Supplement to *Real Analysis* by Folland

Krishna Chebolu
University of Missouri-Columbia

Abstract

During my graduate studies at the University of Missouri-Columbia, to study analysis, our course used the popular text *Real Analysis: Modern Techniques and Their Applications, 2nd Edition* by Gerald B. Folland. Although a classic, I found the text to be quite terse. To better understand this material, and aid others, I compiled this document as a supplement to the textbook– it attempts to make the material more digestible. Often times, mathematics is masked by notation. Though, at the level you presumably are now, you are expected to be comfortable with daunting notation. I have no such expectation from you. Keep in mind that I am not attempting to recreate the material– not everything is elaborated upon.

Notify the author[1] of typos.

1 Measures

I suppose size does matter.

1.1 Introduction

The purpose of this section is to provide motivation and help in the development of a *measure*, though, this term has not been used yet.

In 1-dimension, to measure *size*, say of an interval, you would just view the length of the interval. In 2-dimensions, you think of *area*– in integral calculus, we learn to compute the area under a curve. Similarly, in 3-dimensions, the notion of size is now *volume*. However, something we have always assumed or worked with are *nice things*– what do I mean by nice? I mean that all the objects we are *measuring* (title drop– spoiler alert!) are not necessarily complicated.

Length, area, and volume are low-dimensional ways of a *measure*; a generalization of size. What about measuring stuff in higher dimensions? Another question is– can we measure anything? Surely, we cannot. Not all mathematical objects are easy to measure, and you are right in thinking this. However, in an ideal world, what would a measure look like for n-dimensional objects?

Let us view a measure as a mapping. So, consider $E \subset \mathbb{R}^n$. A measure μ, upon inputting E, must spit out a real a number $\mu(E) \in [0, \infty]$. What are some other properties we would like?

[1]kscydv[@]missouri[.]edu

i If I take the union of sets, how must the measure behave? In 3D, if I have a cube of volume x^3 in front of me and put another cube of the same size beside it– what is the volume of both cubes? $2x^3$! Similarly, take the measure of the union of sets must equal the sum of the measure of each individual set.

ii Take the same cube (of volume x^3). If I rotate or translate that cube (flip it, turn it on its side, move it around, ...), does the volume of the cube change? No. So, two congruent sets should have the same measure.

iii The a measure of a unit-cube (or unit-anything) is 1.

Having that said, we are asking for something that is too good to be true! These conditions contradict themselves. No such function μ can satisfy all three criteria. To convince you, the Folland constructs a non-measurable set in 1D; using *Vitali* construction[2].

The construction: consider an equivalence relation[3] on $[0, 1)$ such that two variables are related ($x \sim y$) if and only if their difference ($x - y$) is rational. Choose N and R as in the text. In N, there is no canonical way to pick a representative of each equivalent class, but N includes only one from each such class; this is where the *Axiom of Choice* is used: it says that some choice exists.

Then the text defines translated copies of N called N_r. You have the interval $[0, 1)$, you shift it to the right by r. You will have a part on the right that sticks out (by $(1+r)-1 = r$), take the part that sticks out and insert it in the gap created by the initial shift to the right (which is of size r, so it fits snug). Voila, you are back in the interval $[0, 1)$. Also, $N_{r_1} \neq N_{r_2}$ for all $r_1 \neq r_2$. Why? because no two elements of N differ by a rational; if they did differ by a rational amount, they would already be accounted for since N contains only ONE representative of each equivalence class. So, all the N_r are disjoint.

Further, the disjoint union of N_r for all $r \in R$ is the interval $[0, 1)$. Why? Since for any $x \in [0, 1)$, we can find a rationally equivalent element $n \in N$, so $x = n + r,\ r \in R$. So, $x \in N_r$. This is true even when x is irrational.

By criteria *iii*, the measure of [unit] interval is 1, and by criteria *i*, the measure must be equal to the sum of the measure of each copy of N_r. That is not possible! Via our construction of N_r (recall the whole wrap-around the part that sticks out and fitting snug back into the unit interval), we have that the measure of N and any translated copy N_r must be the same. Then the measure of the sum of each copy N_r is either 0 (if $\mu(N) = 0$) or ∞ (if $\mu(N) > 0$; adding infinite copies of r).

So what do we do with all these inconsistencies?[4] It is clear that we can define special sets that cannot be *measured*. This raises the question: where are the nice things that are dying to be measured? Do not be afraid, they are here, and they are called...

[2]Classic method to construct a set that is not measurable; depends on the *Axiom of Choice*. It ultimately shows that not all subsets of the real numbers can be assigned a meaningful "length."

[3]Remember that this is a binary relation that is reflexive (I am my own friend), symmetric (I am your friend and you are my friend), and transitive (I am your friend and he is your friend so he is my friend).

[4]Vsauce has a great video on the Banach-Tarski paradox. Click here

1.2 σ-Algebras

These are the mathematical objects that can be measured. The definition of an *algebra* and $\sigma-$*algebra* is straightforward. Though, something to make explicit is the *disjointification of sequences*. You can replace sequence of sets with a disjoint sequence as follows.

Consider $\{E_j\}_1^\infty \subset \mathcal{A}$. Then define $F_1 = E_1$, $F_2 = E_2 \backslash E_1$, $F_3 = E_3 \backslash (E_2 \cup E_1)$, and so on. Then you end up with a sequence of disjoint sets $\{F_k\}_1^\infty$. Note that the union of E_js and F_ks are equal:

Proof. The construction makes clear $\cup F_k \subset \cup E_j$. For the other direction, consider $x \in \cup E_j$, then $\exists m$ such that $x \in E_m$. Then define $m_0 := \min\{m : x \in E_m\}$ (in case x is in more than one set; we can use well-ordering principle to do this). Then $x \in E_{m_0}$. Thus, for all $j < m_0$, we have that $x \notin E_j$. So, by definition of the F_ks, we have that $x \in F_{m_0}$, showing $\cup E_j \subset \cup F_k$. $\square$

Another important property: the intersection of σ-algebras is also a σ-algebra. Assuming this property, if $\mathcal{E}$ is a subset of the power set of X, then there is a unique smallest $\sigma-$algebra $\mathcal{M}(\mathcal{E})$ containing $\mathcal{E}$; predictably, $\mathcal{M}(\mathcal{E})$ is called the $\sigma-$algebra *generated* by $\mathcal{E}$.

Borel $\sigma-$algebras are important! They are the σ-algebras generated by the collection of open sets in X (a metric or topological space). Since B_X is a $\sigma-$algebra, it contains complements and the complements of open sets are, you guessed it, closed sets. Thus, $\mathcal{B}_X$ contains closed sets, and you can generate B_X via closed sets as well. The Borel $\sigma-$algebra on $\mathbb{R}$ is denoted $\mathcal{B}_\mathbb{R}$ and is generated by the atom-like sets shown in **proposition 1.2**.

The product $\sigma-$algebra is not intuitive. Concretely, consider an index set $A = \{1, 2\}$. You may expect a product to be defined as in Wikipedia,

$$\mathcal{M}_1 \times \mathcal{M}_2 = \mathcal{M}\left(\{E_1 \times E_2 : E_1 \in \mathcal{M}_1, E_2 \in \mathcal{M}_2\}\right).$$

Makes sense: the product of the two $\sigma-$algebras is the σ-algebra generated by the product of the two sets, E_1 and E_2. So what even is the definition in Folland? Take the same index set, and let's observe what happens:

$$\begin{aligned}\otimes_{\alpha \in A}\mathcal{M}_\alpha := \left\{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{M}_\alpha, \alpha \in \{1,2\}\right\} &= \{\pi_1^{-1}(E_1) : E_1 \in \mathcal{M}_1\} \cup \{\pi_2^{-1}(E_2) : E_2 \in \mathcal{M}_2\} \\ &= \{E_1 \times X_2 : E_1 \in \mathcal{M}_1\} \cup \{X_1 \times E_2 : E_2 \in \mathcal{M}_2\}.\end{aligned}$$

So, the product $\sigma-$algebra is generated by $\{E_1 \times X_2\}$ and $\{X_1 \times E_2\}$, and since $\sigma-$algebras are closed under countable intersections, the intersection of two sets $E_1 \times E_2$ is included. Thus, $\mathcal{M}(E_1 \times E_2) \subset \left\{\pi_\alpha^{-1}(E_\alpha)\right\}$, by **lemma 1.1**.

If the index set A is countable[5], **proposition 1.3** states that the product $\sigma-$algebra is generated by $\{E_1 \times E_2 \times \cdots \times E_\alpha \times \dots\}$. This is precisely the case that the Wikipedia definition above states.

Proposition 1.4 offers a more general way to construct product $\sigma-$algebras. There are two parts to this claim: first, for $\alpha \in A$, suppose $\mathcal{M}_\alpha$ (a $\sigma-$algebra) has a generator

[5] This means that you can create a 1-1 mapping with $\mathbb{N}$.

$\mathcal{E}_\alpha$, then the product is generated by $\mathcal{F}_1 := \{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{E}_\alpha\}$. Note that there are no assumptions on the index set A. Even for an uncountable set, one can generate the product $\sigma-$algebra. Now, why is the countable index set treated differently from the uncountable index set? Because $\sigma-$algebras are not closed under uncountable intersections. What does $\mathcal{F}_1$ look like? It is the set of all tuples in X whose αth coordinate lies in E_α.

Each element of $\mathcal{F}_1$ is a *cylinder*[6] set. We define the product $\sigma-$algebra as such to avoid the fact that we cannot take the intersection of uncountable rectangles; $\mathcal{F}_1$ works by restricting one coordinate at a time, giving it more *safety* as it builds the product. The second claim, on the other hand, is really **proposition 1.3**.

Moving on, **proposition 1.5** says that when you take a finite/countable collectio of metric spaces and form their product equipped with the natural product metric, the product $\sigma-$algebra generated by the Borel $\sigma-$algebras of each component space is always contained within the Borel $\sigma-$algebra of the product space. Further, if metric spaces are separable[7], then we have an equality. This means that the measurable structure on the product space can be completely understood by looking only at coordinate-wise measurable sets, making it easier to analyze and work with measures and integrals on the product.

This section concludes with defining an *elementary family*, a collection of X's subsets which contains the empty set, is closed under intersections, and for any element of this collection, its complement is a finite disjoint union of other elements (an element's complements breaks apart into other elements).

Proposition 1.7 says that if $\mathcal{E}$ is an elementary family of subsets of a set X, then the collection $\mathcal{A}$ of all finite disjoint unions of sets from $\mathcal{E}$ forms an algebra of sets; that is, $\mathcal{A}$ is closed under finite unions, intersections, and complements, and contains the empty set. Let us consider an example: let $X = [0, 3) \subset \mathbb{R}$, and define the elementary family

$$\mathcal{E} = \{[0, 1), [1, 2), [2, 3), \emptyset\}.$$

We check that $\mathcal{E}$ satisfies the conditions for an elementary family: $\emptyset \in \mathcal{E}$ by definition. $\mathcal{E}$ is closed under intersections: for example, $[0, 1) \cap [1, 2) = \emptyset \in \mathcal{E}$, and $[1, 2) \cap [1, 2) = [1, 2) \in \mathcal{E}$, etc. The complement of each set in $\mathcal{E}$ within X can be written as a finite disjoint union of members of $\mathcal{E}$:

$$\begin{aligned} [0,1)^c &= [1,2) \cup [2,3), \\ [1,2)^c &= [0,1) \cup [2,3), \\ [2,3)^c &= [0,1) \cup [1,2), \\ \emptyset^c &= [0,1) \cup [1,2) \cup [2,3). \end{aligned}$$

Then the algebra $\mathcal{A}$ generated by $\mathcal{E}$ consists of all finite disjoint unions of these intervals. Explicitly, $\mathcal{A}$ contains:

$$\emptyset, [0,1), [1,2), [2,3), [0,1) \cup [1,2), [0,1) \cup [2,3), [1,2) \cup [2,3), [0,1) \cup [1,2) \cup [2,3) = [0,3).$$

So $\mathcal{A}$ has $2^3 = 8$ elements — all subsets of $[0, 3)$ that can be formed by choosing any combination of the three intervals in $\mathcal{E}$.

[6] A cylinder set is called that because it generalizes the idea of a cylinder in geometry: a shape that's restricted in one direction, and extends infinitely in all others.

[7] Separable means that a space has a countable dense subset—a set with only countably many points (like $\mathbb{N}$) such that every point in the space can be approximated as closely as you want by points from that subset.

Select Exercises from 1.2

Exercise 3. Let $\mathcal{M}$ be an infinite $\sigma-$algebra.

a. $\mathcal{M}$ contains an infinite sequence of disjoint sets.

Proof. If we let the sets be empty, then the result is trivial (why: because the sequence of disjoint empty sets is infinite). So suppose the sets are nonempty.

We will show this via contraposition: assume there exists a nonempty set $E \in \mathcal{M}$ such that the restriction of $\mathcal{M}$ to E^c is finite. This would mean that the number of sets in $\mathcal{M}$ outside E is finite, and the number of sets within E are also finite (Why: if $\mathcal{M} \cap E^c < \infty$, then remaining sets of $\mathcal{M}$ are in E. Since $\mathcal{M}$ is disjoint, E's subsets are disjoint too, so there can only be a finite many.). If no such E exists, then it suggests that $\mathcal{M}$ must be finite– $\mathcal{M}$ does not contain infinitely many disjoint sets. W

Let E be a nonempty set. Then the restriction of $\mathcal{M}$ to E and E^c respectively are finite. For another set $F \in \mathcal{M}$, we have

$$F = (F \cap E) \cup (F \cap E^c).$$

However, each set in the union lies in one of the restrictions, so there are finitely many decompositions like the one above, so there are finitely many sets $F \in \mathcal{M}$. So, $\mathcal{M}$ is finite, thus, invoking the power of contraposition, the claim is proved.

Then you may construct a sequence of the sort: pick $A \in \mathcal{M}$ as above, then restrict $\mathcal{M}$ to A^c, and continue recursively. $\square$

b. $|\mathcal{M}| \geq c$

Proof. Using the result from above, let $\{A_n\}$ be the sequence. Then we can define a map $\varphi : 2^{\mathbb{N}} \to \mathcal{M}$ given by $\varphi(I) = \cup_{i \in I} A_i$ $(I \subset \mathbb{N})$. Since the sets A_n are disjoint, the union of any subset of these sets results in a unique set in $\mathcal{M}$. Specifically, if $I_1 \neq I_2$, then the sets of indices for which A_i is included in the union will differ, meaning the unions themselves will differ. Thus, $\varphi(I_1) \neq \varphi(I_2)$ whenever $I_1 \neq I_2$. This proves that the map is injective. Then comparing the cardinalities, we have $|\mathcal{M}| \geq |2^{\mathbb{N}}| = c$. $\square$

Exercise 4. An algebra $\mathcal{A}$ is a $\sigma-$algebra $\iff$ $\mathcal{A}$ is closed under countably increasing unions.

Proof. Since $\mathcal{A}$ is an algebra, we inherit $\mathcal{A}$ containing the empty set and being closed under complements and finite unions.

($\implies$) Suppose $\mathcal{A}$ is a $\sigma-$algebra. Then $\mathcal{A}$ is closed under countably infinite unions, by the definition of a $\sigma-$algebra. So, where those unions are increasing or decreasing are not relevant.

($\impliedby$) Let $\mathcal{A}$ be closed under countably increasing unions. This is not as straightforward. Consider a sequence $\{E_i\}_1^\infty \in \mathcal{A}$. Then define $F_j = \cup_{i=1}^j E_i$. Then the sequence $\{F_i\}_1^\infty$ has property $F_1 \subset F_2 \subset \dots$, so it is a sequence of countably increasing unions. So $\{F_i\}_1^\infty \in \mathcal{A}$. To do this, note that the countable union of countable unions is still countable. By construction

$\{E_i\}_1^\infty = \{F_i\}_1^\infty \in \mathcal{A}$. Thus, $\mathcal{A}$ is closed under countably infinite unions. So combining this newly found knowledge with the fact that $\mathcal{A}$ was already an algebra, we have our result. $\square$

Exercise 5. If $\mathcal{M}$ is the $\sigma-$algebra generated by E, then $\mathcal{M}$ is the union of $\sigma-$algebras generated by F as F ranges over all countable subsets of E.

Proof. A closer look at the question: we are asked to show essentially show that $\mathcal{M}(\mathcal{E}) = \cup_{\mathcal{F}\subset\mathcal{E}}\mathcal{M}(\mathcal{F})$. We can break this question into three parts: showing that the union of $\sigma-$algebras is a $\sigma-$algebra, and proving both sides inclusions in $\mathcal{M}(\mathcal{E}) = \cup_{\mathcal{F}\subset\mathcal{E}}\mathcal{M}(\mathcal{F})$.

(i) Show that a countable union of σ-algebras is a σ-algebra.

a. Empty set: Since $\emptyset \in \mathcal{M}(F_a)$, it follows that $\emptyset \in \bigcup_{a\in A}\mathcal{M}(F_a)$ for each $a \in A$, where A is the indexing set over $\mathcal{E}$.

b. Closure under sets in $\mathcal{M}(F_a)$: For each $F_a \in \mathcal{M}(F_a)$, it follows that $F_a \in \bigcup_{a\in A}\mathcal{M}(F_a)$.

c. Closure under complements: Since $F_a \in \mathcal{M}(F_a)$, we have $F_a^c \in \mathcal{M}(F_a)$, which implies that $F_a^c \in \bigcup_{a\in A}\mathcal{M}(F_a)$.

d. Closure under countable unions: If $\{F_{a_b}\}_{b\in B} \subset \mathcal{M}(F_a)$, then $\bigcup_{b\in B} F_{a_b} \in \mathcal{M}(F_a)$, where B is the indexing set. Therefore, the union of countably many sets from $\mathcal{M}(F_a)$ is still in $\mathcal{M}(F_a)$, and hence in the union $\bigcup_{a\in A}\mathcal{M}(F_a)$.

Therefore, the union $\bigcup_{a\in A}\mathcal{M}(F_a)$ is indeed a σ-algebra.

(ii) Show $\mathcal{M}(E) \subseteq \bigcup_{F\subset E}\mathcal{M}(F)$: Let $A \in \mathcal{M}(E)$. Since $B \in \mathcal{M}(E)$, we know that $B \subseteq E$ because all sets in $\mathcal{M}(E)$ are subsets of E. The sigma-algebra $\mathcal{M}(E)$ is closed under subsets, so B must also belong to some $\mathcal{M}(F)$, where $F \subseteq E$. This is because any measurable set A in $\mathcal{M}(E)$ can be described as measurable with respect to a subset $F \subseteq E$. In particular, if we take $F = B$, then $B \in \mathcal{M}(B) \subseteq \mathcal{M}(E)$, via **lemma 1.1**.

(iii) Show $\bigcup_{F\subset E}\mathcal{M}(F) \subseteq \mathcal{M}(E)$: Let $B \in \cup_{F\subset E}\mathcal{M}(F)$, then for some $a \in A$, $B \in \mathcal{M}(F_a)$. Since $F_a \subset E \implies F_a \subset \mathcal{M}(E) \implies \mathcal{M}(F_a) \subset \mathcal{M}(E)$ using **lemma 1.1**.

Therefore, we conclude that $\mathcal{M}(\mathcal{E}) = \bigcup_{\mathcal{F}\subset\mathcal{E}}\mathcal{M}(\mathcal{F})$. $\square$

1.3 Measures

Finally– measures! This section is really about learning the vocabulary; sentences[8] are in the upcoming sections.

Note that a measure always outputs a nonnegative real number. Makes sense, size cannot be negative, right? Some properties right off the bat: the size of nothing is 0 and if you have a sequence of finitely or countably disjoint sets, then the measure of the union of those sets is the sum of the measure of each set (called *countable* or *finite additivity*; former implies the latter).

[8]metaphorically speaking, of course

A *measure space* (different from a *measurable space*, which is just $(X, \mathcal{M})$) is $(X, \mathcal{M}, \mu)$: the main set, a $\sigma-$algebra on the set, and the measure on them. A measure is *semifinite* if the measure of the set in the $\sigma-$algebra is infinite, but has a subset whose measure is finite.

Some examples:

- The *counting measure* counts the no. of elements in a set.
- The *Dirac* or *point mass measure* assigns a point $x_0 \in X$ a measure of 1 and 0 for the rest.

Theorem 1.8 lists the properties of a measure space $(X, \mathcal{M}, \mu)$. *Monotonicity*: if a set is a subset of another, the measures reflect that. *Subadditivity*: if there exists a countable sequence of sets in $\mathcal{M}$, then the measure of the union of all those sets are less than (or equal to) the sum of the measure of each set. *Contininuity from below*: if there exists a countable sequence of sets in $\mathcal{M}$ and the set of a lower index is a subset of the following set ($E_{i-1} \subset E_i,\ \forall i$), then the measure of the union of the sets is $\lim_{j\to\infty} \mu(E_j)$; basically, the largest set's measure (not really the largest, but the *limit*). *Continuity from below*: same as above but $E_{i-1} \supset E_i, \forall i$, then the measure of the intersection of all the sets is the measure of the smallest set (the limit, which goes to the smallest set).

I like this one– *almost everywhere*; when a statement is true about points in the set X except for when x is in a null set ($\mu(E) = 0,\ E \subset \mathcal{M}$). In this case, when using a measure way of conveying this idea, we use $\mu-$*null* or $\mu-$*almost everywhere*.

Recall that the domain of a measure (μ) is the $\sigma-$algebra ($\mathcal{M}$). The $\sigma-$algebra $\mathcal{M}$ is essentially a curated collection of sets upon which we can construct the notion of a measure. Naturally, not all sets in X make the cut, and are disqualified from join the exclusive club that is $\mathcal{M}$. So it is possible that we have a subset $E \subset \mathcal{M}$ such that $\mu(E) = 0$ (a null set) which has a subset $F \subset E$ such that $F \notin \mathcal{M}$. This means that $\mu(F)$ is not defined. However, if $\mathcal{M}$ is a large enough to contain all subsets of null sets, we call that measure a *complete measure*. So, mathematically, this means

$$\text{If } N \in \mathcal{M} \text{ and } \mu(N) = 0 \implies \forall A \subset N,\ A \in \mathcal{M}.$$

Working with complete measures is nicer. We are saved the trouble from having to show where some subset of a measure-zero is actually measurable. It is a way to get around an annoying technicality. The next question you may have is: can we modify $\mathcal{M}$ to make it big enough to contain all the subsets of null sets? Yes, we can.

Theorem 1.9 does this. Start a measure space. The goal is to extend $\mathcal{M}$ to $\overline{\mathcal{M}}$; $\overline{\mathcal{M}}$ is basically all the sets already in $\mathcal{M}$ with some new sets– the subsets of the null sets. It is really that simple. To reiterate, just add those null sets' subsets into $\mathcal{M}$ and we are done. Put a bar over it to denote its *completion*. The measure for $\overline{\mathcal{M}}$ is expectedely denoted by $\overline{\mu}$. Our *complete measure space* is then $(X, \overline{\mathcal{M}}, \overline{\mu})$. The proof is essentially going through why the complete $\sigma-$algebra is indeed a $\sigma-$algebra and the measure is well-defined.

Select Exercises from 1.3

Exercise 7. If μ_j are measures on $(X, \mathcal{M})$ for each j and $a_j \in [0, \infty)$ for each j, show that $\sum_1^n a_j \mu_j$ is a measure on $(X, \mathcal{M})$.

Proof. To show that we have a measure, we must show the following two criteria:

i. The measure of the empty set is 0: We have $\sum_j a_j\mu_j(\emptyset) = \sum_j a_j(0) = 0$ since each μ_j is a measure and $a_j \cdot 0 = 0,\ \forall j$.

ii. Countable additivity: Let $\{E_k\}_1^\infty$ be a set of sequences. Then for a fixed j, each measure satisfies countable additivity, $\mu_j(\cup E_k) = \sum_k \mu_j(E_k)$. Then $a_j\mu_j(\cup E_k) = a_j \sum_k \mu_j(E_k) \implies \sum_j a_j\mu_j(\cup E_k) = \sum_j a_j \sum_k \mu_j(E_k)$.

   Then by the property of sums, we can interchange the order of summation, $\sum_j a_j \sum_k \mu_j(E_k) = \sum_k \sum_j a_j\mu_j(E_k)$. So we have that

$$\sum_{j=1}^{n} a_j\mu_j \left(\bigcup_{k=1}^{\infty} E_k \right) = \sum_{k=1}^{\infty} \sum_{j=1}^{n} a_j\mu_j(E_k).$$

   Changing the order of summation lets us apply **countable additivity** to each μ_j in the sum independently and ensures that the sum behaves properly across the countable union of disjoint sets. This is essential for showing that the sum of the measures $\sum_{j=1}^n a_j\mu_j$ is itself countably additive.

Thus, $\sum_1^n a_j\mu_j$ is a measure on $(X, \mathcal{M})$. □

Exercise 9. If $(X, \mathcal{M}, \mu)$ is a measure space and $E, F \in \mathcal{M}$, then $\mu(E) + \mu(F) = \mu(E \cup F) + \mu(E \cap F)$.

Proof. Observe that the set E can be written as $(E \cap F) \sqcup (E - F)$ and $E \cup F$ can be written as $(E - F) \sqcup F$. Also, since the measure is countably additive, it is also finitely additive. Then

$$\begin{aligned} \mu(E) + \mu(F) &= \mu((E \cap F) \sqcup (E - F)) + \mu(F) \\ &= \mu(E \cap F) + \mu(E - F) + \mu(F) \\ &= \mu(E \cap F) + \mu((E - F) \sqcup (F)) \\ &= \mu(E \cup F) + \mu(E \cap F). \end{aligned}$$

□

Exercise 10. Given a measure space $(X, \mathcal{M}, \mu)$ and $E \in \mathcal{M}$, define $\mu_E(A) = \mu(A \cap E)$ for $A \in \mathcal{M}$. Then μ_E is a measure.

Proof. We must show the following two properties:

i. $\mu_E(\emptyset) = \mu(\emptyset \cap E) = \mu(\emptyset) = 0$.

ii. $\mu_E(\cup_j A_j) = \mu((\cup_j A_j) \cap E) = \mu(\cup_j (A_j \cap E)) = \sum_j \mu(A_j \cap E) = \sum_j \mu_E(A_j)$.

□

Exercise 13. Every $\sigma-$finite measure is semifinite.

Proof. Let μ be a σ-finite measure on the measurable space $(X, \mathcal{M})$.

Firstly, if $\mu(X) < \infty$, μ will trivially be semi-finite. Therefore, suppose μ is σ-finite but not finite. Now, let us arbitrarily pick $E \in \mathcal{M}$ such that $\mu(E) = \infty$ (we know at least one such element exists, namely X, since otherwise μ would be finite).

From the definition of μ being σ-finite, we know that there exists a sequence $\{F_i\}_{i=1}^{\infty} \subset \mathcal{M}$ such that $X = \bigcup_{i=1}^{\infty} F_i$ and $\mu(F_i) < \infty$ for all $i \in \mathbb{N}$. One can easily see the following:

$$\mu(E) = \mu(E \cap X) = \mu\left(\bigcup_{i=1}^{\infty}(E \cap F_i)\right) \leq \sum_{i=1}^{\infty} \mu(E \cap F_i).$$

Since $\mu(E) = \infty$, we have:

$$\infty \leq \sum_{i=1}^{\infty} \mu(E \cap F_i) \quad \Rightarrow \quad \sum_{i=1}^{\infty} \mu(E \cap F_i) = \infty.$$

Furthermore, since $E \neq \emptyset$ (otherwise $\mu(E) = 0 < \infty$) and $\mu(E) = \mu\left(\bigcup_{i=1}^{\infty}(E \cap F_i)\right)$, we know there must exist at least one $k \in \mathbb{N}$ such that $\mu(E \cap F_k) > 0$. On the other hand, since $\mu(F_k) < \infty$ by construction, we also have $\mu(E \cap F_k) < \infty$. Therefore, since trivially $E \cap F_k \subset E$, we have shown that for an arbitrary $E \in \mathcal{M}$ with $\mu(E) = \infty$, there exists $k \in \mathbb{N}$ such that $F_k \cap E \subset E$ and $\mu(F_k \cap E) < \infty$.

□

Exercise 14. If μ is a semi-finite measure and $\mu(E) = \infty$, for any $C > 0$, there exists $F \subset E$ with $C < \mu(F) < \infty$.

Proof. Let $E = \{F \subset E : F \in \mathcal{M}, \mu(F) < \infty\}$, i.e., the set of all finite measurable subsets of E. We define

$$C = \sup\{\mu(F) : F \in E\}.$$

We aim to show that $C = \infty$. Why: The idea is that by showing $C = \infty$, we establish that the measure of the subsets of E can be arbitrarily large, but still finite. In the process, we also show that we can always find a set $F \subset E$ where $\mu(F)$ lies between any arbitrary positive C and infinity, thereby satisfying the condition $C < \mu(F) < \infty$.

Select a sequence $\{E_n\}_{n=1}^{\infty} \subset E$ such that $\lim_{n\to\infty} \mu(E_n) = C$. Now, define $F_n = \bigcup_{i=1}^{n} E_i$ and $F = \bigcup_{n=1}^{\infty} F_n$.

By the continuity of the measure from below, we have:

$$\lim_{n\to\infty} \mu(F_n) = \mu(F).$$

It's easy to see that for all n, $\mu(F_n) \geq \mu(E_n)$, so:

$$\mu(F) = \lim_{n\to\infty} \mu(F_n) \geq \lim_{n\to\infty} \mu(E_n) = C.$$

Now, we address the question of whether $\mu(F)$ is finite. The answer is yes. Suppose, for the sake of contradiction, that $\mu(F) = \infty$. Since $\lim_{n\to\infty} \mu(F_n) = \mu(F) = \infty$, there exists some N such that for all $n > N$, $\mu(F_n) > C$. But $F_n = \bigcup_{i=1}^{n} E_i$ is always finite, and for all n, $F_n \in E$, so $\mu(F_n) \leq C$. This is a contradiction.

Therefore, we must have $\mu(F) < \infty$, which implies that $F \in E$ and thus $\mu(F) \leq C$. Now, combining $\mu(F) \geq C$ and $\mu(F) \leq C$, we get:

$$\mu(F) = C.$$

If $C < \infty$, we can choose a finite measurable set $W \subset E - F$. Then:

$$\mu(W \cup F) > C,$$

which contradicts the definition of C. Thus, we must have $C = \infty$. □

1.4 Outer Measures

We know what a measure is... probably. This section is about how we construct a measure.

An *outer measure* $\mu^* : \mathcal{P}(X) \to [0, \infty]$ on a nonempty set X is generalized as follows (nothing we have not seen before),

i The outer measure of an empty thing is still 0.

ii The outer measure of a subset is less than the outer measure of its parent set.

iii The outer measure of the union of sets is less than (or equal to) the sum of the outer measure of each set.

It is similar to the definition of a measure, but not the same: in the definition of a measure, criteria *iii* is an equality rather than an inequality (but also, the measure definitions asks for the sets to be disjoint) and criteria *ii* is not explicitly stated. Also note that μ^*'s domain the power set of X. So we are attempting to define a measure of the largest possible $\sigma-$algebra; the *outer* in outer measure may make more sense if you think about it that way.

Let us construct; **proposition 1.10**. Here is the idea: you start with simple measurable building blocks, recall that *elementary set* definition we introduced above. Then we take countable union of these building blocks to approximate more complicated sets– use additivity to obtain the measure of the union. We then take the outer measure to be the smallest measure (that's why the inf) that you need to cover it with the elementary sets. This infimum may be slightly confusing, but it is saying that there are many possible sizes to measure a set via the union of elementary sets; so, we take the smallest possible one. An analogy: you can wear shirts larger than your size, but the one that looks best is the one that fits perfectly.

We have the notion of a set being measurable, but what about a notion for *outer measurability*? If E is nice such that $E \supset A$, the outer measure of A, $\mu^*(A) = \mu^*(E \cap A)$ is the same as the *inner measure* of A, $\mu^*(E) - \mu^*(E \cap A^c)$. Draw a venn diagram, you will see what the text means. But what if I loosen the restriction of E being a gentleman? What if E is not well-behaved?

Theorem 1.11, Carathéodory's theorem discusses this. The collection of μ^*-measurable sets is a $\sigma-$algebra, and μ^* when applied only to this collection is a complete measure.

We apply **theorem 1.11** to define a *premeasure*. We use this to extend measures from algebras to $\sigma-$algebras. Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra, then $\mu_0 : \mathcal{A} \to [0, \infty]$ is a *premeasure* if i the premeasure of the empty set is 0, and

ii if a sequence of disjoint sets in $\mathcal{A}$ whose union is contained in $\mathcal{A}$, then the premeasure of the union of the sets is the sum of the premeasure of each set.

Also, a premeasure is finitely additive[9]. A premeasure can induce the definition of an outer measure (of a set E): take the infimum of the sum of the premeasures of set from $\mathcal{A}$ that cover the set E; denoted, appropriately, by μ_0^*

So what even is going on with all these variations of a measure? A premeasure is the starting point for defining a measure. It assigns a "size" to simple, well-behaved sets such as finite unions of intervals, and behaves additively on those sets. From this initial notion of size, we build an outer measure, which extends the idea to all possible subsets of the space by covering each set with simple ones and taking the smallest total size that still covers it. The outer measure gives a consistent way to talk about the size of any set, but it may not behave additively. To fix this, we look at the subsets where the outer measure behaves nicely—those for which the measure of any set can be split exactly into the measures of its parts inside and outside that subset. These "good" sets form a σ-algebra, and restricting the outer measure to them produces a genuine measure.

TLDR; a premeasure defines sizes for simple sets, an outer measure extends this idea to all sets, and a measure refines it to a well-behaved collection of measurable sets.

Proposition 1.13 connects some of these ideas. Given a premeasure on an algebra ($\mathcal{A} \subset \mathcal{P}(X)$), the corresponding outer measure agrees with the premeasure on that algebra– i.e., when restricted to $\mathcal{A}$, we have $\mu_0(\mathcal{A}) = \mu^*(\mathcal{A})$. Moreover, every set in $\mathcal{A}$ is measurable on that algebra w.r.t. the outer measure in the sense of Carathéodory, ensuring that the original elementary sets remain measurable in the extended measure space.

The last idea of this section is presented in **theorem 1.14**. The goal of this theorem is build a full measure from a premeasure. We start with a premeasure μ_0 on simple sets, the algebra $\mathcal{A}$. We obtain the corresponding outermeasure μ^*. on all subsets of X by covering from the outside. Then, we restrict this outer measure to the nice sets of X, which is the $\sigma-$algebra $\mathcal{M}$ generated by the family of simple sets $\mathcal{A}$ from earlier. Bam, we have a measure μ on that $\sigma-$algebra. The second part of this theorem discusses the entry of another measure, $\nu(E) \leq \mu(E)$, which means that the construction we used creates the largest possible extension of the premeasure; no other extension gives a larger value. When our constructed measure is finite, then $\nu(E) = \mu(E)$ for all $E \subset \mathcal{M}(\mathcal{A})$. A further result is that if our premeasure (which we started with) is $\sigma-$finite[10], then the extension measure we constructed from the premeasure is unique.

Select Exercises from 1.4

My exam over this section is complete so I will skip this for now.

[9] Recall that this means that the premeasure of a finite union of sets is the same as the sum of the premeasure of each individual set. *Subadditive* would imply that rather than an equality, we would have an inequality– less than or equal to.

[10] Recall that, given a measure space $(X, \mathcal{M}, \mu)$, this means that X is a countable union of sets of a $\sigma-$algebra $\mathcal{M}$ and the measure of each set is finite. In other words, the premeasure doesn't assign infinite measure everywhere.

1.5 Borel Measures on the Real Line

Recall that the Borel $\sigma-$algebra is generated by the collection of open sets in the parent set X. So, $B_{\mathbb{R}}$ is generated by the open sets in $\mathbb{R}$. This section is about the family of measures whose domain is $B_{\mathbb{R}}$; these measures are *Borel measures* on $\mathbb{R}$.

The text provides an example to motivate this notion. Consider a function $F(x) = \mu((-\infty, x])$.

1. Then for $a > b$, we have that $F(a) = \mu((-\infty, a]) > \mu((-\infty, b]) = F(b)$, which makes it an increasing function.

2. Also, the intersection of all the sets from $-\infty$ to a sequence x_n would result in the smallest set $(-\infty, x]$ (as $x_n \downarrow x$). This matches the definition for *continuity from below* as given in **theorem 1.8**.

Having these properties identified, let $b > a$, then the interval $(-\infty, b]$ can be written as the union of $(-\infty, a] \cup (a, b]$ or $(-\infty, b] \backslash (-\infty, a] = F(b) - F(a)$.

Now, we will try to find the definition of a measure from an increasing, right-continuous function F. For any *half-open interval* (a.k.a., *h-intervals*) of the form $(a, b]$, the measure is given by

$$\mu((a, b]) = F(b) - F(a).$$

These h-intervals will be our building blocks. The intersection and the complement of an h-interval is an h-interval or the disjoint union of h-intervals. Thus, via **proposition 1.7**, the collection $\mathcal{A}$ of finite disjoint unions of h-intervals is an algebra. We can tighten this via **proposition 1.2**, the $\sigma-$algebra generated by $\mathcal{A}$ is $\mathcal{B}_{\mathbb{R}}$.

If $F(x) = \mu(-\infty, x] = x$, we obtain the *length* of the set, which is the usual measure we may be accustomed to.

Why does F need to be increasing and right-continuous? F needs to be increasing because it should not decrease as the interval grows in size; so naturally, every measure will need this. Right-continuity guarantees that the measure behaves well w.r.t. unions of disjoint intervals and ensures that limits of sequences of intervals are handled consistently. Without this property, we ma run into issues while defining the measure on intervals that approach infinity.

Proposition 1.15 defines a premeasure μ_0 on $\mathcal{A}$ (collection of finite disjoint unions of h-intervals). Let F be a real-values increasing, right-continuous[11] function. If we have disjoint h-intervals $(a_j, b_j],\ j = 1, ..., n$, we define

$$\mu_0\left(\bigcup_1^n (a_j, b_j]\right) = \sum_1^n [F(b_j) - F(a_j)],$$

and, a special case, $\mu_0(\emptyset) = 0$. This definition is generalizing what we have above ($\mu((a, b])$); we are then use the additivity property to obtain this definition. To prove this proposition, we must show that this is well-defined, finitely additive, and monotonic.

We can take this to the next level via **theorem 1.16**. The main result is that given an increasing, right-continuous function F, $\exists! \mu_F$ s.t. $\mu_F((a, b]) = F(b) - F(a)$, for all $a < b$. If

[11] Same as continuity from below.

G is another right-continuous, increasing function, then $\mu_F = \mu_G \iff F - G$ is constant. So if the difference is constant, both F and G define the same measure. On the flip side, if we are given a Borel measure μ and wish to recover F, provided μ is finite on all bounded Borel sets, we can define

$$F(x) = \begin{cases} \mu(0, x] & x > 0, \\ 0 & x = 0, \\ -\mu(-\infty, x] & x < 0. \end{cases}$$

In this case, $F(x)$ is right-continuous and increasing, and $\mu = \mu_F$. Remarks:

1. We could have developed all this theory with left-continuous functions with intervals $[a, b)$.

2. There is a difference between the recovered F in **theorem 1.16** and F defined toward the beginning of this section. The key difference is that the recovered F in **theorem 1.16** is normalized at 0. Using the definition in the beginning of the section, we would have $F(0) = \mu((-\infty, 0])$, but in the recovered $F(0) = 0$.

3. For each right-continuous, increasing F, we obtain not only its corresponding Borel measure μ_F, but also the completion $\overline{\mu_F}$; denoted by μ_F– no overline. This is the *Lebesgue-Stieltjes measure* associated to F.

What is a *Lebesgue-Stieltjes measure*? First note that a *Lebesgue measure* is a specific instead of an L-S measure[12]. Flip the script to say that an L-S measure is a generalization of the Lebesgue measure[13]. So for an L-S measure, the measure changes with changes in the distribution function $F(x)$. With this generalization, we certainly have some benefits.

Lemma 1.17 defines a L-S measure μ. For a set $E \in \mathcal{M}_\mu$, the measure is given by

$$\mu(E) = \inf \left\{ \sum_1^\infty \mu((a_j, b_j)) : E \subset \bigcup_1^\infty (a_j, b_j) \right\}.$$

This is similar to how we defined the outer measure. The measure of the set is the smallest possible sum of the measures such that the set E is contained in them. This lemma is also showing us that we can approximate the measure of any set E via open intervals rather than just h-intervals. Open intervals are nicer to work with.

Theorem 1.18 follows this up. It provides two different ways to discuss the measure of a set E. It is a discussion on regularity, and how we can exploit this to compute the measure in different ways.

a. You can measure them from the *outside*: infimum over open sets. The measure of E can be approximated from above by open sets U that cover E. Take the smallest such measure. This exploits the fact that the measure is continuous from above.

b. You can measure them from the *inside*: supremum over compact sets. You can approximate from the inside by compact sets K that are contained in E. Take the largest measure. This exploits the fact that the measure is continuous from below.

[12] A rectangle is really just a special parallelogram.

[13] When $F(x) = x$, the L-S measure becomes the Lebesgue measure.

This regularity property is important in measure theory because it allows you to compute the measure of complex sets by approximating them with simpler sets (open and compact sets) that are easier to handle.

Let us take a deeper look at the *Lebesgue measure*: the complete measure[14] μ_F associated to the function $F(x) = x$. For an interval, the Lebesgue measure is simply its length. Denote this measure by m[15]. Naturally, the domain of m is called the class of *Lebesgue measurable sets*, and we denote that via $\mathcal{L}$. One of the most important properties is that m is invariant under translations and behaves linearly under scaling. So, if $E \subset \mathbb{R}$ and $s, r \in \mathbb{R}$, we may define

$$E + s = \{x + s : x \in E\}, \quad rE = \{rx : x \in E\}.$$

The first definition implies that we just shift all the elements within the set by s, and the second is scaling a set by r means that we scale each element by r. We use this in **theorem 1.21**.

Theorem 1.21 essentially says that the measure does not change, and if it does, it behaves like a linear transformation. If $E \in \mathcal{L}$ (E is a Lebesgue measurable set), then $E + s$ is also a Lebesgue measurable set, so is rE; for all $s, r \in \mathbb{R}$. What about the measure? In the first case (translation), the measure does not change (think of moving an interval to the right– sure you moved it, but the length does not change). In the second case (scaling/dilution), the measure is scaled by r. If you multiply everything in an interval by 2, then the end points change: $(1, 2) \mapsto (2, 4)$; observe that the length of the interval also doubled ($\times 2$). Mathematically, $m(rE) = |r| \cdot m(E)$.

Now, let us highlight some properties.

1. Every singleton set ($\{x\}$) has Lebesgue measure 0. Makes sense: every singleton set contains one point, so there is no notion of length.

2. Every countable set also has Lebesgue measure 0. Why? Exploit countable additivity. Consider a set $A := \{x_1, x_2, ...\}$. Then $m(A) = \sum m(\{x_i\}) = 0 \cdot n = 0$ provided $|A| = n$. We do not even need countable additivity; countable subadditivity should suffice if the subsets are not disjoint.

3. $m(\mathbb{Q}) = 0$. Crazy!

4. A nonempty open set cannot have Lebesgue measure 0.

5. A topologically large set can have a small measure and vice versa. So take Lebesgue measures with a pinch of salt. You see it above– $m(\mathbb{Q}) = 0$. Just because the set is dense does not mean the measure will reflect that. So one tool cannot alone tell you about a set. You need to topological, as well as the measure-theoretic insight of a set.

6. The Lebesgue null sets (measure 0) include countable sets and sets having the cardinality of the continuum. What even is *continuum*? It refers to the cardinality (size) of the set of real numbers $\mathbb{R}$– it is not countable. So, this statement says that we can

[14] Recall that a complete measure is one that includes all the subsets of null sets.

[15] m for measure

have sets that are as "large" as $\mathbb{R}$, but still have a Lebesgue measure 0. A popular example is the *Cantor set*.

"Hmm, I have heard of Cantor sets before," you may think to yourself. Let me remind you what they are: start with the unit interval $[0, 1]$. Remove the middle third: $[0, 1]\backslash(\frac{1}{3}, \frac{2}{3}) = [0, \frac{1}{3}]\cup[\frac{2}{3}, 1]$. Remove the middle third from each interval; then you will remain with $[0, 1/9]\cup[2/9, 1/3]\cup[2/3, 7/9]\cup[8/9, 1]$. Keep removing the middle third from each interval repeatedly to obtain the Cantor set.

Regardless of how many middle-chunks we remove, we end up with a set that has the same cardinality as $\mathbb{R}$. Why? The Cantor set is uncountable. If we try to compute the measure of this new set, call it C for Cantor, we get

$$m(C) = 1 - \sum \frac{2^j}{3^{j+1}} = 1 - \frac{1}{3} \cdot \frac{1}{1 - (2/3)} = 1 - 1 = 0.$$

So, the measure is 0! You may not trust me that the sum above is correct. I profess the following,

- Step 0 (start): The interval is $[0, 1]$, so the total length is 1.
- Step 1: Remove the open middle third $(\frac{1}{3}, \frac{2}{3})$ from $[0, 1]$. Length removed: $\frac{1}{3}$.
- Step 2: Remove the middle third from each of the 2 remaining intervals. That is, remove 2 intervals, each of length $\frac{1}{9}$:
$$2 \cdot \frac{1}{9} = \frac{2}{9}$$
- Step 3: Remove the middle third from each of the 4 remaining intervals. That is, remove 4 intervals, each of length $\frac{1}{27}$:
$$4 \cdot \frac{1}{27} = \frac{4}{27}$$

Keep going and you will see a pattern emerge.

Combining the properties we have looked at, we obtain **proposition 1.22**: the measure is 0, the cardinality is that of $\mathbb{R}$, and C is compact, nowhere dense, and disconnected. Let me expand on the topological properties: clearly the interval $C \subset [0, 1]$ is closed (we always remove open sets, so we are left with closed ones) and bounded so via *Heine-Borel*, the interval is compact. A set is nowhere dense if its closure has empty interior — that is, it does not contain any open interval, no matter how small, and we have this– At each step, we're removing the middle third of every interval. So what's left becomes infinitely fragmented — no chunk of it contains a full open interval. Finally, disconnected-ness: At each stage, the construction splits the remaining pieces into smaller and smaller intervals with gaps between them. In the limit, any two points in the Cantor set are separated by a gap — you cannot move continuously from one point to another within C.

Select Exercises from 1.5

My exam over this section is complete so I will skip this for now.

2 Integration

We spent a decent amount of time discussing sizes of objects, let us now elaborate on the notion of area. But like most of mathematics, this area, too, is abstract.

In case you forgot, integration *was the whole point*[16].

2.1 Measurable Functions

[16] A nod to a popular internet trend mocking folks who post cliché posts with captions indicating that their purpose in life. In our case, integration is why we defined a measure— it was the whole point.