ by

$$\int^{**} f = \inf \left\{ \int^* g \mid f \leq g \in C_c(X)^m \right\}$$

$$\int_{**} f = \sup \left\{ \int_* g \mid f \geq g \in C_c(X)_m \right\}$$

and then to say that $f \in L^1(X)$ if and only if $\int^{**} f = \int_{**} f$ and this number is in $\mathbb{R}$ (that is, not $\pm\infty$. We define $\int : L^1(X) \to \mathbb{R} \cup \{\pm\infty\}$ by $\int f = \int^{**} f = \int_{**} f$. It's only a little work to show $C_c(X) \subset L^1(X)$, and our newly defined $\int$ agrees with the original one where it should. Moreover $\int$ on $L^1(X)$ is a positive linear functional.

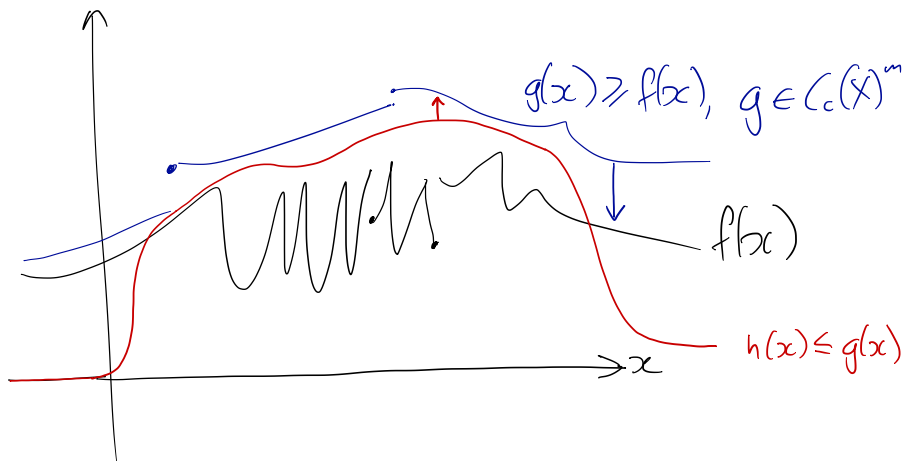*Exercise.* Verify that these definitions of $L^1(X)$ really are the same!

Figure 1: We calculate $\int^{**} f$ as an infimum over slightly larger functions g in the class $C_c(X)^m$ of the supremum over compactly supported continuous functions $h \leq g$ of $\int h$.

Lemma A.21. *A function $f \in C_c(X)^m$ is integrable if and only if $\int^* f < \infty$, in which case $\int f = \int^* f$. Similarly a function $f \in C_c(X)_m$ is integrable if and only if $\int_* f > -\infty$.*

Proof: If $\int^* f < \infty$, for any $\epsilon > 0$ we can find some $h \in C_c(X)$ so $h \leq f$ and $\int^* f - \int h < \epsilon$. (That is, in the definition of being integrable we observe we can take $g = f$, and $h \in C_c(X)$.) □

In not too long we'll discover that the characteristic functions of compact Borel sets are automatically in $L^1(X)$, and that we can define a Borel measure from a Radon integral using the integrals of these characteristic functions. Indeed, we'll also see that $L^1(X)$ as we've just defined it agrees with the $L^1(X)$ that one might more conventionally define starting from that measure.

First, however, we need to prove a number of technical statements about $C_c(X)^m$, $C_c(X)_m$, the upper and lower integrals, and the class $L^1(X)$. In the next section, we'll establish the basic convergence theorems (the monotone convergence theorem, Fatou's lemma, and the dominated convergence theorem).

Lemma A.22. *Given an increasing sequence $\{f_n\} \subset C_c(X)^m$, $\sup\{f_n\}$ is in $C_c(X)^m$ too.*

Proof: Choose $\{g_{nm}\} \subset C_c(X)$ with $g_{nm} \nearrow f_n$. By replacing each g_{nm} with $g'_{nm} = \max(g_{1m}, g_{2m}, \ldots g_{nm})$ (notice we still have $g'_{nm} \in C_c(C)$, and $g'_{nm} \nearrow f_n$) we may assume that g_{nm} is increasing with respect to n as well. We claim $g_{nn} \nearrow \sup\{f_n\}$. □

Lemma A.23. *If $\{f_n\} \subset C_c(X)$ and $f_n \nearrow f$ for some $f \in C_c(X)^m$, then $\int f_n \nearrow \int^* f$.*

We defer the proof to Appendix A.8.

Lemma A.24. *The linear functionals $\int^*$ and $\int_*$ are actually linear!*

Proof: Homogeneity, that is, the condition that $\int^* cf = c \int^* f$ for all $c \geq 0$, is straightforward.

Unfortunately additivity is a bit tricky. It's clear that $\int^* (f + g) \geq \int^* f + \int^* g$, as if we have functions $f', g' \in C_c(X)$ whose integrals are very close to $\int^* f$ and $\int^* g$ respectively, we can use $f' + g'$ to produce a lower bound for $\int^* (f + g)$.

To go the other way we need Lemma A.23: pick sequences $\{f_n\}, \{g_n\} \subset C_c(X)$ so $f_n \nearrow f$ and $g_n \nearrow g$. One see that $f_n + g_n \nearrow f + g$. Then $\int f_n \nearrow \int^* f$, $\int g_n \nearrow \int^* g$, $\int (f_n + g_n) \nearrow \int^* (f + g)$, and the linearity of $\int$ on $C_c(X)$ itself allows us to conclude that $\int (f_n + g_n) = \int f_n + \int g_n \nearrow \int^* f + \int^* g$. Uniqueness of limits in $\mathbb{R}$ gives the result. □

That's a relief.

Corollary A.25. *On $L^1(X)$, $\int$ is linear.*

Lemma A.26. *If $\{f_n\} \subset C_c(X)^m$ and $f_n \nearrow f$, by Lemma A.22 $f \in C_c(X)^m$ and we have*

$$\int^* f_n \nearrow \int^* f.$$

Again we defer the proof to Appendix A.8.

Lemma A.27. *If $f \in C_c(X)_m$ and $g \in C_c(X)^m$, and $f \leq g$, then $\int_* f \leq \int^* g$.*

Proof: Since $g - f \geq 0$,

$$
\begin{aligned}
0 &\leq \int^* g - f \\
&= \int^* g + \int^* (-f) \\
&= \int^* g - \int_* f.
\end{aligned}
$$

□

Corollary A.28. *For any f, $\int_{**} f \leq \int^{**} f$, since for any g in the set $\int_{**}$ is a supremum over, and any g' in the set $\int^{**}$ is an infimum over, we have $g \leq f \leq g'$, so the previous lemma applies.*

Lemma A.29. *The class $L^1(X)$ forms a vector space.*

Proof: This is pretty easy using the definition and the fact that $\int^*$ and $\int_*$ are linear (Lemma A.24). □

Lemma A.30. *The class $L^1(X)$ is closed under* max *and* min.

Proof: If $f_i \in L^1(X)$ for $i = 1, 2$, we can find $\overline{f_i} \in C_c(X)^m$ and $\underline{f_i} \in C_c(X)_m$ so

$$\int^* \overline{f_i} - \int_* \underline{f_i} < \epsilon/2.$$

Now

$$\max(\overline{f_1}, \overline{f_2}) - \max(\underline{f_1}, \underline{f_2}) \leq (\overline{f_1} - \underline{f_1}) + (\overline{f_2} - \underline{f_2})$$

(prove this by checking all four cases in the left hand side), and both sides of this inequality are in $C_c(X)^m$. Thus we have

$$\int^* \max(\overline{f_1}, \overline{f_2}) - \max(\underline{f_1}, \underline{f_2}) \leq \int^* (\overline{f_1} - \underline{f_1}) + (\overline{f_2} - \underline{f_2})$$

and using linearity and Equation (A.1)

$$\int^* \max(\overline{f_1}, \overline{f_2}) - \int_* \max(\underline{f_1}, \underline{f_2}) \leq \int^* \overline{f_1} + \int^* \overline{f_2} - \int_* \underline{f_1} - \int_* \underline{f_1}$$

$$< \epsilon$$

and so the pair of functions $\max(\overline{f_1}, \overline{f_2})$ and $\max(\underline{f_1}, \underline{f_2})$ provide the necessary witnesses to show $\max(f_1, f_2) \in L^1(X)$. A similar argument handles the minimum. $\qquad\square$

Theorem A.31 (The monotone convergence theorem)**.** *If $\{f_n\} \subset L^1(X)$ and $f_n \nearrow f$, and moreover* $\sup \int f_n < \infty$, *then $f \in L^1(X)$ and $\int f = \lim \int f_n$.*

Proof: Since $L^1(X)$ is a vector space, by subtracting off f_0 from f and f_n we may assume that $f_0 = 0$. Define $g_n = f_{n+1} - f_n \in L^1(X)$. By definition, for any $\epsilon > 0$ we may pick $h_n \in C_c(X)^m$ with $h_n \geq g_n$, but still $\int^* h_n < \left(\int g_n\right) + 2^{-n}\epsilon$. With $h = \sum h_n$, we can apply Lemma A.26 to see $h \in C_c(X)^m$ and $h \geq \sum_{i=0}^{n-1} g_i = f_n$, so $h \geq f$. Now we calculate

$$\int f_m = \int_{**} f_m$$

$$\leq \int_{**} f$$

$$\leq \int^{**} f \qquad\qquad\qquad \text{(by Corollary A.28)}$$

$$\leq \int^{**} h$$

$$= \int^* h$$

$$= \lim_N \int^* \sum_{n=0}^{N} h_n \qquad \text{(by Lemma A.26)}$$

$$= \sum \int^* h_n \qquad \text{(by linearity of } \int^*, \text{ Lemma A.24)}$$

$$\leq \sum \int g_n + 2^{-n}\epsilon \qquad \text{(by the discussion above)}$$

$$= \lim \int f_n + \epsilon \qquad \text{(making use of Corollary A.25)}$$

This hold for every m, so taking $m \to \infty$ we obtain $\lim \int f_m \leq \int_{**} f \leq \int^{**} f \leq \lim \int f_n$, showing that f all these numbers are equal and f is integrable. □

Theorem A.32 (The dominated convergence theorem). *If $\{f_n\} \subset L^1(X)$, and $f_n \to f$ pointwise (not necessarily monotonically!), and if $|f_n| \leq g$ for some $g \in L^1(X)$, then $f \in L^1(X)$ and $\int f = \lim \int f_n$.*

Proof: Define $h_n = \inf(f_n, f_{n+1}, \ldots)$. We see $h_n \nearrow f$, and using the bound $|f_n| \leq g$, we have $\int h_n \leq \int f_n \leq \int |f_n| \leq \int g$, so $\sup \int h_n < \infty$ and we can apply the monotone convergence theorem to show $f \in L^1(X)$.

Next we look at the sequence $a_n = 2g - |f - f_n|$. Again, let $b_n = \inf(a_n, a_{n+1}, \ldots)$, so $b_n \nearrow 2g$ and since $0 \leq a_n \leq 2g$ and $0 \leq b_n \leq 2g$, we have $\sup \int b_n < \infty$. We can thus apply the monotone convergence theorem a second time, obtaining

$$2 \int g = \lim \int b_n \leq \lim \int a_n = 2 \int g - \lim \int |f - f_n|$$

This gives $\lim \int |f - f_n| \leq 0$, and the positivity of the integral gives us $\lim \int |f - f_n| = 0$. As $\left| \int f - \int f_n \right| \leq \int |f - f_n|$, we obtain the result. □

I've intentionally left out a proof of Fatou's lemma here (essentially rolling the idea into the proof of the dominated convergence theorem) for the sake of an assignment question. (When you do that assignment question, I'm happy for you to prove Fatou's lemma in the context of measures or integrals. There's a proof of a weak version of Fatou's lemma in the measure theory notes.)

A.4 Radon integrals from measures

We let $\mathcal{F}$ denote the Borel measurable functions on a space X, that is, those functions $f : X \to \mathbb{R}$ such that $f^{-1}(t, \infty)$ is Borel for every $t \in \mathbb{R}$. As usual $\mathcal{F}_+$ denotes the cone of non-negative elements in $\mathcal{F}$.

Theorem A.33. *Suppose μ is a Borel measure on X. There is a unique positive linear functional*

$$\Phi : \mathcal{F}_+ \to [0, \infty]$$

such that

- *$\Phi(\chi_A) = \mu(A)$ for every Borel set A, and*
- *$\Phi(f_n) \nearrow \Phi(f)$ whenever $f_n \nearrow f$ in $\mathcal{F}_+$.*

Proof: Let $\mathcal{F}_s$ denote the simple functions in $\mathcal{F}_+$,

$$\left\{ \sum \alpha_n \chi_{A_n} \mid \alpha_n \geq 0, A_n \in \mathcal{B} \right\}.$$

We first define Φ in $\mathcal{F}_s$ by $\Phi(\alpha_n \chi_{A_n}) = \sum \alpha_n \mu(A_n)$. It's pretty straightforward (mucking about with linear combinations of characteristic functions) to show this is a positive linear functional on $\mathcal{F}_s$.

We now establish a few preparatory results, whose proofs we postpone for a moment.

Lemma A.34. *If $f_n \nearrow f$ in $\mathcal{F}_s$, then $\Phi(f_n) \nearrow \Phi(f)$.*

Lemma A.35. *If $\{f_n\}, \{g_n\} \subset \mathcal{F}_s$, and $f_n \nearrow f$, $g_n \nearrow f$ for some $f \in \mathcal{F}_+$, then $\lim \Phi(f_n) = \lim \Phi(g_n)$.*

Lemma A.36. *If $f \in \mathcal{F}_+$, then there exists a sequence $\{f_n\} \subset \mathcal{F}_s$ so $f_n \nearrow f$.*

Using these facts, we define Φ on all of $\mathcal{F}_+$ by $\Phi(f) = \lim \Phi(f_n)$, using some approximating sequence provided by Lemma A.36. Using the uniqueness result from Lemma A.35 is is straightforward to show that Φ is a positive linear functional. Certainly $\Phi(\chi_A) = \mu(A)$, because we can approximate χ_A with a constant sequence.

We still need to show that $\Phi(f_n) \nearrow \Phi(f)$ whenever $f_n \nearrow f$ in $\mathcal{F}_+$. Pick approximating sequences $\{g_{nm}\} \subset \mathcal{F}_s$, so $g_{nm} \nearrow f_n$. We'd like to construct from this two parameter family some functions $h_n \in \mathcal{F}_s$, with that property that $h_n \nearrow f$, and still $h_n \leq f_n$. If we can do this, we see $\lim \Phi(f_n) \leq \Phi(f) = \lim \Phi(h_n) \leq \lim \Phi(f_n)$, giving the result. You might hope to use a diagonal

argument, and try $h_n = g_{nn}$, but the problem here is that there's no guarantee[1] that $g_{nn} \nearrow f$. Unfortunately while g_{nm} is increasing in m, it is not necessarily increasing with n. To get around this, we take

$$h_n = \max(g_{1n}, g_{2n}, \ldots, g_{nn}).$$

Now, certainly h_n is increasing

$$\begin{aligned}
h_{n+1} &= \max(g_{1,n+1}, g_{2,n+1}, \ldots, g_{n,n+1}, g_{n+1,n+1}) \\
&\geq \max(g_{1,n}, g_{2,n}, \ldots, g_{n,n}, g_{n+1,n+1}) \\
&\geq \max(g_{1,n}, g_{2,n}, \ldots, g_{n,n}) \\
&= h_n.
\end{aligned}$$

To see that $h_n \nearrow f$, for any $x \in X$ and $\epsilon > 0$, we see that we can find some n' so $|f(x) - f_{n'}(x)| < \epsilon/2$. We can then find some m, which we may assume to be at least n', so $|f_{n'}(x) - g_{n'm}(x)| < \epsilon/2$. Now taking $n = m$, we have $h_n(x) \geq g_{nm}$, and so $f(x) - h_n(x)| < \epsilon$.

We're now all done, except for the proofs of the lemmas!

Proof of Lemma A.34: We need to show that if $f_n \nearrow f$ in $\mathcal{F}_s$, then $\Phi(f_n) \nearrow \Phi(f)$.

We first consider the case $f = \chi_A$ for some Borel set A. The general case is an easy consequence. Pick $\epsilon > 0$, and let $B_n = \{f_n \geq 1 - \epsilon\}$. Observe $(1 - \epsilon)\chi_{B_n} \leq f_n$. Since $f_n \nearrow \chi_A, \bigcup B_n = A$, and so $\mu(B_n) \nearrow \mu(A)$ by countable additivity of the measure.

For each n, f_n is a linear combination, with coefficients in $[0, 1]$, of characteristic functions of disjoint sets all contained inside A. Thus $\Phi(f_n) \leq \mu(A)$.

Now we have

$$\begin{aligned}
\mu(A) &\geq \lim \Phi(f_n) \\
&\geq (1 - \epsilon) \lim \Phi(\chi_{B_n}) \\
&= (1 - \epsilon) \lim \mu(B_n) \\
&= (1 - \epsilon)\mu(A).
\end{aligned}$$

Taking $\epsilon \to 0$ gives the result. □

Proof of Lemma A.35: We need to show that $\lim \Phi(f_n) = \lim \Phi(g_n)$. Since $\min(f_n, g_n) \nearrow f_n$, $\lim \Phi(g_m) \geq \lim_n \Phi(\min(f_n, g_n)) = \Phi(f_n)$ by Lemma A.34. Thus $\lim \Phi(g_m) \geq \lim \Phi(f_n)$, and by symmetry we're done. □

---

[1](The problem is that at each point x, we can find some n so $|f(x) - f_n(x)| < \epsilon/2$, but then the sequence $\{g_{nm}(x)\}$ might take too long — in particular until $m > n$ to come within $\epsilon/2$ of $f_n(x)$.)

Proof of Lemma A.36: We need to constructing an approximating sequence $\{f_n\} \subset \mathcal{F}_s$ so $f_n \nearrow f$ for any $f \in \mathcal{F}_+$.

Define

$$A_{nk} = \{(k-1)2^{-n} < f \le k2^{-n}\}$$

and let

$$f_n = \sum_{k \ge 1}(k-1)2^{-n}\chi_{A_{nk}}.$$ □

That concludes the proof of Theorem A.33. □

Definition A.37. Suppose μ is a Radon measure on X.

Suppose $f \in C_c(X)$. We can write $f = f_+ + f_-$, where $f_+ = \max(f, 0) \in C_c(X)_+$ and $f_- = \min(f, 0)$, with $-f_- \in C_c(X)_+$. We have some $\alpha \ge 0$ and a compact set K so that $f_+ \le \alpha\chi_K$. Since μ is a Radon measure, $\mu(K) < \infty$, so $\Phi(f_+) < \infty$. (Here Φ is the positive linear functional defined in Theorem A.33 from μ.) Similarly $\Phi(-f_-) < \infty$.

We then check that if we had picked some other g_+ in $C_c(X)$ and g_- with $-g_- \in C_c(X)$ so $f = g_+ + g_-$, then $\Phi(f_+) - \Phi(-f_-) = \Phi(g_+) - \Phi(-g_-)$. This follows easily from the observations that in this case $g_+ = f_+ + k$ and $g_- = f_- - k$ for some $k \in C_c(X)_+$.

We define the integral $\int_\mu$ by $\int_\mu f = \Phi(f_+) - \Phi(-f_-)$. We easily see $\int_\mu$ is a Radon integral: positivity and linearity follow from the same conditions for Φ, and the invariance condition of the last paragraph.

A.5 Measures from Radon integrals

We first define

$$\mathcal{M}' = \left\{B \subset X \mid \chi_B \in L^1(X)\right\}$$

and then

$$\mathcal{M} = \{A \subset X \mid A \cap K \in \mathcal{M}' \text{ for all compact } K\}$$

Lemma A.38. *If K is a compact set, χ_K is integrable.*

Proof: If K is compact, $\chi_K \in C_c(X)_m$ by Lemma A.19. Clearly $\int_* \chi_K \ge 0$, so by Lemma A.21 χ_K is integrable. □

It is immediate from the definitions that we then have every compact set in $\mathcal{M}'$, and hence also in $\mathcal{M}$.

Lemma A.39. *The collection of subsets $\mathcal{M}$ is a σ-algebra.*

Proof: Certainly the zero function is integrable, so $\emptyset \in \mathcal{M}$.

If $A \in \mathcal{M}$, for any compact set K we have

$$\chi_{(X\setminus A)\cap K} = \chi_K - \chi_{A\cap K}$$

which is in $L^1(X)$ since by Lemma A.38 χ_K is integrable, and $A \in \mathcal{M}$ ensures $\chi_{A\cap K}$ is integrable. (Of course here we're using that $L^1(X)$ is a vector space, per Lemma A.29.) This shows that $X \setminus A$ is also in $\mathcal{M}$, so we see $\mathcal{M}$ is closed under taking complements.

We showed in Lemma A.30 that $L^1(X)$ is closed under max and min, so $\mathcal{M}'$ is closed under finite unions and intersections. By the monotone convergence theorem $\mathcal{M}'$ is also closed under countable unions: $\max\{\chi_{E_i}\}_{i=0}^n \nearrow \chi_{\bigcup E_i}$. $\qquad\square$

Lemma A.40. *The σ-algebra $\mathcal{M}$ contains all Borel sets.*

Proof: Given any closed set C, we see $C \cap K$ is compact if K is compact, so Lemma A.38 ensures $C \in \mathcal{M}$. $\qquad\square$

Theorem A.41. *If we define $\mu(A) = \int_{**} \chi_A$ for every $A \in \mathcal{M}$, then μ is a Radon measure.*

Proof: We first prove that μ is inner regular, that is,

$$\int_{**} \chi_A = \sup\left\{\int \chi_K \mid K \text{ a compact subset of } A\right\}.$$

Consider $h \in C_c(X)_m$, with $h \leq \chi_A$. As h is upper semicontinuous, the sets $C_n = \{h \geq n^{-1}\}$ are closed, and in fact compact, as h is dominated by some function with compact support. Since $h \leq 1$, in fact $h \leq \chi_{\bigcup C_n}$, and so $\int h \leq \sup \int \chi_{C_n}$. By definition $\int_{**} \chi_A$ is the supremum of such $\int h$, so

$$\int_{**} \chi_A \leq \sup \int \chi_{C_n} \leq \sup\left\{\int \chi_K \mid K \text{ a compact subset of } A\right\}.$$

The other inequality is easy, so we have proved inner regularity.

Next we check countable additivity of μ. Suppose we have a countable collection $\{A_n\}$ of disjoint sets in $\mathcal{M}$. Consider K a compact subset of $\bigcup A_n$. Then

$$\mu(K) = \int_{**} \chi_K$$
$$= \int \chi_K \qquad\qquad\qquad \text{(since } K \text{ is integrable)}$$

17

$$= \int \sum \chi_{K \cap A_n}$$

$$= \sum \int \chi_{K \cap A_n} \qquad \text{(by the MCT)}$$

$$= \sum \int_{**} \chi_{K \cap A_n} \qquad \text{(as } A_n \in \mathcal{M}, \chi_{K \cap A_n} \text{ is integrable)}$$

$$= \sum \mu(K \cap A_n). \qquad\qquad\qquad (\text{A.2})$$

We now prove that $\mu\left(\bigcup A_n\right) \leq \sum \mu(A_n)$ and then that $\mu\left(\bigcup A_n\right) \geq \sum \mu(A_n)$.

Easily from Equation (A.2) we have $\mu(K) \leq \sum \mu(A_n)$. Since this holds for any compact K inside $\bigcup A_n$, it holds for the supremum too, and inner regularity gives

$$\mu\left(\bigcup A_n\right) = \sup\left\{\mu(K) \mid K \text{ a compact subset of } \bigcup A_n\right\}$$
$$\leq \sum \mu(A_n).$$

In the other direction, we write

$$\mu\left(\bigcup A_n\right) = \sup_{\text{cpct } K \subset \bigcup A_n} \mu(K)$$
$$= \sup_{\text{cpct } K \subset \bigcup A_n} \sum \mu(K \cap A_n)$$

and interchanging supremums and sums,

$$= \sum \sup_{\text{cpct } K \subset \bigcup A_n} \mu(K \cap A_n)$$
$$\geq \sum \sup_{\text{cpct } K \subset A_n} \mu(K \cap A_n)$$

(as this is a supremum over a more restrictive family)

$$= \sum \sup_{\text{cpct } K \subset A_n} \mu(K)$$
$$= \sum \mu(A_n)$$

by inner regularity again. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

A.6 These constructions are inverses!

We start with a Radon integral $\int$, define a Radon measure μ following §A.5, and then define a new Radon integral $\int_\mu$ following §A.4. First, if $f = \sum_{n=0}^{N} \alpha_n \chi_{A_n}$, with $\alpha_n \geq 0$ and A_n Borel and

compact, we see

$$\int_\mu f = \sum \alpha_n \mu(A_n) = \sum \alpha_n \int_{**} \chi_{A_n} = \int_{**} f = \int f.$$

(We used the assumption that f had compact support in the last step.)

Next we can approximate any $f \in C_c(X)_+$ by such functions, with $f_n \nearrow f$. Then

$$\int_\mu f = \Phi(f)$$

$$\qquad = \lim \Phi(f_n) \qquad\qquad \text{by the second condition in Theorem A.33}$$

$$\qquad = \lim \int_\mu f_n$$

$$\qquad = \lim \int f_n \qquad\qquad\qquad \text{by the previous paragraph}$$

$$\qquad = \int f \qquad\qquad\qquad\quad \text{by the monotone convergence theorem}$$

Easily this gives the result that $\int_\mu = \int$.

In the other direction, we start with a Radon measure μ, construct a Radon integral $\int_\mu$ according to §A.4, and then construct another Radon $\widehat{\mu}$ following §A.5. As both measures are inner regular, it's enough to show that $\mu(K) = \widehat{\mu}(K)$ for every compact set K.

Approximate χ_K using Lemma A.19, obtaining $f_n \in C_c(X)$ with $f_n \searrow \chi_K$. Then $\widehat{\mu}(K) = \int_{**} \chi_K = \lim \int f_n$ by Lemma A.23, and each of these is at least $\mu(K)$, so we have $\widehat{\mu}(K) \geq \mu(K)$. In fact, by inner regularity we now know that $\widehat{\mu}(B) \geq \mu(B)$ for any Borel set B.

To obtain the other inequality, we consider some $f \in C_c(X)$, with $0 \leq f \leq 1$, and $f \geq \chi_K$. Observe that $f^n \searrow \chi_L$, where $L = \{f = 1\}$. Now,

$$\widehat{\mu}(L) = \int_{\mu**} \chi_L = \lim \int_\mu f^n$$

and since $f - f^n \nearrow f - \chi_L$, we can apply the second condition in Theorem A.33 to obtain

$$\lim \int_\mu f^n = \int_\mu \chi_L = \mu(L).$$

Finally, if we had $\mu(K) < \widehat{\mu}(K)$ we would obtain a contradiction from $\widehat{\mu}(L) = \widehat{\mu}(K) + \widehat{\mu}(L \setminus K) > \mu(K) + \mu(L \setminus K) = \mu(L)$.

Phew! That completes the proof of *Riesz' representation theorem*:

Theorem A.42. *On a nice[2] space X, there is a bijection between Radon measures and Radon integrals, given by the constructions of §A.4 and §A.5.*

A.7 Stieltjes integrals

We now construct all Radon integrals on $\mathbb{R}$. (This is very specific to $\mathbb{R}$, using the order structure on $\mathbb{R}$ is an essential manner; it's much harder to describe all Radon integrals on anything bigger!)

Let m be a monotone non-decreasing function on $\mathbb{R}$.

We'll now construct a Radon integral $\int_m$ from m. For each $f \in C_c(\mathbb{R})$, pick some interval $[a, b]$ containing its support. Given a partition $\lambda = \{a = \lambda_0 < \lambda_1 < \cdots < \lambda_n = b\}$, define upper and lower Riemann-Stieltjes sums by

$$\sum_{\lambda}^{*} f = \sum_{k=0}^{n-1} \left(\sup_{x \in [\lambda_k, \lambda_{k+1})} f(x) \right) (m(\lambda_{k+1}) - m(\lambda_k))$$

$$\sum_{\lambda*} f = \sum_{k=0}^{n-1} \left(\inf_{x \in [\lambda_k, \lambda_{k+1})} f(x) \right) (m(\lambda_{k+1}) - m(\lambda_k)) \,.$$

(These are all Riemann-Stieltjes sums because when $m(x) = x$ these are just the Riemann sums for a partition.)

Clearly $\sum_{\lambda}^{*}$ is decreasing as the partition λ becomes finer, and similarly $\sum_{\lambda*}$ is increasing. Clearly $\sum_{\lambda}^{*} - \sum_{\lambda*} f \geq 0$, and we next show that this difference converges to zero. With that, we define the Stieltjes integral $\int_m f = \lim_\lambda \sum_{\lambda}^{*} = \lim_\lambda \sum_{\lambda*}$.

As f is *uniformly* continuous on $[a, b]$, for any $\epsilon > 0$ we have a $\delta > 0$ so $\sup_{x \in [\lambda_k, \lambda_{k+1}]} f(x) - \inf_{x \in [\lambda_k, \lambda_{k+1}]} f(x) < \epsilon$ as long as $\lambda_{k+1} - \lambda_k < \delta$. Thus for any partition in which every interval is of length at most δ, with have $\sum_{\lambda}^{*} - \sum_{\lambda*} f < \epsilon(m(b) - m(a))$.

Exercise. Show the the Stieltjes integral is actually linear.

It's immediate that the Stieltjes integral is positive, so we have constructed a Radon integral.

Before we show that all Radon integrals are of this form, we make some observations about monotone functions. One can see that such a function has at most countably many points of discontinuity. If m is not already lower semicontinuous (for a monotone increasing function this is the same as continuous from the left), we can modify its values at those discontinuities so that it becomes lower semicontinuous. Explicitly, we can define a new function $m'(x) = \sup_{y < x} m(y)$, which is still monotone increasing, is lower semicontinuous, and agrees with m except at countable many points.

---

[2](locally compact, Hausdorff, second countable)

Exercise. Check that the Stieltjes integral is not affected by this modification: that is $\int_m = \int_{m'}$ as positive linear functionals on $C_c(X)$.

In fact, every Radon integral of $\mathbb{R}$ is a Stieltjes integral, for some lower semicontinuous monotone function $m : \mathbb{R} \to \mathbb{R}$. Fix some Radon integral $\int$, and construct the corresponding class of integrable functions $L^1(X)$. Define

$$m(x) = \begin{cases} \int \chi_{[0,x)} & \text{if } x \geq 0 \\ -\int \chi_{[x,0)} & \text{if } x < 0. \end{cases}$$

The monotone convergence theorem ensures that m is lower semicontinuous (and it's easy to see it's monotone non-decreasing, from the positivity of $\int$).

We need to show $\int = \int_m$. Easily, $f \leq \sum_{k=0}^{n-1} \left(\sup_{x \in [\lambda_k, \lambda_{k+1})} f(x) \right) \chi_{[\lambda_k, \lambda_{k+1})}$ for any partition λ, so

$$
\begin{aligned}
\int f &\leq \sum_{k=0}^{n-1} \int \left(\sup_{x \in [\lambda_k, \lambda_{k+1}]} f(x) \right) \chi_{[\lambda_k, \lambda_{k+1}]} \\
&= \sum_{k=0}^{n-1} \int \left(\sup_{x \in [\lambda_k, \lambda_{k+1}]} f(x) \right) (m(\lambda_{k+1}) - m(\lambda_k)) \\
&= \sum_{\lambda}^{*} f
\end{aligned}
$$

and similarly $\int f \geq \sum_{\lambda *} f$. Since these quantities converge to $\int_m f$, we have the desired result.

Exercise. If $m : \mathbb{R} \to \mathbb{R}$ is monotone non-decreasing, and absolutely continuous, so in particular m' exists and is integrable, show that

$$\int_m \chi_E = \int_E F'(x) dx$$

where the right hand side is Lebesgue integration, and more generally

$$\int_m f = \int f(x) m'(x) dx.$$

Exercise. If

$$m = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

what is $\int_m$?

Exercise. Show that $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}$ if and only if m is continuous.

A.8 Appendix: some technical details

Definition A.43. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is *lower semicontinuous* if $f^{-1}(t, \infty]$ is open for every real t.

Exercise. A function is lower semicontinuous if and only if for every convergent sequence $x_n \to x$, $f(x) \leq \liminf f(x_n)$.

Exercise. The supremum of an arbitrary family of lower semicontinuous functions is still lower semicontinuous.

Lemma A.44. *Every function $f \in C_c(X)^m$ is lower semicontinuous.*

Proof: Such a function is the supremum of a sequence of continuous functions. ☐

Lemma A.45. *If $f_n \searrow 0$, and each f_n is upper semicontinuous, positive, and compactly supported, then $||f_n||_\infty \searrow 0$.*

Proof: For any ϵ, the set $\{f_n \geq \epsilon\}$ is closed (this is direct from f_n being upper semicontinous). Since $\{f_n \geq \epsilon\}$ is contained in the support of f_n, which has been assumed to be compact, these sets are compact as well.

Now $\bigcap_{n \geq 0} \{f_n \geq \epsilon\} = \emptyset$, and by a standard compactness argument this means only finitely many of $\{f_n \geq \epsilon\}$ (as n varies) are non-empty. That is, eventually $||f_n||_\infty \leq \epsilon$, and since this argument held for any ϵ, we have the result. ☐

Lemma A.46. *If $\{f_n\} \subset C_c(X)_m$, and $f_n \searrow 0$, then $\int_* f_n \searrow 0$.*

Proof: We first check that we can apply Lemma A.45. The functions f_n are upper semicontinuous by Lemma A.44. Since each $f_n \geq 0$, it has compact support. Thus Lemma A.45 shows that $||f_n||_\infty \searrow 0$.

Note that $\lim \int_* f_n$ exists, since this is a sequence of decreasing non-negative numbers. We now pass to a subsequence $\{f_{n_k}\}$ so that $||f_{n_k}|| < 2^{-k}$, and for each k pick $g_k \in C_c(X)$ with $f_{n_k} \leq g_k \leq 2^{-k}$. Since the f_n are decreasing, we can further assume the supports of g_k are decreasing (by replacing g_k with $\min(g_1, \ldots g_k)$, say). Then define $g = \sum g_n$, which is in $C_c(X)_+$, and we have $\sum \int g_n \leq \int g < \infty$.

Finally $\sum \int_* f_n \leq \sum \int g_n < \infty$, and so we must have $\int_* f_n \searrow 0$. ☐

Proof of Lemma A.23: Recall we want to show that if $\{f_n\} \subset C_c(X)$ and $f_n \nearrow f$ for some $f \in C_c(X)^m$, then $\int f_n \nearrow \int^* f$.

For any compactly supported continuous g with $g \le f$, we have $g - \min(f_n, g) \searrow 0$, so by Lemma A.46

$$\int_* g - \min(f_n, g) \searrow 0$$

which we can re-express as

$$\int_* g = \lim \int \min(f_n, g)$$

which in turn is less that $\int f_n$. Since this held for any such $g \le f$, we also have $\int^* f \le \lim \int f_n$. Since $f_n \le f$, the other direction is trivial, and we have the result. $\square$

Proof of Lemma A.26: Recall that we want to show that if $\{f_n\} \subset C_c(X)^m$ and $f_n \nearrow f$, then $\int^* f_n \nearrow \int^* f$.

By the monotonicity of $\int^*$, we have $\lim \int^* f_n \le \int^* f$.

Now suppose $g \in C_c(X)$ and $g \le f$. We have $g - \min(f_n, g) \searrow 0$, so

$$\int_* (g - \min(f - n, g)) \searrow 0$$

by Lemma A.46. However by linearity we can write

$$\int_* (g - \min(f - n, g)) = -\int^* -g + \min(f_n, g)$$
$$= \int g - \int^* \min(f_n, g)$$
$$\ge \int g - \int^* f_n$$

Putting these facts together, $\lim \int g - \int^* f_n \le 0$, so $\int g \le \int^* f_n$. As this holds for every continuous compactly supported $g \le f$, $\int^* f \le \int^* f_n$, completely the claim.

$\square$

A.9   Appendix: Some point set topology

Definition A.47. A topological space X is *Hausdorff* if for every two points x, y, there exist disjoint open sets U, V with $x \in U$ and $y \in V$.

Definition A.48. A topological space X is *locally compact* if every point has a compact neighbourhood, i.e. for every $x \in X$, there exists an open set U and a compact set K so that $x \in U \subset K$.

When X is Hausdorff, an equivalent formulation of local compactness is that "every neighbourhood contains a compact neighbourhood".

Definition A.49. A topological space X is *separable* if it contains a countable dense set.

Definition A.50. A topological space X is *second countable* if there is a countable collection (U_i) of open sets, such that every open set is the union of some of the U_i.

Every second countable space is separable, but not vice versa! In one important way second countable spaces behave like metric spaces — convergence of sequences (rather than of nets) determines all the topological information. It is a consequence of this that in second countable spaces (just as in metric spaces) sequential compactness and compactness are equivalent.

We collect here a few useful facts.

Lemma A.51. *If X is a locally compact metric space, and K is a compact set, there is an $\epsilon > 0$ so the set*

$$\{x \mid d(x, K) \leq \epsilon\}$$

is still compact.

Proof: Using local compactness choose $r_k > 0$ and $K_k \subset X$ for each $k \in K$ so that $B(k, r_k) \subset K_k$ and K_k is compact. The balls $\{B(k, r_k/2)\}_{k \in K}$ form an open cover of K, so there is a finite subcover indexed by $I \subset K$. Let $r = \min_{i \in I} r_i$. Now, if $d(x, K) \leq r/2$, then $d(x, y) \leq r/2$ for some $y \in K$, and that y is in turn inside $B(i, r_i/2)$ for some $i \in I$. Thus $x \in B(i, r_i) \subset \bigcup_{i \in I} K_i$. Since the set $\{x \mid d(x, K) \leq r/2\}$ is closed, and a subset of the compact set $\bigcup_{i \in I} K_i$, we see that it is compact. $\qquad \square$