Everything You Need To Know About MAT457

Nathanael Chwojko-Srawley

November 8, 2021

Contents

The goal of this document is to encapsualte all the notes and intuition I will gain taking MAT457. The preliminary chapter will be a quick reminder on hy we care to generlizatize our notion of integrability.

Assumptions in reading this text

1. basic definition and notions in topology are assumed, covering at least:
   (a) open/closed sets, closure and interior, Hausdorff space, limit definition of closure (given the space if Hausdroff, or at least T_1), continuity
   (b) Compactness and connectedness. Heine-Borel Bolazno-iestass in $\mathbb{R}^n$. Path-connectedness
   (c) subspace topology, product topology, Quotient topology
   (d) Metric spaces, many equivalent metric on product spaces,
2. A good understanding of Riemann integration
3. A basic understanding of point-wise convergence and uniform convergence. For point-wise convergence, we'll use $f_n \to f$ notation. For uniform convergence, we'll use $f_n \rightrightarrows f$.
4. Defining and understanding basic properties of sup, inf, lim sup, and lim inf.

Since we are doing analysis, the axiom of choice is assumed. There are some fascinating results we get by assuming the axiom of choice which we will point out.

Preliminary Integration

In pre-university education, it is taught how to find the area of some simple shapes like rectangles, triangles, cubes, spheres, and so on. Later on, You learn how many of these shapes can be captured as the area under a particular graph (the graph usually being continuous). This means that the notion of area can be represented by functions (in particular the graph of a functions), and so it becomes a meaningful question to ask about how to find the area under the graph.

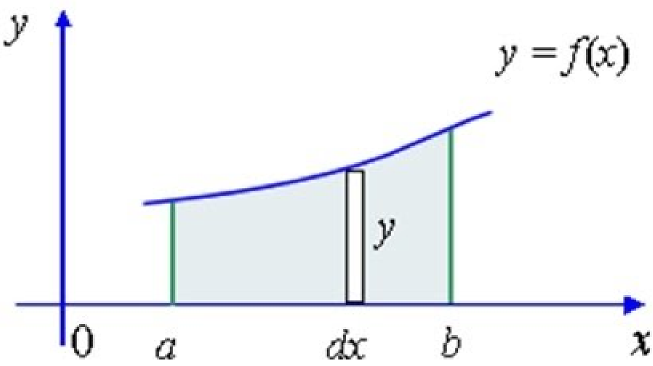


Figure 1: integrating under a graph

Let's say that f is the 1 (or more) dimensional continuous function over $\mathbb{R}$ (or $\mathbb{R}^n$), and we are trying to find the area under the graph. The first attempt at this might be the following (called Darboux sums): Given a partition the domain $P = \{x_1, x_2, ..., x_n\}$, we can make many rectangles which are all easy to measure. Then, as the partition gets finner (size$(P) \to 0$, $|P| \to \infty$), we will get a better an better approximation of the area.

If we take this approach, we will introduce some level of choice: there are many ways in which this partition can be formed, as well as a choice of the "height" of the rectangle. For example, you can choose the lowest or highest point:

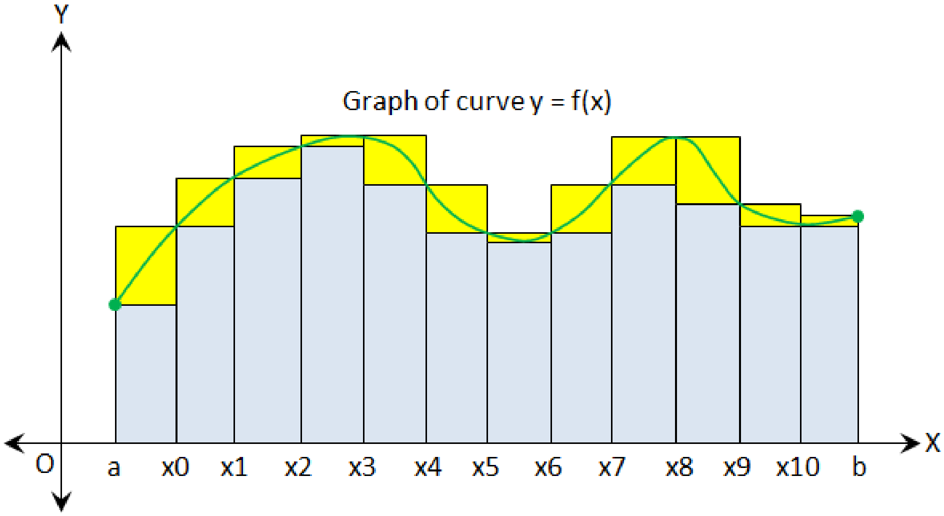


Figure 2: lower and upper sums

You can also choose any points in between, though what "in between" means starts to get hard to define in higher dimensional spaces where different orders can be put onto the "surfaces" (or hyper-surfaces) of the top of the rectangles (i.e. hyper-rectangles). However, maximal and minimal point is always well defined on these surfaces (or better, inf and sup always being well defined on these bounded surfaces), and even better, choosing the maximal point for each rectangle height upper-bounds all possible other choices of Darboux sums, and the minimal point for each rectangle's height lower-bounds all possible choices of Darboux sums. To be more precise, we'll replace the maximum and minimum condition with sup and inf, in which case the resulting sum is called the *Riemann sum.*

Intuitively, the upper and lower Riemann sums must match up and exist as size$(P) \to 0$ and $|P| \to \infty$. It would be strange if they don't match up, and so we can (as a start) define a function that is not "too weird" as one where the upper and lower Darboux sums match as size$(P) \to 0$, that is

$$\lim_{n\to\infty} \sum_{i=1}^{n-1} \left(\inf_{[x_i, x_i+1]} f(x) \right) \cdot (x_{i+1} - x_i) = \lim_{n\to\infty} \sum_{i=1}^{n-1} \left(\sup_{[x_i, x_i+1]} f(x) \right) \cdot (x_{i+1} - x_i)$$

If this is the case, f is said to be *Riemann integrable.* Fortunately, it turns out that if f is Riemann integrable on one partition, it is Riemann integrable on all partitions, and so we don't need to "remember" our initial choice of partition hen we say f is Riemann integrable (we don't say "f is Riemann integrable on partition p, while f is not Riemann integrable on partition q").

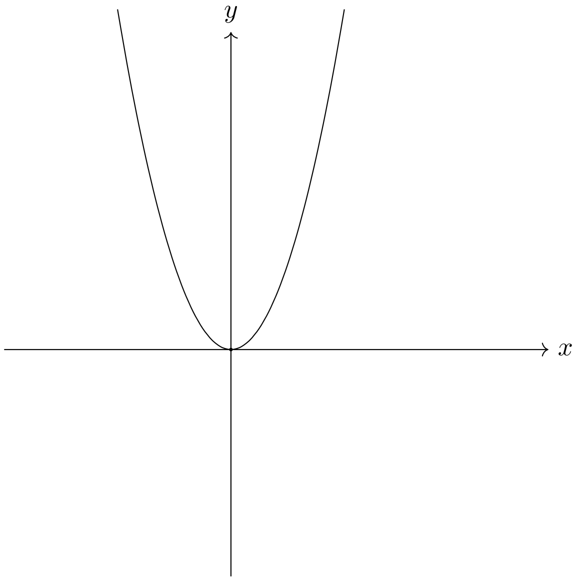
All continuous functions over a bounded domains will be integrable (so trigonometric functions, polynomials, log, $\sqrt{\quad}$ over the appropriate domain, and so forth). Since any geometric shape learnt in pre-university education can be bounded by a continuous function, we have just generalized the idea of measuring those areas[1]. In fact, being continuous *almost* encapsulates all possible functions.

[1]though we won't go into the details of finding those areas

To understand the generalization, consider

$$f_{\text{no }1} : [-1,1] \to \mathbb{R}, \qquad f_{\text{no }1}(x) \begin{cases} x^3 & x \neq 1 \\ 0 & x = 1 \end{cases}$$

Notice that there is a gap in the line



however, this gap has "zero measure", we would think that the result should be the same. And in fact it is! The function $f_{\text{no }1}$ is also Riemann integrable (as can be checked) and will have the same value as x^3 on $[-1,1]$. In general, if finitely (or even countably) many points are ommited, then f will still have the same measure. Since $f_{\text{no }1}$ is no longer continuous, but it as just off by some finite number of points, we say that f is continuous *almost everywhere* [2]. It should be reviewed that f is Riemann integrable if and only if f is continuous almost everywhere.

Stepping aside from measuring area under a graph for a moment, we turn to the study of another very common notion in analysis: limits. Limits are of fundamental importance in Analysis. They are the tool used to construct or define many different types of analytical objects (ex. $\mathbb{R}$, continuity, differentiability, Fourier series, sequences and series). More generally, the concept (and formalisation) of a limit is how mathematicians get to study infinites which are "well-defined". In my mind, limits are one of the distinguishing feature between algebra and analysis: many algebraic properties "break down" or are different when we try to bring in more than finitely many objects (ex. products of infinite groups cannot be direct products in **Grp** since $1 + 1 + \cdots$ is not defined (hence they are direct sums), the dimension of $\mathbb{Q}[x]$ and $\mathbb{Q}[[x]]$ as $\mathbb{Q}$-algebras differ, etc.). A good definition of limit is important to be able to define objects "at infinity", since if it is not well defined, then we can make the "result" non-unique.

Here is a silly example of that: what is $\infty - \infty$? Maybe we can say that the answer is 0? or ∞? If we are not careful with our definition, it can actually be any number. First, "notice" that $1 + 1/2 + 1/3 + \cdots$ "$=$" ∞, or more precisely, the series $S_n = \sum 1/n$ is unbounded from above. Since

[2] Later, almost every continuous will mean continuous up to a zero measure set once we properly define a measure

these two are "equal", we can treat them as the same. Then assuming the usual algebraic properties of distributivity work over countably many components, we hvae

$$(1 + 1/2 + 1/3 + \cdots) - (1 + 1/2 + 1/3 + \cdots) = (1 + 1/2 + 1/3 + \cdots) - 1 - 1/2 - 1/3 - \cdots$$

Now, pick any $x \in \mathbb{R}$. Then this sequence can be "rearranged" to converge to that number. Start by picking 1. If $x > 1$, pick $1/2$, if it's less pick -1. Keep doing this, and this will evidentially converge to x! Thus $\infty - \infty$ can be any real number we want!

This is why the idea of a limit is important: If the limit exists, we mean that there is a unique value for which the limit will converge to. This is why limits are used to define objects with some "countably infinite" properties [3].

One important consequence of limits being unique is when we use limits to construct new functions from a limit of functions! As review, you should check the construction of a *continuous but nowhere differentiable function*. Another example is with Fourier series and Fourier transformations which are used to encode and decode signals. Another example the definition of compact exhaustion for imperfect integrals. There are many more yet to come as well!

Now, returning to integrals for a moment, we can ask ourselves if integrals commute with limits, that is

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n$$

where the limit here is *pointwise limit*. Note that uniform limits (if it exist) will sort of commute:

$$\text{uniflim}_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n$$

However, point-wise limit doesn't generally do

Example 0.1: Limit Doesn't Commute

Take $\chi_{\mathbb{Q}} : [0, 1] \to [0, 1]$ to be the indicator function on $\mathbb{Q}$, that is

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \cap [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

Notice that $\chi_{\mathbb{Q}}$ is nowhere continuous, and hence is not Riemann integrable. We can see this more clearly by noticing that the upper Riemann sum is 1 and lower Riemann sum is 0. However

$$g_n(x) = \begin{cases} 1 & \text{if } \frac{p}{q},\ p, q \in \mathbb{N}, p \le n \\ 0 & \text{otherwise} \end{cases}$$

is integrable, and $g_n \to \chi_{\mathbb{Q}}$, that is $\lim_{n\to\infty} g_n = \chi_{\mathbb{Q}}$. Notice that each g_n are integrable, but $\chi_{\mathbb{Q}}$ is not, that is

$$\lim_{n\to\infty} \int g_n \ne \int \lim_{n\to\infty} g_n$$

Since the right hand side is not defined.

[3] Object with some "uncountably infinite" or larger properties use a generalisation of a limit called an *ultra-filter* which we don't need to study in this course

Thus, our notion of integrability does not work will with limits! Thus, since integrals gives us useful ways of analyzing our functions, and many functions are constructed as the limit of functions, we are going to spend lots of time upgrading our notion of integrability to make it commute (under appropriate circumstances) with limits.To upgrade the notion of integration, there must be some requirements we must keep in mind while finding a suitable replacement. In particular

1. The area under the graph should be the same as the Riemann Integral
2. Geometric intuition's should still be consistent, in particular
   (a) if $f \leq g$ then $\int f \leq \int g$
   (b) $\int af + bg = a \int f + b \int g$
3. A set A can be thought to be "integrable" if we define

$$\chi_A(x) \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

   and $\int \chi_A$ is defined.

This last criterion is actually very interesting as it allows us to generalize the question from finding the area's under graphs (or measuring the area under the graph) to figuring out how to measure sets in general! In fact, starting by defining how to measure sets and then defining how to integrate function turns out to be a fruitful order in which we can learn about this generalisation of integration. The reason this order is as actually an insight by Lebesgue: defining the properties of something being "measurable" will allow us to re-define the integral with more rigour. Even better, many different spaces can have different notions of a "measure" (for example, finite spaces where we measure points, or probability spaces where sets represent the probability of an event, or physics where we might want to measure the average density of a set), which the integral is the wrong concept (it's no longer strictly about "area"), but follows the same geometric intuitions as these other spaces (as we will show). Therefore, we will start by defining the spaces over which we will want to define a way to measure sets.

Chapter 1

Measures

As a start, we might want to define a reasonable definition of what it means to measure any subset of $\mathbb{R}^n$. Such a function would be of the form $\mu : P(\mathbb{R}^n) \to [0, \infty]$. If μ were to exist, we would want our common geometric intuitions to hold, that is:

1. If $E_1, E_2, ..., E_n, ...$ is a countable collection of pair-wise disjoint subsets, then
$$\mu\left(\bigcup_i E_i\right) = \sum_i \mu(E_i)$$
2. If E and F are sets such that there is a function that only translates, rotates, and reflects E to get to F (i.e. a rigid motion), then $\mu(E) = \mu(F)$
3. If C is a unit cube in $\mathbb{R}^n$, then $\mu(C) = 1$

However, these three axioms are inconsistent with one another! We have actually shown this by showing if μ is defined on $P(\mathbb{R}^n)$ (even when $n = 1$), then we can construct *Vitali sets* that will have to be both 0 and ∞ in measure – a contradiction.

Theorem 1.0.1: Vitali Sets Are Unmeasruable

Vitali sets are un-measurable

***Proof* :**
Take $[0, 1]$. Partition this set as follows:
$$x \sim y \iff x - y \in \mathbb{Q}$$
In a sense, this is the set of cosets of $\mathbb{R}/\mathbb{Q}$ restricted to $[0, 1]$; we will rely on this intuition to label the equivalence classes. Since $\mathbb{R}/\mathbb{Q}$ is the set of irrational numbers modulo rational numbers, let N_q be an equivalence class of contains x that is equivalent to the irrational number q. Notice that there are uncountably many such equivalence relations.

Now, pick one $p \in N_q$ for each equivalence class, and let N be the collection of all these p's, so N contains one point from each equivalence class (notice that the axiom of choice must be used to create a non-empty N). We'll show that N is unmeasurable.

First, for each rational number $r \in \mathbb{Q} \cap [0, 1)$, define:

$$N + r = \{x + r \mid x \in N \cap [0, 1 - r)\} \cup \{x + r \mid x \in N \cap (1 - r, 1)\}$$

Then clearly, $N + r \subseteq [0, 1)$ for each $r \in \mathbb{Q} \cap [0, 1)$, and furthermore, for each $x \in [0, 1)$ there exists an r such that $x \in N + r$. To show this, let's say $y \in N$ is the element that was picked from the equivalence class of of $[x]$. Then by our choice of y, $x \in N + r$ since $r = x - y$ if $x \geq y$, or $r = x - y + 1$ if $x < y$. Next, notice that the $N + r$ also forms a partition: if $(N + r) \cap (N + s) \neq \emptyset$ then $x - r$(or $x - r + 1$) would equal $x - s$ (or $x - s + 1$), but they belond to distinct equivalence classes, and so $(N + r) \cap (N + s) = \emptyset$.

We have no reached the point where we can declare our contradiction. If μ satisfies the three wanted properties, by (1) and (2):

$$\mu(N) = \mu(N \cap [0, 1 - r)) + \mu(N \cap (1 - r, 1)) = \mu(N + r)$$

for any $r \in \mathbb{Q} \cap [0, 1)$. Since $\mathbb{Q} \cap [0, 1)$ is countable and $[0, 1)$ is the disjoint union of the $N + r$'s:

$$1 = \mu([0, 1)) = \sum_{r \in \mathbb{Q} \cap [0,1)} \mu(N + r)$$

where $1 = \mu([0, 1))$ comes from (3). Now, we have two possibilities:

1. Either $\mu(N) = 0$, in which case each $\mu(N + r) = 0$ but then $0 = 1$ – a contradiction
2. Either $\mu(N) \neq 0$ (say k, so $\mu(N + r) = k$, but then $0 = \infty$ – a contradiction

In other words, no mater what $\mu(N)$ is, we will run into a contradiction!

As you can verify, this construction required that we take advantage of the countability condition (the 1st property). You might ask if weakening it to only finitely many will fix the issue, but sort of crazily it does not! In 1924, Banach and Tarski proved the following paradox now known as the *Banach Tarski Paradox*:

> Let U and V be open subsets of $\mathbb{R}^n$, $n \geq 3$. Then there exists a $k \in \mathbb{N}$ and subsets $E_1, E_2, ..., E_k$, $F_1, F_2, ..., F_k$ such that
>
> 1. all E_j's are disjoint and their union is U
> 2. all F_j's are disjoint and their union is V
> 3. There exists a rigid motion from E_j to F_j for $j = 1, ..., k$

This is kind of insane: this means that we can take a pea-size ball, and make into the size of the earth! This means that the notion of "geometry" is not actually a consistent concept in general set theory!

Fortunately for us, The sets E_j and F_j are very strange sets; most sets that are in fact measurable. To be more precise, it is consistent to say that the axiom of choice fails and that all sets are Lebesgue measurable (which will be the measure that will replace the Riemann integrability). For the purposes of this course, that means that any set we construct without the axiom of choice (and just doing normal ZF without any sneaky logical cheating) will be measurable. For more details, see `https://math.stackexchange.com/questions/1847052/most-functions-are-measurable`

1.1 Sigma-Algebras and Properties

Since the axiom of choice is the source of the problem of unmeasurability in the previous problems, we start by defining a space on which we avoid unmeasurable sets, namely, we want to limit ourselves to *countable* choices. The definition should be pretty general, as it should accommodate measures in more general spaces. It should also be "arithmetically expandable", as in we should be able to measure smaller parts of the space (perhaps in more convenient orientations) and add them all up. Thus, we define an *algebra*:

Definition 1.1.1: Algebra and σ-Algebra

Let X be an arbitrary set. Then let $\mathcal{A} \subseteq P(X)$ be a non-empty collection of subsets such that

1. for any finite set of sets $A_1, A_2, ..., A_n \in \mathcal{A}$, $\bigcup_i A_i \in \mathcal{A}$
2. For any $A \in \mathcal{A}$, $A^c \in \mathcal{A}$.

If it is closed under countable union, we call it a σ-algebra. We will usually denote σ-algebra on X as $\mathcal{M}(X)$.

The σ-algebra will be our natural space in which we will show in the next section we can define a "measure function" without the previous contradiction.

I like to think of the finite closure as a more general way of saying that a space is "closed under a function", the function here being $\cup$. In this way, it's sort of the most general algebraic structure. Adding the fact that it's closed under countable union brings us further way from the realm of algebra, and hence it's a σ-algebra. The term sigma is very common to mean countable, for example, a space σ-compactness if it is the countable union of compact subspaces.

Definition 1.1.2: Measurable Space

Let X be a set and $\mathcal{M} \subseteq P(X)$ be a σ-algebra. Then $(X, \mathcal{M})$ is called a *measurable space* and a set in $\mathcal{M}$ is called a *measurable set*

Proposition 1.1.1: Properties of σ-algebra

Some immediate properties of a σ-algebra:

1. It is closed under countable intersection, since $\bigcap_i X_i = (\bigcup_i X)^c$, and by difference, since $X \setminus Y = X \cap Y^c$.
2. $\emptyset$ and $X \in \mathcal{A}$, since for any $E \in \mathcal{A}$, $E \cap E^c = \emptyset \in \mathcal{A}$ and $E \cup E^c = X \in \mathcal{A}$
3. Any countable collection of subsets $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ can be turned into a countable disjoint union, where
$$F_k = E_k \setminus \left(\bigcup_{i=1}^{k-1} E_i \right) = E_k \cap \left(\bigcup_{i=1}^{k-1} E_i \right)^c$$
and by what we've just shown, every $F_k \in \mathcal{A}$, and by construction $\cup_i E_i = \cup_i F_i$. We will use this technique often very soon (note that not all sets are nonempty, that proof requires more set-theory manipulation).
4. Given any family of σ-algebras on X, their intersection is also a σ-algebra. Because of this, given $Y \subseteq X$, there is always a unique smallest σ-algebra $\mathcal{M}(Y)$ that contains Y (there is at least always one, $P(Y)$, and so this concept is well-defined). $\mathcal{M}(Y)$ is called the σ-algebra generated by Y.

A lemma we'll point out right now to simplify future proofs is the following:

Lemma 1.1.1

If $Y \subseteq \mathcal{M}(X)$, then $\mathcal{M}(Y) \subseteq \mathcal{M}(X)$

***Proof* :**
Since $\mathcal{M}(X)$ is a σ-algebra that contains Y, then it will contain all countable unions and compliment of Y, and hence it will contain $\mathcal{M}(Y)$ (i.e. proposition 1.1.1[4]).

Example 1.1: σ-Algebras

1. If X is any set, then $\{\emptyset, X\}$ and $P(X)$ are trivially σ-algebras. Once we define measurable function, these will be the initial and final object in measurable spaces.
2. If X is any set, then $P(X)$ is also a σ-algebra. Usually, we will use proposition 1.1.1[5] and other tools to construct a smaller σ-algebra to not work with $P(X)$ since it's too big for most cases.
3. If X is an uncountable set, define
$$\mathcal{A} = \{E \subseteq X \mid E \text{ is countable or } E^c \text{ is countable}\}$$
this set is called the σ-algebra of countable and co-countable sets. It is useful when we want to make sure that the countable union of sets is still countable (or it's compliment is) [a]

[a]Sneakily, we use the axiom of choice for countably many choices to prove this is true

Another common σ-algebra generated on a set X is related to a topology on X, that is, if $\mathcal{T}(X)$ is a topology, of X, we will define the following σ-algebra:

Definition 1.1.3: Borel σ-Algebra

Let X be a set and $\mathcal{T}(X)$ be a topology on X. Define the *Borel σ-algebra* to be

$$\mathcal{B}_X := \mathcal{M}(\{U \subseteq X \mid U \text{ is an open set contained in } X\}$$

The elements of $\mathcal{B}_X$ are called *Borel sets*

The set $\mathcal{B}_X$ will include the open sets, closed sets, countable union and intersection of closed sets, and the countable union of countable intersections of open/closed sets, et cetera.

Since the countable intersection of open sets and countable union of closed sets are not necessarily open or closed sets respectively[1], but are members in $\mathcal{B}_X$, we will give them names: A countable intersection of open sets is called a G_δ set , and a countable union of closed sets is called a F_σ set (F for *fermé* in french, G because it's after F in the alphabet). The σ is for *Summe* (sum in German) since it's a union and the δ is for *Durchschnitt* (intersection in German) since it's an intersection. We can continue this pattern and define $G_{\delta\sigma}$ to be the countable union of countable intersection of open sets, $F_{\sigma\delta}$ to be the countable intersection of countable union, and so and so on and so forth for any length of interchanging σ's and δ's. In general, this "chain" of unions and intersections does not need to terminate, and a $F_{\sigma\delta}$ (as a simple example for a chain of length 2, $\mathbb{R} \setminus \mathbb{Q}$ is a $F_{\sigma\delta}$ set, but not a F_σ set; use Baire's Categories theorem). There will be cases in which it will terminate (in fact, the majority of measure's we'll work with will have this terminate after at just F_σ or G_δ "up to a zero set", see theorem 1.4.3, and example ref:HERE for a measure that does not terminate)

One of the most common Borel sets we will be working with is $\mathcal{B}_\mathbb{R}$ with the usual euclidean topology on $\mathbb{R}$. It is useful to have an of different generating sets for $\mathcal{B}_\mathbb{R}$:

Proposition 1.1.2: Generators For $\mathcal{B}_X$

Let $\mathbb{R}$ be the real numbers and $\mathcal{B}_\mathbb{R}$ be the Borel σ-aglebra on $\mathbb{R}$ given the standard euclidean topology $\mathcal{T}(\mathbb{R})$. Then the following generate $\mathcal{B}_\mathbb{R}$:

1. The set of open intervals
2. the set of closed intervals
3. the set of half-open intervals
4. the set of open rays
5. the set of closed rays

Proof **:**
These proofs are quite easy to prove, and should be done as an exercise. A fact that you might need to recall is that every open set in $\mathbb{R}$ can be written as a countable union of open intervals (hint: $\mathbb{R}$ has a countable basis).

[1] it is possible that the countable intersection/union of open/closed sets are open/closed, and so the word "necessarily" is required

Products and σ-Algebras

Recall that the product between two sets is defined as $A \times B := \{(a, b) \mid a \in A, b \in B\}$, or more generally $\{X_\alpha\}_{\alpha\in A}$ for a collection of non-empty sets X_α[2], with $X = \prod_{\alpha\in A} X_\alpha$,. As we know from Category Theory, the product comes equipped with set functions π_α that satisfy the universal property of products (or more generally, this is a limit over the discrete category). Though we have yet to give a category where σ-algebras are the objects (we need to define what the morphisms are[3]), we can already construct our set $\prod_\alpha X_\alpha$ that will be the product of the X_α in the category of measurable spaces.

Definition 1.1.4: Product σ-Algebra

Let $\{X_\alpha\}_{\alpha\in A}$ be a collection of non-empty sets and $X = \prod_{\alpha\in A} X_\alpha$. Let

$$\bigotimes_{\alpha\in A} M_\alpha = \mathcal{M}(\{\pi_\alpha^{-1}(E_\alpha) \mid E_\alpha \in M_\alpha,\ \alpha \in A\})$$

to be the *product σ-Algebra* on $\{M_\alpha\}_{\alpha\in A}$, i.e., it is the σ-algebra generated by all of the pre-images of the sets in M_α.

It is not difficult to check that the product σ-algebra is indeed a σ-algebra. If $|A|$ is countable we usually write $\bigotimes_{i=1}^\infty M_i$, and if $|A|$ is finite, we can write $\bigotimes_{i=1}^n M_i$ or $M_1 \bigotimes M_2 \bigotimes \cdots \bigotimes M_n$. Notice that since we allow for countable unions, it is not so easy to represent $\bigotimes_{\alpha\in A} M_\alpha$ as the product of the elements of M_α, (just like for Groups in the category of **Grp**), and hence why we do not use the $\prod$ or $\bigoplus$ notation. On the other hand, the fact that only a countable union of elements is permitted will still limit the product in the number of non-trivial components, hence justifying the $\bigotimes$ notation instead of $\prod$. In fact, it is the same as $\bigoplus$, but instead of "all but finitely many", we have "all but countably many":

Proposition 1.1.3: Countable Product σ-algebra

Let $M_n = \mathcal{M}(\mathcal{E}_n)$ be a σ-algebra for $n \in \mathbb{N}$ where $\mathcal{E}_n \subseteq P(X_n)$. If $X_n \in \mathcal{E}_n$, then

$$\bigotimes_{n=1}^\infty M_n = \mathcal{M}\left(\left\{\prod_{i=1}^n E_i \,\middle|\, E_i \in M_i\right\}\right)$$

***Proof*:**
The $\supseteq$ direction is easy (Take $\bigcap_{i=1}^n \pi_i(E_i)$ for appropriate values). For $\subseteq$, Since $X_i \in M_i$, we have that every element is inside the right hand side. Make sure you see why $X_n \in \mathcal{E}_n$ makes a difference. By lemma 1.1.1, the result follows.

We can further simplify our representation of $\bigotimes_{\alpha\in S} M_\alpha$ given a generating set for each M_α (such a set always exists, namely M_α itself is always a generating set)

[2]If they were empty, then the product would be empty
[3]this is done in chapter 2

Proposition 1.1.4: Generating set for Product σ-Algebra

Let $\{M_\alpha\}_{\alpha\in A}$ be a non-empty collection of σ-algebras, and let $\mathcal{E}_\alpha \subseteq P(X_\alpha)$ be a generating set for M_α. If $\mathcal{F} = \cup_{\alpha\in A}\mathcal{E}_\alpha$, then:

$$\bigotimes_{\alpha\in A} M_\alpha = \mathcal{M}(\{\pi_\alpha^{-1}(E_\alpha) \mid E_\alpha \in \mathcal{E}_\alpha,\ \alpha\in A\}) = \mathcal{M}(\mathcal{F})$$

Furthermore, if $|A|$ is countable and $X_\alpha \in \mathcal{E}_\alpha$ for each $\alpha \in A$, then

$$\bigotimes_{\alpha\in A} M_\alpha = \mathcal{M}\left(\left\{\prod_{\alpha\in A} E_\alpha \,\middle|\, E_\alpha \in \mathcal{E}_\alpha,\ \alpha\in A\right\}\right)$$

Proof **:**
Dealing with the first case, the $\supseteq$ is clear since $E_\alpha \in \mathcal{E}_\alpha$ implies $E_\alpha \in M_\alpha$, so $E_\alpha \in \bigotimes_{\alpha\in A} M_\alpha$. For the $\subseteq$ direction, If $E_\alpha \notin \mathcal{E}_\alpha$, then since $\mathcal{M}(\mathcal{E}_\alpha) = M_\alpha$, some countable unions and intersections of elements of $\mathcal{E}_\alpha$ equal E_α, proving the $\subseteq$ direction.

The countable case follows proposition 1.1.3

Borel Spaces and Products

A topological space of much study in Analysis is metric spaces. As a reminder, if $\{X_\alpha\}_{i=1}^n$ is a collection of metric spaces, then $X = \prod_{i=1}^n X_i$ has the product metric which is the generalization of the euclidean metric. As was seen in a topology class, the following metrics on a product space all define the same product topology (for notational simplicity, let's work with two metric spaces X and Y and consider $X \times Y$):

1. $d((x_1, x_2), (y_1, y_2)) = \max\{d_X(x_1, y_1), d_Y(x_2, y_2)\}$
2. $d((x_1, x_2), (y_1, y_2)) = \sqrt{d_X(x_1, y_1)^2 + d_Y(x_2, y_2)^2}$
3. $d((x_1, x_2), (y_1, y_2)) = d_X(x_1, y_1) + d_Y(x_2, y_2)$

We can ask about establishing Borel sets $\mathcal{B}_X$ on $X = \prod_{i=1}^n X_i$ with X having the product measure. It is tempting to say that this naturally splits into $\bigotimes \mathcal{B}_{X_i}$, however, this is not always the case! This is mainly due to the fact that not all topological spaces have a countable basis, and since we're allowing countable intersection of open sets, this will actually create more sets than the if we take the product of the Borel spaces:

Proposition 1.1.5: Borel sets on product Metric Space

Let $X_1, ..., X_n$ be metric spaces, and $X = \prod_{i=1}^n X_i$ be the product metric space. Then

$$\bigotimes_{i=1}^{n} \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$$

If the spaces X_i are separable (i.e. has a countable dense subset, and since it's a metric space it means there is a countable basis), then

$$\bigotimes_{i=1}^{n} \mathcal{B}_{X_i} = \mathcal{B}_X$$

Proof **:**
The proof of the $\subseteq$ direction rather simple. By proposition 1.1.4, $\bigotimes_{i=1}^n \mathcal{B}_{X_i}$ is generated by $\mathcal{E} = \left\{\pi_i^{-1}(U_j) \;\middle|\; U_j \text{ is open in } X_j,\ 1 \leq j \leq n\right\}$. Since the sets of $\mathcal{E}$ are open in X, by lemma 1.1.1, $\bigotimes_{i=1}^n \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$.
For the converse, since each X_j is separable, each have a countable dense subset. Since X_j is a metric space, we can form a countable basis with open balls with rational radii around each points in the countable dense subsets.

As a consequence, $\bigotimes_{i=1}^n \mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^n}$, since $\mathbb{R}$ has a countable dense subset.

Elementary Sets

This is a technical result needed for later (maybe introduce it as exercises?)

Definition 1.1.5: Elementary Family

Let X be a set and $\mathcal{E} \subseteq P(X)$. Then $\mathcal{E}$ is called an *elementary family* if

1. $\emptyset \in \mathcal{E}$
2. If $E, F \in \mathcal{E}$ then $E \cap F \in \mathcal{E}$
3. If $E \in \mathcal{E}$, then E^c is a finite disjoint union of elements of $\mathcal{E}$

Proposition 1.1.6: Elementary Families and Algebra

Let $\mathcal{E}$ be an elementary family. Then the collection of finite disjoint unions of element of $\mathcal{E}$ forms an algebra

Proof **:**
(TBD) Let $U, V \in \mathcal{E}$, and consider $V^c = \bigsqcup_{j=1}^{J} F_j$ (for disjoint $F_j \in \mathcal{E}$). Then

$$U \setminus V = U \cap V^c = \bigsqcup_{j=1}^{J} (U \cap F_j)$$

notice that each $U \cap F_j \in \mathcal{E}$. So

$$U \cup V = (U \setminus V) \cup V = \bigsqcup_{j=1}^{J} (U \cap F_j) \bigsqcup V$$

where $\bigsqcup_{j=1}^{J} (U \cap F_j) \in A$
Continuing inductively,

(also, remember that separable means countable dense subset) Once the course starts

1. There is an interseting example ith logical statements!!

Remark. If A is a set. and $\varphi(x)$ is a logical statement about x. Then $\{x : \varphi(x) \text{ holds and } x \in A\}$ is a set.

Let $\mathcal{E} \subset P(X)$, define:

$$\mathcal{E}_1 = \left\{ \bigcup_{i=1}^{\infty} \bigcap_{j=1}^{\infty} E_{i,j} \ , \ \text{either } E_{i,j} \in \mathcal{E} \text{ or } E_{i,j}^c \in \mathcal{E} \right\}$$

$$\mathcal{E}_2 = \left\{ \bigcup_{i=1}^{\infty} \bigcap_{j=1}^{\infty} E_{i,j} \ , \ \text{either } E_{i,j} \in \mathcal{E}_1 \text{ or } E_{i,j}^c \in \mathcal{E}_1 \right\}$$

etc.

$\bigcup \mathcal{E}_n$ is in general not a σ-alg. and there is no constructive way to understand $\mathcal{M}(\mathcal{E})$.

Figure 1.1: fascinating!

1.2 Measure

We can finally define the function which will give us an idea of the "size" of sets:

Definition 1.2.1: Measure

Let X be a set along with a σ-algebra $\mathcal{M}$. A function $\mu : \mathcal{M} \to [0, \infty]$ is called a *measure* on $\mathcal{M}$ (or $(X, \mathcal{M})$ or even X if the measure is clear from context) if

1. $\mu(\emptyset) = 0$
2. **(Countably Additivity)** If $\{E_i\}_{i=1}^{\infty}$ is a sequence of disjoint sets, then $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$

Given that $(X, \mathcal{M})$ is a measurable space, along with a μ we can say the space has a measure:

Definition 1.2.2: Measure Space

Let $(X, \mathcal{M})$ be a measurable space. Then if μ is a measure on X, then $(X, \mathcal{M}, \mu)$ is a *measure space*

If the 2nd condition of a measure was limited to a finite sequence (or more generally, all but countably many of the sets in the sequence or nonempty), then we would say that it respects finite additivity. If μ respects finite but not countable additivity, μ is called a *finite additive measure*.

Most measure's that we will work with have some finiteness condition associated with it:

Definition 1.2.3: Finiteness Condition on μ

Let $(X, \mathcal{M}, \mu)$ be a measure space.

If $(X) < \infty$, then μ is called a *finite measure*

if there exists sets $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$ such that $\bigcup_i^{\infty} E_i = X$ and $\mu(E_i) < \infty$ for all E_i, then μ is called a *σ-finite measure*. More generally, if $E = \bigcup_{i=1}^{\infty} E_i$ for some sequence of E_i, then E is called σ-finite on μ (some say that E is said to be σ-finite on μ)

if for for all $E \in \mathcal{M}$ such that $\mu(E) = \infty$, there exists a $D \subsetneq E$, $D \in \mathcal{M}$ such that $\mu(D) < \infty$, then μ is called a *semi-finite measure*

If $\mu(X) < \infty$, then this implies for all $S \in \mathcal{M}$, $\mu(S) < \infty$ (as can quickly be checked). Naturally, all finite measure's are σ-finite, and all σ-finite measure are semi-finite, but the converses are not true (as we'll explore in the following examples). Note too that most measure's we will be working over will be σ-finite (think of $\mathbb{R}$ with the intuitive notion of length of an interval).

Example 1.2: Measures

1. Let X be nonempty, $\mathcal{M} = P(X)$, and $f : X \to [0, \infty]$. Then we can define

$$\mu_f(E) = \sum_{e \in E} f(e)$$

   where the sum can be infinite (the domain of f is the positive part of the extended real numbers). Notice that f is semifinite if and only if $f(x) < \infty$ for all $x \in X$, and μ is σ-finite if and only if μ is semi-finite and $\{x \mid f(x) > 0\}$ if countable.

   If $f(x) = 1$ for all $x \in X$, then μ_f is called the *counting measure*. On the other hand, if for some $x_0 \in X$, $f(x_0) = 1$ and $f(x) = 0$ if $x \neq x_0$, then μ_f is called a *point mass* or *Dirac measure* at x_0

2. Let X be an uncountable set an $\mathcal{M}$ be the σ-algebra of countable or co-countable sets (recall example 1.1). Define

$$\mu(E) = \begin{cases} 0 & E \text{ is countable} \\ 1 & E \text{ is co-countable} \end{cases}$$

   Then μ is a measure!

3. An example of an finite additive measure that is not a measure is the following: let X be

some infinite set and $\mathcal{M} = P(X)$. Define μ to be

$$\mu(E) = \begin{cases} 0 & E \text{ is finite} \\ \infty & E \text{ is infinite} \end{cases}$$

Then μ is an finite additive measure, but not a measure

Proposition 1.2.1: Properties of Measures

Let $(X, \mathcal{M}, \mu)$ be a measure space. Then

1. **(Monotonicity)** Let $E, F \in \mathcal{M}$ and $E \subseteq F$. Then $\mu(E) \leq \mu(F)$
2. **(Subadditivity)** if $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$, then $\mu\left(\cup_i \mu(E_i)\right) \leq \sum_i \mu(E_i)$
3. **(Continuity from Below)** if $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$, and $E_1 \subseteq E_2 \subseteq \cdots$, then
$$\mu\left(\cup_i E_i\right) = \lim_{i \to \infty} \mu(E_i)$$
4. **Continuity from Above** if $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$ and $E_1 \supseteq E_2 \supseteq$ *and* $\mu(E_1) < \infty$. Then
$$\mu\left(\cap_i E_i\right) = \lim_{i \to \infty} \mu(E_i)$$

Proof **:**
here. Do them as exercises!

Notice that for (d), we can instead limit it to $\mu(E_i) < \infty$ for some i, and the proof remains the same. We still require that some i exists such that $\mu(E_i) < \infty$, for it's possible that all $\mu(E_i)$ is infinite, but $\mu(\cap_i E_i) < 0$. For example, take the measurable space $(\mathbb{N}, P(\mathbb{N}))$ with the counting measure. Then the sets

$$E_i = \{n \in \mathbb{N} \mid n \geq i\}$$

then every E_i is infinite but their intersection is empty and thus has measure 0 in the counting measure.

zero measure sets

Frequently, there are sets in our measure that have a "trivial size" in the current measure. This is akin to when we integrate, the border has no effect on the value of the integral. Such sets are said to be zero sets or null set:

Definition 1.2.4: Null Set

Let $(X, \mathcal{M}, \mu)$ be a measure space and $E \in \mathcal{M}$. Then if $\mu(E) = 0$, E is said to be a *null set* or *zero set*.

By countably subadditivity, the countable union of null sets is a null set, hence null sets are "invariant" under countable union. Due to this, we can work with measurable sets in $\mathcal{M}$ *up to null sets*. We shall often state in proofs that a statement is true *almost everywhere* (often abbreviated to a.e.) if it is true on the sets except for the null sets.

If we take $E \in \mathcal{M}$ such that $\mu(E) = 0$, then by monotonicity, if $F \subseteq E$ the $\mu(F) = 0$, *provided that* $F \in \mathcal{M}$. This is not always the case:

Example 1.3: Not Complete Space

Partition $[0, 1]$ into 10 (really any number) of subintervals, $\left[\frac{i}{n}, \frac{i+1}{n}\right]$ for $0 \leq i \leq 9$, and let each of their measures be 0. Take $\mathcal{M}$ to be the σ-algebra defined on this set. Then is clearly an uncountable number of sets that do not have a measure defined on them that are subset of zero-measure sets.

However, adding all of these F's that are subsets of null sets does not break our measure. A measure who's domain includes all subsets of null sets is a *complete measure* (i.e. the σ-algebra is closed under subsets of null sets). Proving that this is still a measure is useful to eliminate future "trivial obstructions" to our proofs:

Theorem 1.2.1: Completing a Measure

Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $\mathcal{N} = \{N \in \mathcal{M} \mid \mu(N) = 0\}$ and $\overline{\mathcal{M}} = \{E \cup F \mid E \in \mathcal{M} \text{ and } F \subseteq N \text{ for some } N \in \mathcal{N}\}$. Then $\overline{\mathcal{M}}$ is also a σ-algebra, and there is a *unique* extension $\overline{\mu}$ of μ that is a complete measure on $\overline{\mathcal{M}}$ that restricts to the measure of μ on $\mathcal{M}$

***Proof* :**

First of all, since $\mathcal{M}$ and $\mathcal{N}$ are both closed under countable union, so $\overline{\mathcal{M}}$ is closed under countable union. To show it's closed under compliments, consider some $E \cup F \in \overline{\mathcal{M}}$ ($E \in \mathcal{M}$, $F \subseteq N \in \mathcal{N}$). Without loss of generality, assume $E \cap N = \emptyset$ (if not, take $F \setminus E$ and $N \setminus E$ and continue the proof, then this will have he same value as our original $E \cap N$). Then

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

So we get

$$(E \cup F)^c = (E \cup N)^c \cup (N \setminus F)$$

Since $E \cup N \in \mathcal{M}$, $(E \cup N)^c \in \mathcal{M}$. Since $N \setminus F \in \mathcal{N}$, Since it's closed under union, $(E \cup F)^c \in \overline{\mathcal{M}}$. Thus $\overline{\mathcal{M}}$ is a σ-algebra.

Since $\overline{\mathcal{M}}$ is a σ-algebra, we can try to define a measure, and indeed

$$\overline{\mu}(E \cup F) = \mu(E)$$

is a well-defined measure. If

$$E_1 \cup F_1 = E_2 \cup F_2$$

Then $E_1 \subseteq E_2 \cup F_2$, and so by monotonicity:

$$\mu(E_1) \leq \mu(E_2 \cup F_2) \leq (E_2 \cup N_2) = \mu(E_2) + \mu(N_2) = \mu(E_2)$$

Similarly, $\mu(E_1) \geq \mu(E_2)$. Thus $\overline{\mu}(E_1 \cup F_1) = \overline{\mu}(E_2 \cup F_2)$, and so $\overline{\mu}$ is well-defined. Since $\overline{\mu}$ has all subsets of zero-sets in its domain, it is a complete measure.

Furthermore, if $\overline{\nu}$ was another complete measure on $\mathcal{M}$, then it is easy to see that $\overline{\nu} = \overline{\mu}$, (since if it were different, then some subset of a zero-set must have non-zero measure, a contradiction to monotonicity) showing that $\overline{\mu}$ is unique.

Definition 1.2.5: Completion of Measure

Let $(X, \mathcal{M}, \mu)$ be a measurable space. Then $\overline{\mathcal{M}}$ and $\overline{\mu}$ is called the *completion of* $\mathcal{M}$ *with respect to* μ and the *completion of* μ respectively

1.3 Outer Measure

We now move on to constructing a particular type of measure which is draws inspiration from how we measured area's when we calculated Riemann Integrals. In the abstract, when we tried defining a Jordan measure (or content[4]), we approximated the inner area of a shape E in $\mathbb{R}^n$ by the limit of sums of rectangles contained in E, and the outer-area of E in $\mathbb{R}^n$ by the limit of the sums of rectangles containing E. The two numbers are called the inner and outer area of E. If the two matched, we would call the result the area of E. This intuition can be generalized to the countable case, except we can replace rectangles with some other "basic sets" that we might want that form an algebra. Furthermore, the intuition of having the "inner area" and "outer-area' is actually a bit restricting to work with in the general measure case, and so that condition will be replaced with a different measureability condition we will soon introduce.

For the sake of generality, we will start by definition a more general notion of the measure called the outer-measure, and then build-up on how form the outer-measure, we can get a measure that satisfies this generalization of Riemann integration:

Definition 1.3.1: Outer Measure

A function $\mu^* : P(X) \to [0, \infty]$ is called an *outer measure* if

1. $\mu^*(0) = \emptyset$
2. if $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$
3. $\mu^* \left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$ (even if the sets are disjoint!)[a]

[a] We will return to see when equality holds in definition 1.3.2

Note that we have not yet proven that an outer-measure is a measure; in particular notice that the domain is the powerset. What we will do is define a generalization of the Riemann integral that will be our "base-line" for the outer-measure:

[4] since "Jordan measures" only allow finite additivity, some don't like to use the word measure and opt for the word content. In this book, we would say *finite additive measure* as has already been stated earlier

Proposition 1.3.1: Outer Measure Construction

Let $\mathcal{E} \subseteq P(X)$ where $\emptyset, X \in \mathcal{E}$, and let $\rho : \mathcal{E} \to [0,1]$ where $\rho(\emptyset) = 0$. For any $A \subseteq X$, define

$$\mu^*(A) = \inf \left\{ \sum_1^\infty \rho(E_j) \;\middle|\; E_j \in \mathcal{E} \text{ and } A \subseteq \bigcup_1^\infty E_j \right\}$$

Then μ^* is an outer measure

***Proof* :**
First, this measure is well-defined since for any $A \subseteq X$, If we take $E_i = X$, then $A \subseteq \cup_i E_i$. Next, we verify the 3 properties

1. Clearly, $\mu^*(\emptyset) = 0$ by taking $E_i = \emptyset$ for all i
2. If $A \subseteq B$, then let $\mathcal{A}$ and $\mathcal{B}$ sets that correspond to $\mu^*(A)$ and $\mu^*(B)$. Then since all of the elements of $\mathcal{A} \supseteq \mathcal{B}$, by the property of the infimum $\inf \mathcal{A} \leq \inf \mathcal{B}$, which transferring notations gives us $\mu^*(A) \leq \mu^*(B)$
3. Let $\{E_i\}_{i=1}^\infty \subseteq P(X)$ be a collection of (not necessarily disjoint) subsets. Let $\epsilon > 0$. Consider $\mu^*(E_k)$ for some $E_k \in \{E_i\}_{i=1}^\infty$. By the property of the infimum, there must exist some $\{F_j^k\}_{j=1}^\infty$ such that $E_i \subseteq \cup_j F_j^k$ where $\sum_{j=1}^\infty \rho(F_j^k) \leq \mu^*(E_k) + \epsilon 2^{-k}$.

   Now, let $E = \cup_i E_i$. By construction, $E \subseteq \cup_{i,k=1}^\infty F_j^k$, and

   $$\sum_{j,k} \rho(F_j^k) \leq \mu^*(A) + \epsilon$$

   Since ϵ can be arbitrarily small, this completes the proof.

As mentioned before, an outer-measure is not necessarily a measure. What we will do to allow this to be a measure is single out sets that have the following property

Definition 1.3.2: μ^*-measureable

Let $E \subseteq X$ be a set and μ^* be an outer measure on X. Then E is called μ^**-measurable* if

$$\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c) \quad \forall S \subseteq X$$

In other word, if E can naturally "split" a set S into a part that's "inside" E or "outside" E. The set $S \subseteq X$ is called the *test set*.

By definition of the outer measure, since $S = (S \cap E) \cup (S \cap E^c)$:

$$\mu^*(S) \leq \mu^*(S \cap E) + \mu^*(S \cap E^c) \quad \forall S \subseteq X$$

Therefore, what is more important is to check the $\geq$ direction. Furthermore, if $\mu^*(A) = \infty$. Then $\geq$ holds trivially, so to show a set is μ^* measurable, it is equivalent to show that for all finite measure sets A, $\mu^*(S) \geq \mu(S \cap E) + \mu(S \cap E^c)$.

As mentioned in the definition, this gives us a sense of being able to measure the "inside" and "outside" of a set. If $E \subseteq A$, that is the μ^*-measurable set belong in a better set, then $\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \setminus E)$ which can be thought of as saying that we are measuring S from the "inside" of E and the "outisde" of E, and getting the same result. In fact, if $\mu^*(X) < \infty$, then we can define $\mu_*(E) = \mu^*(X) - \mu(E^c)$, and have that E is μ^*-measurable if and only if $\mu_*(E) = \mu^*(E)$ aligning with our intuition for inside equalling outside (see exercise ref:HERE). We do not do this method since it requires some finiteness condition on μ. We can conversely define μ_* in terms of supremums and need that E contains the collection of $\cup_i A_i$, and that the elements of $\{A_i\}_{i=1}^{\infty}$ must be disjoint (see exercise ref:HERE).

The big thing that is in the way of saying that an outer measure is a measure is that its domain is $P(X)$, which can cause problems when X is uncountable. However, all sets that satisfy μ^*-measureability turn out form a σ-algebra and make μ^* a complete measure!

Theorem 1.3.1: Carathéodory's Theorem

Let μ^* be an outer measure on X, and let $\mathcal{M}$ be the collection of μ^*-measurable sets. Then $\mathcal{M}$ forms a σ-algebra, and the restriction of μ^* to $\mathcal{M}$ is a complete measure

Notice that in the process of defining this measure, we started with the powerset, and then simply choose all sets that restrict (which we will show forms a σ-algebra). In this way proving a function is an outer-measure is a little less finicky since we don't need to be so careful in defining over which σ-algebra we want to define it over: the σ-algebra comes with μ^*

***Proof* :**
Let $\mathcal{M}$ be the collection of μ^*-measurable sets. Clearly, $\emptyset \in \mathcal{M}$ since for all $S \subseteq X$:

$$\mu^*(S \cap \emptyset) + \mu^*(S \cap X) = 0 + \mu^*(S)$$

implying μ^*-measureability. Next, $\mathcal{M}$ is closed under compliments, since if $E \in \mathcal{M}$, then

$$\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c) = \mu^*(S \cap E^c) + \mu^*(S \cap E)$$

which also shows that $X \in \mathcal{M}$. Finally, we need to show that $\mathcal{M}$ is closed under countable unions. We'll start by showing finite unions, which then means it suffices to show the union of two sets is closed (due to induction). Let $A, B \in \mathcal{M}$. As we remarked, it suffices to show that for any $E \subseteq X$ such that $\mu^*(E) < \infty$ that $\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$. Since A and B are μ^* measurable, we have

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) && E \text{ is the test set} \\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) \\ &\quad + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) && E \cap A \text{ and } E \cap A^c \text{ are the test sets} \end{aligned}$$

Then notice we can write

$$A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$$

Thus, using subadditivity, we can re-write 3 of the terms in the previous expression as:

$$\mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) \geq \mu^*(E \cap A \cup B)$$

Since $E \cap A^c \cap B^c = E \cap (A \cup B)^c$, we get:

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$

showing that $\mathcal{M}$ is closed under finite union. Since $\mathcal{M}$ is closed under compliment and contains $\emptyset$, $\mathcal{M}$ is an algebra. Even better, If we let A and B be disjoint, and let $A \cup B$ be the test set, then since A is μ^* measurable:

$$\mu^*(A \cup B) = \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) = \mu^*(A) + \mu^*(B)$$

showing that $\mu^*|_{\mathcal{M}}$ is a finitely additive measure.

Next, we must show that $\mathcal{M}$ is a σ-algebra, which as a consequence will show that $\mu^*|_{\mathcal{M}}$ is a measure. Let $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$. To simplify our task, re-construct $\{E_i\}_{i=1}^{\infty}$ to be the set of disjoint sets $\{D_i\}_{i=1}^{\infty}$. Let $B_n = \bigcup_{i=1}^{n} D_i$ and $B = \bigcup_{i=1}^{\infty} D_i$. Our goal is to show that for any $E \subseteq X$ such that $\mu^*(E) < \infty$ that $\mu^*(E) \geq \mu^*(E \cap B) + \mu^*(E \cap B^c)$. We will start by considering $\mu^*(E \cap B_n)$:

$$\begin{aligned} \mu^*(E \cap B_n) &= \mu^*(E \cap B_n \cap D_n) + \mu^*(E \cap B_n \cap D_n^c) && E \cap B_n \text{ is the test set} \\ &= \mu^*(E \cap D_n) + \mu^*(E \cap B_{n-1}) \end{aligned}$$

Notice that on the right hand side we have diminished to the case of $E \cap B_{n-1}$ from $E \cap B_n$. Using induction, we can see that the expression simplifies to

$$\mu^*(E \cap B_n) = \sum_{i=1}^{n} \mu^*(E \cap D_i)$$

Continuing, since B_n is measurable

$$\mu^*(E) = \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c) \overset{!}{\geq} \mu^*(E \cap B_n) + \mu^*(E \cap B^c) = \sum_{i=1}^{n} \mu^*(E \cap D_i) + \mu^*(E \cap B_n^c)$$

where $\overset{!}{\geq}$ comes from monotonicity after switching B_n^c to B^c. Since this equality holds true for all n, we get:

$$\begin{aligned} \mu^*(E) &\geq \sum_{i=1}^{\infty} \mu^*(E \cap D_i) + \mu^*(E \cap B^c) \\ &\geq \mu^*\left(\bigcup_i E \cap D_i\right) + \mu^*(E \cap B^c) && \text{subadditivity} \\ &= \mu^*(E \cap B) + \mu^*(E \cap B^c) && \text{by construction of } E \cap B \\ &\geq \mu^*(E) && \text{subadditivity} \end{aligned}$$

Since we got $\mu^*(E) \geq \mu^*(E)$, we see that in fact all inequalities are equalities, getting us that

$$\mu^*(E) = \mu^*(E \cap B) + \mu^*(E \cap B^c)$$

showing us that $B \in \mathcal{M}$. Furthermore, replacing $E = B$ in the previous proof shows us that

$$\mu^*(B) = \sum_{i=1}^{\infty} \mu^*(D_i)$$

giving us that μ^* is a measure on $\mathcal{M}$.

Finally, we must show that $\mathcal{M}$ is complete, that is, if $E \in \mathcal{M}$ such that $\mu^*(E) = 0$, then all subsets of E are in $\mathcal{M}$. Notice that since μ^* is defiend on all of $P(X)$, it is equivalent to show that if $\mu^*(A) = 0$, then $A \in \mathcal{M}$. Then for any test set E, by monotonicity:

$$\mu^*(E) \le \mu^*(E \cap A) + \mu^*(E \cap A^c) = 0 + \mu^*(E \cap A^c) \le \mu^*(E)$$

meaning equality holds, showing that $A \in \mathcal{M}$. Thus, $\mu^*|_{\mathcal{M}}$ is a complete measure, as we sought to show.

Carathéodory's Theorem is quite useful being able to state the existence of a measure given an outer-measure, but it's not good to construct a measure (for example, can you tell me a non-trivial element in $\mathcal{M}$ without directly appealing to the definition?). The following concept is used to more systematically construct the set $\mathcal{M}$ and the measure μ induced by μ^*:

Definition 1.3.3: Premeasure

Let $(X, \mathcal{A})$ be an algebra. Then μ_0 is called a *premeasure* if

1. $\mu_0(\emptyset) = 0$
2. **"Respects" countable union**: If $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ is a collection of disjoint such that $\cup_i A_i \in \mathcal{A}$, then $\mu_0(\cup_i A_i) = \sum_i \mu_0(A_i)$

Note that the premeasure is automatically finitely additive since we can let all but finitely many of the terms in (2) be $\emptyset$ (and hence are a stronger condition than finitely additive measure). However, it is not yet a measure since it's domain is not necessarily closed under countable union. However, if it is closed for countable union for a particular collection, that collection must satisfy the "countable" condition. Interestingly, by the definition of an algebra, this implies that the countable union can be represented as a finite union and intersection of elements from $\mathcal{A}$.

As before, we can define semifinite and σ-finite in a similar way. Using μ_0, we can define the following outer-measure:

$$\mu^* : P(X) \to [0, \infty] \quad \mu^*(E) = \inf\left\{ \sum_i^{\infty} \mu_0(A_i) \;\middle|\; A_i \in \mathcal{A},\ E \subseteq \cup_i A_i \right\}$$

Proposition 1.3.2: Premeasure To Measure

Let μ_0 be a premeasure on $\mathcal{A}$, and let μ^* be defined as above. Then

1. $\mu^*|_{\mathcal{A}} = \mu_0$
2. Every set in $\mathcal{A}$ is μ^*-measurable

Essentially, μ^* is an extension of μ_0 and the associated restriction of μ^* to a measure will contain $\mathcal{A}$

Proof **:**

1. Let's show $\le$ and $\ge$ for any $E \in \mathcal{A}$. The $\le$ direction is obvious since $\mu^*(E) \le \mu_0(E)$ since $E \subseteq \bigcup_{i=1}^{\infty} A_i$ where $A_1 = E$ and $A_i = \emptyset$ for $i > 1$, and so it is part of the set over which we take the infimum, and so the result *has* to be smaller.

For the $\geq$ direction, let's say $E \subseteq \bigcup_{i=1}^{\infty} A_i$ for any $\{A_i\}_{i=1}^{\infty} \in \mathcal{A}$. Let $B_n = E \cap \big(A_n \setminus \cup_1^{n-1} A_i\big)$. Then the collection $\{B_i\}_{i=1}^{\infty}$ is a disjoint collection of members of $\mathcal{A}$ whose union is E. Thus, by the property of premeasure

$$\mu_0(E) = \sum_1^{\infty} \mu_0(B_i) \overset{!}{\leq} \sum_1^{\infty} \mu_0(A_i)$$

where the $\overset{!}{\leq}$ inequality comes from the fact that $\mu_0(B_i) \leq \mu_0(A_i)$. Since this is true for any arbitrary collection $\{A_i\}_{i=1}^{\infty}$, it follows that

$$\mu_0(E) \leq \mu^*(E)$$

Thus establishing equality

2. We need to show that for any $E \subseteq X$ and $A \in \mathcal{A}$, $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$. As mentioned earlier, it suffice to prove $\geq$ where $\mu^*(E) < \infty$.

   Let $\epsilon > 0$. By the property of the infimum, there exists a subset $\{B_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ where $E \subseteq \cup_i B_i$ such that $\sum_i^{\infty} \mu_0(B_i) \leq \mu^*(E) + \epsilon$. By part (1), μ^* is [countably] additive on $\mathcal{A}$ (since it equals the pre-measure)

$$\begin{aligned} \mu^*(E) + \epsilon &\geq \sum_{i=1}^{\infty} \mu_0(B_i \cap A) + \sum_{i=1}^{\infty} (B_i \cap A^c) && \text{additivity} \\ &\geq \mu^*(E \cap A) + \mu^*(E \cap A^c) && \text{monotinicity} \end{aligned}$$

   and since ϵ was arbitrary, A must be μ^*-measurable, completing the proof

We can in fact be more precise with the measures that extend a premeasure:

Theorem 1.3.2: Premeasure Extensions

Let $\mathcal{A} \subseteq P(X)$ be an algebra, μ_0 a premeasure on $\mathcal{A}$, and $\mathcal{M} = \mathcal{M}(\mathcal{A})$ (i.e. $\mathcal{M}$ is the σ-algebra generated by $\mathcal{A}$). Let μ be the extension of μ_0 from proposition 1.3.2. Then if ν is another measure on $\mathcal{M}$ that extends μ_0, then

$$\nu(E) \leq \mu(E) \qquad \forall E \in \mathcal{M}$$

with equality when $\mu(E) < \infty$ (i.e., it possible that μ is infinite while ν is finite).
If μ_0 is σ-finite, then μ is the unique extension of μ_0 to a measure on $\mathcal{M}$

In other words, we can extend a pre-measure to a measure by knowing what it does on the algebra (which in most cases will be a collection of simpler sets, like intervals or rectangles), and then what happens on the rest of the measurable sets we use the infimum definition. If we define another extension of μ_0, then that measure will be *finer* than the μ given by μ^*, i.e, μ is the *coarsest* extension. If μ_0 is σ-finite, then this will in fact be the only extension of μ_0

Proof **:**
Let's say ν is an extension of μ_0 so that $(X, \mathcal{M}(\mathcal{A}), \nu)$ is a measure space. To show that $\nu(E) \leq \mu(E)$, let's say $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ where $E \subseteq \cup_i A_i$. Then since ν is a measure, it is sub-additive on non-disjoint sets, and so

$$\nu(E) \leq \sum_i \nu(A_i)$$

Since ν extends μ_0, by proposition 1.3.2 $\nu|_{\mathcal{A}} = \mu_0$, so

$$\nu(E) \leq \sum_i \nu(A_i) = \sum_i \mu_0(A_i)$$

Since $\{A_i\}_{i=1}^{\infty}$ was an arbitrary choice, this works for any collection, and so

$$\nu(E) \leq \mu(E) \quad [= \mu^*(E)]$$

We show that $\nu(E) = \mu(E)$ if $\mu(E) < \infty$. Let $A = \cup_i A_i$ for some collection $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$. Then

$$\nu(A) = \lim_{n\to\infty} \nu\left(\bigcup_{i=1}^{n}(A_i)\right) = \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n}(A_i)\right) = \mu(A)$$

If μ is not σ-finite, then there can be more than one extension:

Example 1.4: 2 extensions of pre-measure

1. Let $\mathcal{A}$ be the collection of finite unions of the sets of the form $(a, b] \cap \mathbb{Q}$ with $-\infty \leq a \leq b \leq \infty$. Then $\mathcal{A}$ is an algebra on $\mathbb{Q}$

***Proof* :**
Recall proposition 1.1.6 where we can construct an algebra using an elementary family. Thus, we will show that the collection $\mathcal{A}$ satisfies the 3 properties of an elementary family:

(a) To get the empty-set in $\mathcal{A}$, we will understand the notion of "finite union" to also include "no union". Otherwise, since $\mathbb{Q}$ is dense in $\mathbb{R}$, all intervals $(a, b]$ intersected with $\mathbb{Q}$ is nonempty (by MAT157 knowledge). Since $a < b$, the singleton interval is not admissible, and so we will have to assume this "no union" property

(b) Let $E, F \in \mathcal{A}$. We want to show that $E \cap F \in \mathcal{A}$. By definition, $E = (a, b] \cap \mathbb{Q}$ and $F = (c, d] \cap \mathbb{Q}$. Then by some basic set-theory properties, we know that the intersection of two left-open/right-closed intervals is again a left-open/right-closed interval or $\emptyset$. If their intersection is $\emptyset$, then we're done. If not, let $(e, f] = (a, b] \cap (c, d]$. Then $E \cap F = (e, f] \cap \mathbb{Q}$. Thus, $E \cap F \in \mathcal{A}$

(c) Let $E \in \mathcal{A}$, so that $E = (a, b] \cap \mathbb{Q}$. Let $i \in \mathbb{N}$ ($0 \notin \mathbb{N}$) and:

$$E_i = (a - i, a - (i + 1)] \cap \mathbb{Q} \quad F_i = (b + (i - 1), b + i] \cap \mathbb{Q}$$

then each $E_i, F_i \in \mathcal{A}$, $E_i \cap E_j = F_i \cap F_j = \emptyset$ if $i \neq j$ and $E_i \cap F_j = \emptyset$ for all $i, j \in \mathbb{N}$. Let $E = \{E_i\}_{i=1}^{\infty}$ and $F = \{F_i\}_{i=1}^{\infty}$ and let $G = E \cup F$. Then

$$\bigcup_{g \in G} G = E^c$$

since G is a countable union of disjoint elements of $\mathcal{A}$, this completes the proof.

Thus, by proposition 1.7, $\mathcal{A}$ is an algebra.

2. The σ-algebra generated by $\mathcal{A}$ is $P(\mathbb{Q})$

***Proof* :**
Let $x \in \mathbb{Q}$ Notice that

$$\{x\} = \bigcap_{k=1}^{\infty} \left(x - \frac{1}{k}, x\right] \cap \mathbb{Q}$$

hence, every singleton is inside the σ-algebra of $\mathbb{Q}$. Since every subset of $\mathbb{Q}$ is countable, every subset is the countable union of singletons. Since every singleton is inside the σ-algebra, we have that the σ-algebra is $P(\mathbb{Q})$.

3. Define μ_0 on $\mathcal{A}$ by $\mu_0(\emptyset) = 0$ and $\mu_0(A) = \infty$ for $A \neq \emptyset$. Then μ_0 is a premeasure on $\mathcal{A}$, and there is more than one measure on $P(\mathbb{Q})$ whose restriction to $\mathcal{A}$ is μ_0

***Proof* :**
μ_0 is trivially a pre-measure. By definition, $\mu_0(\emptyset) = 0$. For any nonempty $A \in \mathcal{A}$, we have $\mu_0(A) = \infty$. When checking countable additivity, unless all the sets are empty, we get that both sides are infinite.

Thus, we can define μ to be the μ^* induced by μ_0 restricted to all measurable sets. By construction, all sets are measurable. If a sets is non-empty, then it will always be covered by a nonempty set, which has measure ∞. Thus, If $E \in P(X)$ for any nonempty test set $S \subseteq \mathbb{Q}$

$$\infty = \mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c) = \infty$$

and if $S = \emptyset$

$$0 = \mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c) = 0$$

Thus, $\mathcal{M}_\mu = P(\mathbb{Q})$.

We can define another extension of μ_0 using the counting measure, in particular, let μ_C be the counting measure on $P(\mathbb{Q})$. Then since every $A \in \mathcal{A}$ has countable cardinality

$$\mu_C(A) = \infty$$

Thus, $\mu_C|_{\mathcal{A}} = \mu_0$. However

$$\mu_C(\{1\}) = 1 \neq \infty = \mu(\{1\})$$

Thus, we have defined two separate extensions of μ_0.

1.4 Borel Measure on the Real Line

We now focus on the case of $X = \mathbb{R}$ and $\mathcal{M} = \mathcal{B}_\mathbb{R}$. I would say that this is a "special case", however, as far as I can tell by what I know now, this is essentially were we will be spending most of our time in Analysis. Other measure's like the Haar Measure and the Hausdroff Measure are related to this case (though it might be that $\mathbb{R}$ is replaced with more general spaces, usually locally compact topological groups, which $\mathbb{R}$ is a special example of).

Another way of looking at it is that the euclidean topology is central to the study of many fields of mathematics and physics, and is the basis for our notion of manifolds (which are locally euclidean). Therefore, getting a sense of how to define notions of "size" in $\mathbb{R}$ (and then generalize to $\mathbb{R}^n$ or $\mathbb{R}^\infty$) gives us a solid first step in gaining an intuition on measures.

One way to define a measure in $\mathbb{R}$ is by using our intuition of the fact that the length of (a, b) is $b - a$. This is a good intuition and will be the basis to the generalization of this idea we'll do in order to be able to define a larger class of measures based on this idea which will also have more uses outside of $\mathbb{R}$. Here is how we will motivate the construction of these measures: let μ be a finite measure on Borel sets (a measure define on Borel sets is called the *Borel Measure*). Define:

$$F(x) = \mu((\infty, x])$$

F is sometimes called the *distribution function* of μ. Intuitively, it shows the "accumulated size" of the space. Clearly, F is an increasing function, and is right-continuous since

$$(\infty, x] = \bigcap_k (\infty, x + 1/k]$$

or any decreasing function, $x_n \searrow x$ (i.e. limits are preserved from the right). Furthermore, if we take $a < b$, then $(-\infty, b] = (-\infty, a] \cup (a, b]$, thus

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

Notice how similar this is to the usual result from Riemann integration! In the following, instead of defining F in terms of μ, we will define μ from an increasing right-continuous function F by showing how F define's a pre-measure and using Carathéodory. The special case of $F(x) = x$ will give us our usual notion of length in $\mathbb{R}$.

Throughout this section, we will use sets of the form $(a, b]$ where $a < b$, along with the edge cases of (a, ∞) and $\emptyset$. Let all intervals of this type be called h-intervals (h for half open). It is clear that the union or compliment of two h-intervals are h-intervals (or two disjoint h-intervals) and that h-intervals are closed under finite operations, and so h-intervals from a algebra $\mathcal{A}$. By proposition 1.1.2, the σ-algebra by the set $\mathcal{A}$ is $\mathcal{B}_{\mathbb{R}}$. Using this, we'll first show that we can define a premeasure μ based off of an F:

Proposition 1.4.1: Distribution and Premeasure

Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right continuous. If $(a_i, b_i]$ is a set of disjoint h-intervals. If we define μ_0 to be:

$$\mu_0 \left(\bigcup_i^{\infty} (a_i, b_i] \right) = \sum_{i=1}^{\infty} F(b_i) - F(a_i)$$

and $\mu_0(\emptyset) = 0$, then μ_0 is a premeasure on the algebra $\mathcal{A}$

***Proof* :**
We will first show that μ_0 is well-defined, then we simply must show that $\mu\left(\bigcup_i (a_i, b_i]\right) = \sum_i \mu_0((a_i, b_i])$.

To show it's well-defined, we must show that for any finite representation of $(a, b]$, $\mu_0((a, b]) = F(b) - F(a)$. To that end, let $\{(a_i, b_i]\}_{i=1}^n$ be disjoint where $(a, b] = \sqcup i = 1^n [a_i, b_i)$. Since the co-domain is commutative, we can re-arrange the order that we are taking the union (i.e., re-index the set) so that

$$a < a_1 < b_1 = a_2 < b_2 = a_3 < \cdots < b_n = b$$

Then the resulting μ_0 is simply an alternating series:

$$\sum_{j=1}^{n} F(b_j) - F(a_j) = F(b) - F(a)$$

showing that it is independent of representation of the intervals. Thus, if we have any disjoint collection $\{I_i\}_{i=1}^n$ and $\{J_j\}_{j=1}^n$ such that $\cup_i I_i = \cup_j J_j$, since we can always form alternative series:

$$\sum_i \mu_0(I_i) = \sum_{i,j} \mu_0(I_i \cap J_j) = \sum_j \mu_0(J_j)$$

and so μ_0 is well-defined. Since it is well-defined, then we also get that μ_0 is finitely additive (essentially by definition), and since $\mu_0(\emptyset) = 0$, it is at least a finite measure.

Next, we need to show that it is a premeasure. Since $\mu_0(\emptyset) = 0$, it remains to check that if $\{A_i\}_{i=1}^\infty \subseteq \mathcal{A}$ where $\cup_i^\infty A_i \in \mathcal{A}$, then

$$\mu_0\left(\bigcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu_0(A_i)$$

Since $\cup_i^\infty I_i = I \in \mathcal{A}$, there exists (at least one) subsets $\{A_i\}_{i=1}^n \subseteq \mathcal{A}$ such that $\cup_i^\infty I_i = \cup_i^n A_i$. Therefore the elements from I can be partitioned so that the union of each partition represents a single h-interval. By finite additivity of μ_0, it suffices to consider what happens at one of these partitions. So, let J be the union of one of the families so that $(a,b]J = \cup_{j=1}^\infty I_{i_j}$. We will show that $\mu_0(\cup_j^\infty J_j) = \sum_j^\infty \mu_0(J_j)$ by showing $\leq$ and $\geq$. For the $\geq$ direction, notice that

$$\mu_0(J) = \mu_0\left(\bigcup_j^n J\right) + \mu_0\left(J \setminus \bigcup_j^n J\right) \geq \mu_0\left(\bigcup_j^n J\right) = \sum_j^n \mu_0\left(J\right)$$

Thus, as $n \to \infty$ we have $\mu_0(J) \geq \sum_j^\infty \mu_0(J_j)$.

For the $\leq$ direction, we will take advantage of the right-continuity of F.

(Might not be the case TBD p.34 Folland) First, if either b is infinite, the result holds automatically. If a is infinite, then b would have to also be infinite, and so we have the zero-intervals. So, let a and b be finite.

Let $\epsilon > 0$. Then since F is right-continuous, there exists a δ such that $F(a+\delta) - F(a) < \epsilon$ or similarly $-F(a) < -F(a+\delta) + \epsilon$. Similarly, for the same epsilon, there exists a δ_j such that $F(b_j + \delta_j) - F(b_j) < \frac{\epsilon}{2^n}$. Notice that $(a_j, b_j] \subseteq (a_j, b_j + \delta_j)$, so the open intervals cover $[a,b]$, and since $[a,b]$ is compact, there exists a finite sub-cover! If $\{(a_j, b_j + \delta_j)\}_{i=1}^N$ is the finite subcover of $[a,b]$, if we discard all sets that are contained in a bigger set in this family, then we get that

$$b_j + \delta_j \in (a_{j+1}, \beta_{j+1} + \delta_{j+1}) \quad \text{for } j = 1, ...N-1$$

With this, we get the following sequence of inequalities:

$$\begin{aligned}
\mu_0(J) &= F(b) - F(a) \\
&\leq F(b) - F(a+\delta) + \epsilon \\
&\leq F(b_N + \delta_N) - F(a_1) + \epsilon && \text{extrema of cover} \\
&= F(b_N + \delta_N) - F(a_N) + \sum_{j=1}^{N-1} (F(a_{j+1}) - F(a_j))\epsilon && \text{fancy 0} \\
&= F(b_N + \delta_N) - F(a_N) + \sum_{j=1}^{N-1} (F(b_{j+1} + \delta_j) - F(a_j))\epsilon && \text{expanding} \\
&< \sum_{j=1}^{N} [f(b_j - f(a_j) + \epsilon 2^{-n}] + \epsilon \\
&< \sum_{j=1}^{\infty} \mu_0(J_j) + 2\epsilon
\end{aligned}$$

and since ϵ was arbitrary, the result follows.

This result let's us prove that *any* increasing right-continuous function defines a measure!

Theorem 1.4.1: Distributions define Measure

If $F : \mathbb{R} \to \mathbb{R}$ is an increasing right-continuous function, there is a unique Borel measure μ_F on $\mathbb{R}$ such that $\mu_F((a,b]) = F(b) - F(a)$ for all $a, b \in \mathbb{R}$. This function is unique up to the following condition: if G is another such function, then $F - G$ is constant.
Conversely, if μ is a Borel measure that is bounded on all finite Borel sets, we can define

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

Then F is an increasing right-continuous function

Proof **:**
First, by the previous proposition, F induces a premeasure on $\mathcal{A}$. If $F - G$ is constant, then the sums will cancel out in $F(b) - F(a)$, and so $\mu_F = \mu_G$. Similarly, if the two are equal, F and G can only differ up to a constant and that these measures are σ-finite (recall that $\mathbb{R} = \bigcup_{-\infty}^{\infty}(-i, i+1]$).

For the converse, note that if μ is monotonic, then so is F. Next, above/bellow continuity of μ will imply right-continuity of F for $x \geq 0$ and $x < 0$. Therefore, by construction, $\mu = \mu_F$ on $\mathcal{A}$, and so by proposition 1.4.1, $\mu = \mu_F$ on $\mathcal{B}_{\mathbb{R}}$ by uniqueness in theorem 1.3.2.

Note that F need not be left-continuous. Something with the point-density measure that proves the counter-case.

We have been using intervals of the form $(a, b]$ and right-continuous function. The same theory can be developed with $[a, b)$ intervals and left-continuous functions. If we restrict μ to being a finite Borel measure, then we get $\mu = \mu_F$ where $F(x) = \mu((-\infty, x])$ (i.e. it differs up to a constant of the μ_F when μ is a Borel measure). Next, by the construction of the measure from the pre-measure, the resulting measure is in fact a complete measure. Even better, taking completion of μ_F and then applying Carathéodory is the same as just taking completion of μ after applying Carathéodory to μ_F, and so taking the completion on μ_0 or μ_F does not make a difference! The complete measure is also usually denoted by μ_F and called the *Lebesgue-Stieljes measure* of F.

As a consequence of the Lebesgue-Stieljes measure μ being complete, if we have a measure on a Borel space (which we usually call a Borel measure), then then then the set of measurable sets $\mathcal{M}^*$ might be larger than the number of Borel sets (in particular in that it may contain some zero sets that are not Borel sets).

The Lebesgue-Stieljes measure has a couple of nice properties worth exploring. Let μ be a complete Lebesgue-Stieljes measure on $\mathbb{R}$ associated to F, and let $\mathcal{M}_\mu$ be the associated μ^*-measurable sets. Then for any $E \in \mathcal{M}_\mu$:

$$\mu(E) = \inf\left\{ \sum_1^\infty F(b_i) - F(a_i) \;\middle|\; E \subseteq \cup_i^\infty (a_i, b_i] \right\} = \inf\left\{ \sum_1^\infty \mu((a_i, b_i]) \;\middle|\; E \subseteq \cup_i^\infty (a_i, b_i] \right\}$$

Notice that the contribution of the right end points of every h intervals is clearly negligable and so we can replace it with open intervals

Lemma 1.4.1: h-Intervals to Open Intervals

Let μ be the complete Lebesgue-Stieljes measure on $\mathbb{R}$ associated to F. Then for any $E \in \mathcal{M}_\mu$:

$$\mu(E) = \inf \left\{ \sum_1^\infty \mu((a_i, b_i)) \;\middle|\; E \subseteq \cup_i^\infty (a_i, b_i) \right\}$$

Proof **:**
This is the same epsilon trick you have seen before, right it down here as an exercise later

In fact, we can write down $\mu(E)$ in even simple terms

Theorem 1.4.2: Lebesgue-Stieljes in Terms of Open and Compact sets

Let μ be the complete Lebesgue-Stieljes measure on $\mathbb{R}$ associated to $\mathbb{F}$. Then for any $E \in \mathcal{M}_\mu$:

$$\mu(E) = \inf \{\mu(U) \mid U \supseteq E \text{ and } U \text{ is open}\}$$

$$\mu(E) = \sup \{\mu(K) \mid K \subseteq E \text{ and } U \text{ is compact}\}$$

Proof **:**
By lemma 1.4.1, for all $\epsilon > 0$, there exists open intervals (a_i, b_i) such that $E \subseteq \cup_i^\infty (a_i, b_i)$ and $\mu(E) \leq \sum_i^\infty (a_i, b_i) + \epsilon$. If we let $U = \cup_i (a_i, b_i)$, we have that $\mu(U) \leq \mu(E) + \epsilon$. Conversely, since $E \subseteq U$, $\mu(U) \geq \mu(E)$, and so the result for the first statements follows.

For the 2nd equivalence, let's first suppose E is bounded. If E is closed, then it is compact and so E is contained in the set that we are taking the sup over, and so equality is immediate. If E was open, then we can choose an $\epsilon > 0$ such that $\overline{E} \setminus E \subseteq U$ and $\mu(U) \leq \mu(\overline{E} \setminus E) + \epsilon$. Set $K = \overline{E} \setminus U$. Then K is compact, and so

$$\begin{aligned} \mu(K) &= \mu(E) - \mu(E \cap U) \\ &= \mu(E) - [\mu(U) - \mu(U \setminus E)] \\ &\geq \mu(E) - [\mu(U) - \mu(\overline{E} \setminus E)] \\ &= \mu(E) - \mu(U) + \mu(\overline{E} \setminus E) \\ &\geq \mu(E) - \epsilon \end{aligned}$$

showing that any compact set arbitrarily approximates K.

If E is unbounded, let $E_i = E \cap (i, i+1]$ so that $E = \cup_i E_i>$ Then by what we've just shown, for all $\epsilon > 0$, there exists a $K_j \subseteq E_j$ such that $\mu(K_j) \leq \mu(E_j) - \frac{\epsilon}{2^n}$. Take $H_n = \cup_{-n}^n K_i$. Then H_n is compact, $H_n \subseteq E$, and

$$\mu(H_n) \geq \mu(\cup_{-n}^n E_i) - \epsilon$$

Since this equality holds when $n \to \infty$, the result follows (TBD on how it follows).

This gives us a strikingly easy representation of representing every measurable subset of $\mathbb{R}$!

Theorem 1.4.3: $E \subseteq \mathbb{R}$ Representation

Let $E \subseteq \mathbb{R}$. Then the following are equivalent

1. $E \in \mathcal{M}_\mu$
2. $E = V \setminus N_1$ where V is a G_δ set and $\mu(N_1) = 0$
3. $E = W \cup N_2$ where W is a H_σ set and $\mu(N_2) = 0$

Proof **:**
Clearly, (2) and (3) imply (1). For (1) implies (2) or (3), I would highly recommend you try and solve this in the $\mu(E) < \infty$ case to jog your memory.

For the $\mu(E) = \infty$, TBD (exercise 25 p.37 Folland)

This means that all borel sets, or more generally sets in $\mathcal{M}_\mu$ (also contain all unions of borel sets with measure zero sets). Note that this theorem and proof would be much harrier if we did not assume the measure was complete. ‘

I will leave this section with another way of interpreting finite measurable sets that Folland left as an exercise (we'll use the notation of $A \triangle B = (A \setminus B) \cup (B \setminus A)$)

Proposition 1.4.2: Diminishing Difference

Let $E \in \mathcal{M}_\mu$ and $\mu(E) < \infty$. Then for every $\epsilon > 0$, there exists a set A which is the finite union of open intervals and

$$\mu(E \triangle A) < \epsilon$$

1.4.1 Lebesgue Measure

We will study the case where the measure μ is defined through the distribution $F(x) = x$, i.e., when F represents our usual intuition of length! This is arguably the most important measure on $\mathbb{R}$, and most certainly the most used. For that reason, it has some special symbology: we will denote the measure by m instead of μ and the set of measurable sets by $\mathcal{L}$. If we restrict m to $\mathcal{B}_\mathbb{R}$, we will also write it as m.

The first properties worth establishing is to see that measurable sets in $\mathbb{R}$ inherit some of our intuitions how a "shape" should act under translation and dilation, that it's invariant under translation (just like we assumed for the construction of Vitali sets!) and it scales linearly with dilation:

Proposition 1.4.3: Scaling and Dilation

Let $(X, \mathcal{L}, m)$ be a Lebesgue measure space. Define

$$E + s := \{e + s \mid e \in E\} \qquad rE := \{re \mid e \in E\}$$

Then $E + s, rE \in \mathcal{L}$ for all $r, s \in \mathbb{R}$, and:

$$m(E + s) = m(E) \qquad m(rE) = |r|m(E)$$

Proof **:**
We essentially notice that open intervals are invariant under translation and operate in the form described in the proposition, then apply theorem 1.3.2 (which is a really cool application of this theorem!!)

For any $E \in \mathcal{B}_{\mathbb{R}}$ define $m_s(E) = m(E + s)$ and $m^r(E) = M(rE)$. Then m_s and m^r clearly agree with m and $|r|m$ on finite unions of intervals, and so by theorem 1.3.2 they agree on all of $\mathcal{B}_{\mathbb{R}}$. As a consequence, if $m(E) = 0$, then $m(E + s) = m(rE) = 0$. Thus, for all sets in $\mathcal{L}$ (i.e. all unions of borel sets and Lebesgue zero-sets) must be preserved and linearly scaled (respectively) under translation and dilation:

$$m(E + s) = m(E) \qquad m(rE) = |r|m(E)$$

as we sought to show

Since we are working over $\mathcal{M}(\mathcal{B}_{\mathbb{R}})$, the next interesting questions we may ask about m is how it interacts with the topological sets on $\mathbb{R}$. For starters, it is clear (or should be immediately proven) that any singleton has measure 0. Therefore, by countable additivity, every countable set has measure zero; in particular $m(\mathbb{Q}) = 0$. Notice that the rational are dense which is a "topologically large" concept, however it is measure-theoretically small. In fact, we can even have a collection of *open dense* sets, and still be measure-theoretically small. Take some enumeration $\{r_i\}_{i=1}^{\infty}$ of the rational numbers in $[0, 1]$ and let I be the indexing set for this enumeration $i \in I$. Let I_n be the interval of size $\frac{\epsilon}{2^n}$ which is centered at r_n. Then set $U = (0, 1) \cup_n I_n$. By construction, U must be dense and open in $(0, 1)$, however

$$\mu(U) \leq \sum_{i=1}^{\infty} \mu(I_n) = \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon$$

and hence can be as measure-theoretically small as we want. From this we can also find the opposite, where we have something topologically small but measure-theoretically big, by taking $K = [0, 1] - U$. By monotonicity

$$\mu(K) \geq 1 - \epsilon$$

and yet K is nowhere dense.

There is at least one intuition that can remain: a nonempty *open* set will have non-zero measure. However, the converse, that a non-zero measure set will have an open intervals, is not actually true! For that, we can explore the concept of *cantor sets*

Cantor Set

(I copied this directly from my 357 notes)

(was gonna look through: (commented so that this complies

For those who are Categorically inclined, Cantor sets are almost like the initial objects of the category of comapct metric spaces. They are not initial because there does not exist a unique uniformally continuous function from the cantor set to any compact metric space, but there always will exist at least 1 such function.

Definition 1.4.1: Classical Cantor Set

Let $I = [0, 1]$. From this, we'll construct the cantor set iteratively

1. Start by removing the middle third $I_1 = I - \left(\frac{1}{3}, \frac{2}{3}\right)$
2. Given I_i, remove the open middle third of each connected compoennt

Let $C = \cap_{n=0}^{\infty} I_n$ be the Cantor set. In particular, it is the *standard middle thirds Cantor set.* We'll usually refer to this set as "the" Cantor set.

Notice that C is measurable by continuity from above since $C_1 \subseteq C_2 \subseteq \cdots \subseteq C_n \subseteq$.

Proposition 1.4.4: Properties Of Cantor Sets

Give C the standard subspace metric of $[0, 1]$. Then C is:

1. nonempty and uncountable
2. compact
3. nowhere dense ($\operatorname{int}(\overline{A}) = \emptyset$ for all $A \subseteq C$
4. $m(C) = 0$
5. totally disconnected

Proof **:**
Using some topological properties, since C is the intersection of compact spaces, it is compact.

Next, the cantor set is nonempty since $0 \in C_k$ for all k, and so $0 \in C$. To show C is uncountable, Notice that we can represent every element of the uncountable set $[0, 1]$ in it's base 3 expansion, namely for appropriate $a_k \in \{0, 1, 2\}$:

$$x = \sum_{k=1}^{\infty} a_k 3^{-k}$$

Then take:

$$f : [0, 1] \to C \qquad x = \sum_{k=1}^{\infty} a_k 3^{-k} \mapsto$$

To show total disconnectedness, let's take the same I as before. I is closed in $\mathbb{R}$, and thus in C^n. Note that I^c is the union of finite closed sets, and thus is also closed. Thus, I is clopen in C^n. Thus, $C \cap I$ is also a clopen neigborhood of x in C. Since x was arbitrary, C is totally disconnected.

Finally uncountablenss come from C being compact and compelte, since by theorem **??**, C is uncountable.

This construction does not rely on the size of the middle-third; in fact, we can make it converge to any value we want. To see this, first notice that To do this, we first prove a lemma:

Lemma 1.4.2

Suppose $\{\alpha_i\}_{i=1}^{\infty} \subseteq (0,1)$.

1. $\prod_1^\infty (1-\alpha_i) > 0$ if and only if $\sum_1^\infty \alpha_i < \infty$ (Compare $\sum_1^\infty \log(1-\alpha_i)$ to $\sum_1^\infty \alpha_i$)
2. Given $\beta \in (0,1)$, there always exists a sequence $\{\alpha_i\}_{i=1}^{\infty}$ such that $\prod_1^\infty (1-\alpha_i) = \beta$

***Proof* :**
First, we see that we can convert $\prod_i^\infty (1-\alpha_i)$ to $\sum_i^\infty \log(1-\alpha_j)$ by considering

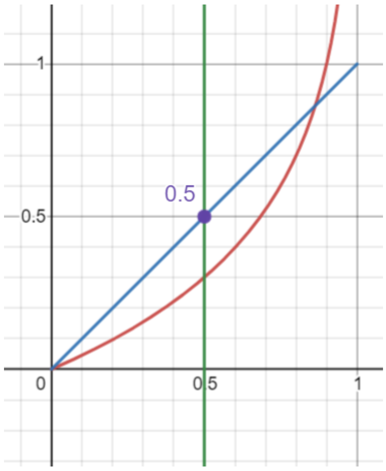
$$\begin{aligned}
\log\left(\prod_i^\infty (1-\alpha_i)\right) &= \log\left(\lim_{n\to\infty}\prod_i^n (1-\alpha_i)\right) \\
&\overset{!}{=} \lim_{n\to\infty} \log(\prod_i^n (1-\alpha_i)) \\
&= \lim_{n\to\infty} \sum_i^n \log(1-\alpha_i) \\
&= \sum_{i=1}^\infty \log(1-\alpha_i)
\end{aligned}$$

where $\overset{!}{=}$ comes from continuity of log for $1-\alpha_i > 0$. Notice that it's impossible that $1-\alpha_i \le 0$ form some α_i terms since $\alpha_i \in (0,1)$, and so this proof works. Thus, we get the equivalent condition of

$$\sum_i^\infty \log(1-\alpha_i) > -\infty$$

With this setup, we proceed with the proof

($\Leftarrow$) Let $\sum_i^\infty \alpha_i < \infty$. Then we know that $\alpha_i \to 0$. Thus, pick an N such that for all $i > N$, $\alpha_i < 0.5$. Then notice that $\log(1-\alpha_i) < \alpha_i$. This is easier to see with a visual aid:



In particular, for all $0 < x < 1/2$, we have $|\log(1-x)| - x > 0$, showing that x is greater than $\log(1-x)$ for all $x \in (0, 1/2)$.

Then from here, this is a simple application of the Weiestrass M-test. Let $K = \sum_i^N \log(1-\alpha_i)$ to eliminate the irregular part. Then since $|\log(1-\alpha_i)| < \alpha_i$, for all $i > N$ and $\sum_i^\infty \alpha_i$ converges, by the Weiestrass M-test the $\sum_{i=N+1}^\infty \log(1-\alpha_i)$ converges. Let's say it converges to L. Then

$$\sum_{i=1}^{\infty} \log(1-\alpha_i) = K + L > -\infty$$

as we sought to show.

(⇒) let $\prod(1-\alpha_i) > 0$ so equivalently $\sum_i^\infty \log(1-\alpha_i) > -\infty$ so $\sum_i^\infty \log(1-\alpha_i) = K$ for appropriate $K \in \mathbb{R}$. Then since the series converges, we have that

$$eK = e\sum_i^\infty \log(1-\alpha_i) = \sum_i^\infty e\log(1-\alpha_i)$$

i.e., we can scale the terms by e. We will show why this scaling is important in a moment; before we do we must setup a little more.

Since the series converges, it must be that $e\log(1-\alpha_i) \to 0$. So pick N such that for all $i > N$, $e\log(1-\alpha_i) < 1/2$. Then notice that $\alpha_i < e\log(1-\alpha_i)$. This is easier to see with a visual aid:

Or, to carry out some of the computations, if you raise $e\log(1-\alpha_i) = \log(1-\alpha_i)e$ to the power of e, you get

$$(e^{\log(1-\alpha_i)})^e = (1-\alpha_i)^e \geq \alpha_i^e \quad \forall \alpha_i \in (0, 1/2]$$

and so the inequality holds. Then the proof continues as for the $\Leftarrow$ direction by setting up an appropriate constant term to bound the irregularities, then using the Weiestrass M test bound every α_i by $\log(1-\alpha_i)$, and since $\sum_i^\infty \log(1-\alpha_i)$ converges, we get that $\sum_{i=N}^\infty \alpha_i$ also converge, and since it will be the sum of a constant plus the convergent value, $\sum_{i=1}^\infty \alpha_i$ converges, as we sought to show.

Now, let β be given. Let $x_i = \prod_{i=1}^n (1-\alpha_i)$. We'll construct a sequence where

$$x_1 = \beta + \frac{1-\beta}{2} \qquad x_2 = \beta + \frac{1-\beta}{4} \quad \cdots \quad x_n = \beta + \frac{1-\beta}{2^n} \cdots$$

to accomplish this, let

$$\alpha_i = 1 - \frac{\beta + \frac{1-\beta}{2^i}}{\beta + \frac{1-\beta}{2^{i-1}}} \tag{1.1}$$

This construction comes from solving the following equation:

$$\left(\beta + \frac{1-\beta}{2}\right) y_2 = \beta + \frac{1-\beta}{4} \ \Rightarrow \ y_2 = \frac{\beta + \frac{1-\beta}{4}}{\beta + \frac{1-\beta}{2}}$$

so $\alpha_2 = 1 - y_2$, and clearly $\alpha_2 \in (0,1)$ since the denominator is bigger than the numerator in the above calculations. Generalizing this pattern, we get the that α_i would be of the from presented in equation (1.1) and $\alpha_i \in (0,1)$. Thus, we get that

$$\lim_{n\to\infty} \prod_i^n (1-\alpha_i) = \lim_{n\to\infty} \beta + \frac{1-\beta}{2^n} = \beta$$

as we sought to show

Now, let's say we are starting with $[0,1]$ but on each iteration eliminates intervals that sum to α_i. Let K_i be the ith step in this itteration. Notice that:

$$m(K_{i+1}) = (1-\alpha_i)m(K_i)$$

Then by what we've just shown, we can make $K = \cap_i^\infty K_i$ be any value up to and including 0 and 1.

Proposition 1.4.5: All Cantor Set Homeomorphism

Let C_n be a cantor set where $m(C_n) = n$. Then $C_n \cong C_0$

***Proof* :**
section 9 of Pugh – will get back to it

You can also prove uncountability directly by making a string represent which side of the split interval you go to. TBD

Next, you might've seen many uncountable sets which are dense in $\mathbb{R}$. Here we have an absolutely facsinating result: C is *nowhere dense in* $[0, 1]$.

Definition 1.4.2: Dense, Somwehre Dense, Nowhere Dense

If $S \subseteq M$ and $\overline{S} = M$, then S is dense in M. The set S is somehwere dense if there exists an open nonemtpy set $U \subseteq M$ such that $\overline{S \cap U} \supset U$. If S is not somwerhe dense, then it is *nowhere dense*

Proposition 1.4.6: C Nowhere Dense

The Cantor set contains no interval and is nowehre dense in $\mathbb{R}$

Proof **:**
To prove it doens't contian an interval, assume it does, then chose an n alrge enough, that is $(\frac{1}{3})^n < b - a$, so that it will not be in the interval.

For anowhere dense, assuem ti's somwerhe then, then $\overline{C \cap U} \supset U \supset (a, b)$ – a contradiction.

We know present how Cantor sets are almost initial objects in compact metric space

Theorem 1.4.4: Cantor Surjection Theorem

Given a compact nonempty metric space M. there is a continuous surjection of C onto M

Proof **:**
In homeowkr, we proved this for $C \to [0, 1]$. Proof in Pugh. starts on p.108

Finally, we present how all cantor sets are homeomorphic.

Theorem 1.4.5: Moore-Kline Theorem

We say M is a Cantor space if it is compact, nonempty, perfect, and totally disconnected (like the cantor set). Every cantor sapce M is homeomorphic to the standard middle-thirds Cantor set C.
In other words, those properties *define* Cantor sets.

Proof **:**
Pugh. p.112

Corollary 1.4.1: Cantor And Fat Cantor Set

The fact Cantor set is homeomorphic to the standard Cantor set

Proof **:**
Using 1.4.5, the result is immediate.

This is intersting, since it shows that meausre is *not* a topological poperty! This should makes sense with $(0,1)$ and $\mathbb{R}$ being homeomorphic with the same measures, but different, but this is an even more extreme example.

Corollary 1.4.2: Product Of Cantor Sets

A cantor set is homeomorhic to its own cartesian product; $C \equiv C \times C$.

***Proof* :**

A fact that Folland added that I thought was cool is that since every subset of the cantor set is Lebesgue Measurable (since the cantor set has measure zero and the Lebesgue measure is complete), we have that

$$|\mathcal{L}| = |P(\mathbb{R})| > \aleph_1 \quad \text{but} \quad |\mathcal{B}_{\mathbb{R}}| = \aleph_1$$

so there are a *tone* of zero-measure sets, however, we can essentially ignore them because they are zero measure sets!!

Chapter 2

Integration

In this chapter, we now work with function's between measure spaces that preserve measure structure! I'll add more intuition here later

Measurable functions are starting to feel like the types of functions we "usually" work with, the same way measurable sets are the sets we "usually" work with. Being continuous can be a bit strong (ex. being a monotone function doesn't imply continuous), but measureability rectifies that by in a sense expanding what are the possible "pre-images" we allow to take (ex. Monotone functions *are* always measurable, as we will demonstrate)

2.1 Measurable Functions

Recall that if $f : X \to Y$ is a function, then $f^{-1} : P(Y) \to P(X)$ is a function where $f^{-1}(E) = \{x \in X \mid f(x) \in E\}$ which preserves unions, intersections, and compliments. Thus, if $\mathcal{N}$ is a σ-algebra on Y, $f^{-1}(\mathcal{N})$ is a σ-algebra on X

Definition 2.1.1: Measurable Function

Let $f : X \to Y$ be a function and $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. Then f is called a $(\mathcal{M}, \mathcal{N})$-*measurable function* (or just measurable) if for all $E \in \mathcal{N}$

$$f^{-1}(E) \in \mathcal{M}$$

Just like in topology, this could be thought of as a surjection of the measurable space on X onto the measurable space on Y. It should be clear that the composition of two measurable functions is measurable, and that the identity function is clearly a measurable function, and so the collection of measurable functions and measurable spaces from a category, usually denoted **Meas**.

Just like lemma 1.1.1 was useful for simplifying proofs, we proof the following:

Lemma 2.1.1: Measurable Function and Generating Sets

Let $\mathcal{N}$ be generated by $\mathcal{E}$. Then $f : X \to Y$ is a $(\mathcal{M}, \mathcal{N})$-measurable function if and only if for all $E \in \mathcal{E}$, $f^{-1}(E) \in \mathcal{M}$

Proof **:**
The $\Rightarrow$ direction a-fortiori true by definition. The $\Leftarrow$ direction comes from f^{-1} preserving unions and compliments, so $\left\{E \subseteq Y \;\middle|\; f^{-1}(E) \in \mathcal{M}\right\}$ is a σ-algebra that contains $\mathcal{E}$, and so contains $\mathcal{N}$.

From this, there is a very nice corrollary when $\mathcal{M}$ is a collection of Borel sets:

Corollary 2.1.1: Continuous Function $\Rightarrow$ Measurable Function

Let $f : X \to Y$ be a continuous function. Then it is a measurable function.

Proof **:**
Since f is continuous, the pre-image of an open set (in particular, open intervals) is open, and hence by proposition 1.1.2 and lemma 2.1.1 it is measurable

Naturally, in analysis, the most important measurable functions will be of the form

$$f : \mathbb{R} \to \mathbb{R} \qquad f : \mathbb{R} \to \mathbb{C}$$

In this case, we will *always* assume the codomain will have $\mathcal{B}_\mathbb{R}$ or $\mathcal{B}_\mathbb{C}$ measure. In fact, if we have a function $f : X \to \mathbb{R}$ or $f : X \to \mathbb{C}$, then we will sometimes say f is $\mathcal{M}$-measurable instead of $(\mathcal{M}, \mathcal{B}_\mathbb{C}$-measurable. We sometimes distinguish between Borel and Lebesgue measurable for functions of the form $f : \mathbb{R} \to \mathbb{C}$ (resp. $f : \mathbb{R} \to \mathbb{R}$. If it's Borel measurable, the domain has the $\mathcal{B}_\mathbb{R}$ measure space, and $\mathcal{L}$ if it is Lebesgue measurable.

Note: Though the composition of two measurable sets is measurable, it does *not* follow that the composition of two *Lebesgue* measurable sets are Lebesgue measurable. see p.44 and homework exercise)

Example 2.1: Measurable Functions
include the composition of two Lebesgue measurable need not be Lebesgue here. Include Monotone function. comment how naturally all continuous functions are also here.

Just like for Borel sets earlier, we have the following equivalences

Proposition 2.1.1: Borel sets and Measurable Functions

Let $(X, \mathcal{M})$ be a measurable space and $f : X \to \mathbb{R}$. Then the following are equivalent:

1. f is $\mathcal{M}$-measurable
2. $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$
3. $f^{-1}([a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$
4. $f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$
5. $f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$

Proof **:**
This was a 357 homework! In this book, it follows from proposition 1.1.2 and lemma 2.1.1.

Similarly, we will soon be encountering co-domains which are product spaces (ex. $\mathbb{R}^n$ or $\mathbb{C}^n$). In particular, let's say we had some set X, and some family of measurable spaces $\{(Y_\alpha, \mathcal{M}_\alpha\}_{\alpha \in A}$, and $f : X \to Y_\alpha$ is a map for each $\alpha \in A$. Then we can impose a unique smallest σ-algebra on X that makes each f_α a measurable function, in particular, $\mathcal{M}_X = \left\{ f_\alpha^{-1}(E_\alpha) \;\middle|\; E_\alpha \in \mathcal{M}_\alpha,\ \alpha \in A \right\}$. This σ-algebra on X is called the *σ-algebra generated by* $\{f_\alpha\}_{\alpha \in A}$.

Proposition 2.1.2: Product of Measurable Functions

Let $(X, \mathcal{M})$ be a measurable space and $(Y_\alpha, \mathcal{N}_\alpha)$ be a collection of measurable spaces, so $Y = \prod_{\alpha \in A} Y_\alpha$, $\mathcal{N} = \bigotimes_{\alpha \in A} \mathcal{N}_\alpha$, and $\pi_\alpha : Y \to Y_\alpha$ is the coordinate map. Then $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable if and only if $f_\alpha = \pi_\alpha \circ f$ is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable.

Proof **:**
Let's say f is $(\mathcal{M}, \mathcal{N})$-measurable. Then since the composition of two measurable functions is measurable, so are each f_α is measurable (since π_α is measurable by proposition 2.1.1).

Conversely, if every f_α is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable, then for each $E_\alpha \in \mathcal{N}_\alpha$, $f^{-1}(\pi_\alpha^{-1}(E_\alpha)) = f^{-1}(E_\alpha) \in \mathcal{M}$, and since these sets generate $\mathcal{M}$, we have that f is $(\mathcal{M}, \mathcal{N})$-measurable by proposition 2.1.1

This proposition also let's us define what it means to take the "product" of two functions. As a consequence, we can give an easy criterion for when complex functions are measurable:

Corollary 2.1.2: Complex Function Measureability

Let $f : X \to \mathbb{C}$ be a complex function. Then f is $\mathcal{M}$-measurable if and only if $\Re f$ and $\Im f$ are both measurable.

Proof **:**
This is simply taking advantage of the topology of $\mathbb{C}$, namely

$$\mathcal{B}_{\mathbb{C}} = \mathcal{B}_{\mathbb{R}^2} \stackrel{!}{=} \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$$

where the $\stackrel{!}{=}$ equality comes from the fact that $\mathbb{R}$ is separable (see proposition 1.1.5). And so proposition 2.1.2 completes the proof.

Sometimes, we want to work with the extended real line (or in general some compactification of a topological space). If $\overline{\mathbb{R}} = [-\infty, \infty]$, Then let

$$\mathcal{B}_{\overline{\mathbb{R}}} = \left\{ E \subseteq \overline{\mathbb{R}} \;\middle|\; E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}} \right\}$$

Notice that is we put the metric ρ on $\overline{\mathbb{R}}$ to be $\rho(x, y) = |\tan^{-1}(x) - \tan^{-1}(y)|$, then $\mathcal{B}_{\overline{\mathbb{R}}}$ is indeed the usual definition of a Borel σ-algebra (exercise)

Next, if the codomain has some notion of $+$ and $\cdot$, we establish that the basic "$\mathbb{C}$-algebra operations" work on measurable functions:

Proposition 2.1.3: Arithmetics on Measurable Functions

Let $f : X \to \mathbb{C}$ and $g : X \to \mathbb{C}$ be $\mathcal{M}$-measurable functions. Then $f + g$, fg, and zf for $z \in \mathbb{C}$ are $\mathcal{M}$-measurable functions

Note that we have limited ourselves to $\mathbb{C}$. This still works with $\overline{\mathbb{R}}$ with the appropriate care of the $\infty - \infty$ and $0 \cdot \infty$ case (exercise 2 in book)

Proof **:**

Since f and g, are measurable, then their "cartesian product" is measurable by proposition 2.1.2

$$F : X \to \mathbb{C} \times \mathbb{C} \qquad F(x) = (f(x), g(x))$$

In particular, F is $(\mathcal{M}, \mathcal{B}_{\mathbb{C}} \otimes \mathcal{B}_{\mathbb{C}})$-measurable. Next, consider $p : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$, $p(x, y) = x + y$ and $t : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$, $t(x, y) = xy$ (p for plus, t for times). Then since p and t are continuous, by corrollary 2.1.1 p and t are $(\mathcal{B}_{\mathbb{C}\times\mathbb{C}}, \mathcal{B}_{\mathbb{C}})$-measurable. Thus, $p \circ F$ and $t \circ F$ is $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$-measurable, and by definition

$$f + g = p \circ F \qquad fg = t \circ F$$

completing the proof

Next, we show that sup, inf, lim sup and lim inf are all preserved under measurability

Comment here for future reference:

Proposition 2.1.4: Measurable functions and Limit Bounding

Let $f_i : X \to \overline{\mathbb{R}}$ be $\mathcal{M}$-measurable functions for each i and consider $\{f_i\}_{i=1}^{\infty}$. Then

$$g_1(x) = \sup_i f_i(x) \qquad g_2(x) = \inf_i f_i(x)$$

$$g_3(x) = \limsup_{i\to\infty} f_i(x) \qquad g_4(x) = \liminf_{i\to\infty} f_i(x)$$

The codomain $\overline{\mathbb{R}}$ was used so that the sup and inf is still considered defined if sup $= \pm\infty$ or inf $= \pm\infty$.

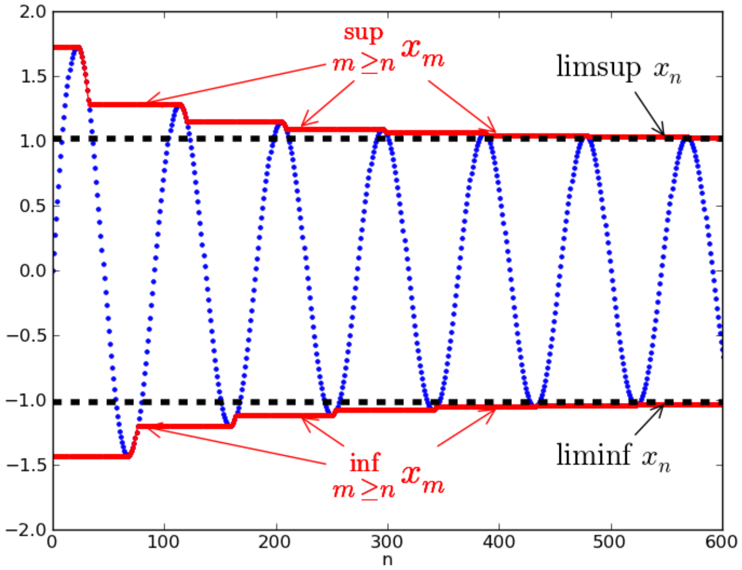
Also, recall that

$$\sup_i f_i(x) = \sup\{f_i(x) \mid i \in \mathbb{N}\}$$

and so this expression means we're taking the set $f_i(x)$ for each i and taking the sup of that set. In other words, for every point x we're taking the supremum of the set $\{f_1(x), f_2(x), ...\}$. Symmetrically the same for inf. For lim sup, recall that:

$$\limsup_i f_i(x) = \lim_{i\to\infty} \sup \{f_k(x) \mid k \geq i\}$$

that is, we are constantly shrinking the set over which we are doing the supremum, which in a sense means we're trying to "hug" the value from above as close as possible. Symmetrically the same for lim inf. If you're visual, I like to keep this image in mind: Let's say the blue line is the value of $f_i(x_0)$ for a specified x_0. Then:



Proof **:**
To show $\sup_i f_i$ and $\inf_i g_i$ are measurable, we will do a very clever trick: notice that:

$$g_1^{-1}((a,\infty]) = \bigcup_i^\infty f_i^{-1}((a,\infty]) \qquad g_2^{-1}(-\infty,a])) = \cup_i^\infty f_i^{-1}([-\infty,a))$$

since the pre-image of rays going off to infinity must contain and so g_1 and g_2 are measurable by proposition 1.1.5. Next, let

$$h_k(x) = \sup_{i>k} f_i(x) = \sup \{f_k(x) \mid k > i\}$$

then by what we've just established, h_k is measurable for each k. Then by construction:

$$g_3 = \inf_k h_k$$

and so g_3 is measurable (notice that we applied inf to h_k). Similarly for g_4.

as an immediate corrollary

Corollary 2.1.3

Let f, g be measurable functions. Then so is $\max\{f, g\}$ and $\min\{f, g\}$

Corollary 2.1.4: Measurable Preserved under Pointwise limit

Let $\{f_i\}$ be a sequence of measurable function. If $\lim_{i\to\infty} f_i(x)$ exists for every x, and we define $f : X \to \overline{\mathbb{R}}$ to be

$$f(x) = \lim_{i\to\infty} f_i(x)$$

then f is measurable (i.e. point-wise convergence preserves measureability)

Recall that point-wise convergence *does not preserve continuity* (the steeple function's are the usual example). However, measureability is sufficiently general to be preserved under this weaker form of convergence! Naturally, if the sequence of functions $\{f_i\}$ are all $\mathcal{M}$-measurable and uniformly convergent, they are measurable

> ***Proof*** **:**
> If f exists (i.e., every $f_n(x)$ converges pointwise to $f(x)$), then by proposition 2.1.4, $f = g_3 = g_4$, completing the proof.

With this in mind, we will write introduce some new notation which will soon become useful when defining integrable functions.

Definition 2.1.2: Real and Complex Decomposition

Let $f : X \to \mathbb{R}$ and $g : X \to \mathbb{C}$ be real and complex measurable functions respectively. Then

$$f^+ = \max\{f, 0\} \qquad f^- = \max\{-f, 0\}$$

and

$$\operatorname{sgn}(g(x)) = \begin{cases} \frac{g(x)}{|g(x)|} & g(x) \neq 0 \\ 0 & g(x) = 0 \end{cases} \qquad |g(x)| = |z| = |x + iy| = \sqrt{x^2 + y^2}$$

which gives the decomposition of:

$$f = f^+ - f^-$$

, $g = (\operatorname{sgn} g)|g|$

The functions f^+ and f^- are clearly well-defined and measurable. For the complex case, $|\cdot|$ is continuous everywhere and $z \mapsto \operatorname{sgn} z$ is continuous everywhere except 0. It is measurable since if $U \subseteq \mathbb{C}$ is open, then $\operatorname{sgn}^{-1}(U)$ is either V for some open set V, or $V \cup \{0\}$, meaning sgn is Borel measurable.

In some textbooks, arg is used instead of sgn.

2.1.1 Standard Representation

We now how that every measurable function is in fact simply the limit of a bunch of easy to work with function, in particular, linear combination of characteristic functions! As a reminder, if $\chi_E : X \to \mathbb{R}$ (or $\{0, 1\}$) where $E \subseteq X$ and

$$\chi_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

Then χ_E is called the characteristic (or indicator) function. It is sometimes also denoted as 1_E or I_E.

Definition 2.1.3: Simple function

Let $E_i \subseteq \mathbb{C}$ be some finite collection of sets. Then the finite linear combination of characteristic functions on E_i multiplied by a complex constant:

$$f = \sum_{i=1}^{n} z_i \chi_{E_i}$$

is called a *simple function*

Example 2.2: Simple Functions

1. Show that $|f(X)| < \infty$ then f is in fact a simple function (hint: let $E_i = z_i$ for each $z_i \in f(X)$). In fact, each coefficient can also be distinct (let it be the respective z_i from the range)
2. Show that if f and g are simple functions, so are $f + g$ and fg

Using these, we can now show that any measurable function is simply the limit of simple functions!

Theorem 2.1.1: Simple Function Approximation

Let $(X, \mathcal{M})$ be a measurable space. Then

1. If $f : X \to [0, \infty]$ is measurable, there exists a sequence of functions $\{\phi_i\}_{i=1}^{\infty}$ such that

$$0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq f$$

   where $\lim_{n\to\infty} \phi_i(x) = f(x)$ (i.e. $\phi_n \to f$ pointwise). and $\phi_n \rightrightarrows f$ on a bounded domain (i.e. uniformally converges)

2. If $f : X \to \mathbb{C}$ is measurable, there exists a sequence of functions $\{\phi_i\}_{i=1}^{\infty}$ such that

$$0 \leq |\phi_1| \leq |\phi_2| \leq \cdots \leq |f|$$

   where $\lim_{n\to\infty} \phi_i(x) = f(x)$ (i.e. $\phi_n \to f$ pointwise). and $\phi_n \rightrightarrows f$ on a bounded domain (i.e. uniformally converges)

Proof **:**
This was a 357 hw! Essentially, you will get indicator functions to converge to the value like so: In particular, for $n \in \mathbb{N}_{\geq 0}$ and $0 \leq k \leq 2^{2n} - 1$, let

$$E_n^k = f^{-1}\left(\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right) \qquad F_n = f^{-1}((2^n, \infty])$$

and

$$\phi_n = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n}\chi_{E_n^k} 2^n \chi_{F_n}$$

then we get that $\phi_n \leq \phi_{n+1}$ for all n and $0 \leq f - \phi_n \leq \frac{1}{2^n}$ where $f \leq 2^n$. The rest is left to check.
Generalizing to $\mathbb{C}$, decomposing to $f = g + ih = (g^+ + g^-) = i(h^+ + h^-)$, where the $+$ and $-$ functions are the positive and negative parts of f and g, then we can repeat the process for each of these and get $\psi_n^+, \psi_n^-, \zeta_n^+, \zeta_n^-$ and let $\phi_n = (\psi_n^+ - \psi_n^-) + i(\zeta_n^+ - \zeta_n^-)$.

Given this, we get a powerfull way of being able to deduce one function is measurable it is equal to a measurable function up to a zero-measure set (if the measure μ is complete)

Proposition 2.1.5: Measurability up to Zero Set

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two measure with ν being a complete measure, and let $f : X \to Y$ be $(\mathcal{M}, \mathcal{N})$-measurable. Then:

1. If $f = g$ up to a zero measure set with respect to μ (shorthanded to μ-a.e.), then g is measurable
2. If f_n is measurable for each $n \in \mathbb{N}$ and $f_n \to f$ μ-a.e, then f is measurable

Proof **:**

1. Let's say N were the points in X such

   Secondly,

2. Since $f_n \to f$ a.e. then $\limsup f_n = \liminf f_n$ a.e., i.e. for every point such Taft $\limsup f_n(x) \neq \liminf f_n(x)$, put them in the set N, and $\nu(N) = 0$. Since N is measurable,

exercise 10 in book

However, if μ is not complete, it turns out to make little difference since we can complete the measure, produce the result, and restrict back to the original measure without changing the result:

Proposition 2.1.6

Let $(X, \mathcal{M}, \mu)$ be a measure sapce and $(X, \overline{\mathcal{M}}, \overline{\mu})$ be its completion. if f is a $\overline{\mathcal{M}}$-measurable function on X, there is a $\mathcal{M}$-measurable function g such that $f = g$ $\overline{\mu}$-a.e.

***Proof* :**
If f is a simple function, then we see that the completion will at most remove finitely many points when we restrict every χ_E $(E \in \overline{\mathcal{M}})$ to g, and so $f = g$ $\overline{\mu}$-a.e.

Now let f be a $\overline{\mathcal{M}}$-measurable function, and by theorem 2.1.1 choose a sequence of $\overline{\mathcal{M}}$-measurable functions $\{\phi_i\}_{i=1}^{\infty}$ that converge pointwise to f. The idea is that each $\phi_i = \psi_i$ $\overline{\mu}$-a.e., that is, there exists a E_i where $\phi_i(x) \neq \psi_i(x)$ for all $x \in E_i$ but $\overline{\mu}(E_i) = 0$, and the countable union of these will still be measure 0.

In particular Let $E = \bigcup_i^{\infty} E_i$. By definition of completion, there must exist some N such that $\mu(N) = 0$ and $E \subseteq N$. Let $g = \lim_{\chi_{X-N}} \psi_n$. Then g is measurable by corollary 2.1.4, and clearly $f = g$ on N^c, completing the proof.

2.2 Integration of non-negative Function

We are now in a position to how to integrate functions (integrable coming soon)! Throughout this section, let $(X, \mathcal{M}, \mu)$ be a measure space

Definition 2.2.1: L^+

Let
$$L^+ = \{f : X \to [0, \infty] \mid f \text{ is a } \mathcal{M}\text{-measurable function}\}$$

Sometimes, L^+ is written as $L^+(X)$, $L^+(\mu)$, or $L^+(\mu, X)$ to emphasize which space of functions we are dealing with. If it is unambiguous, we use L^+.

Definition 2.2.2: Integral of Simple Function

Let $f = \sum_1^n a_i \chi_{E_i}$ be a simple function. Then define the *integral of* f *with respect to* μ to be

$$\int f d\mu = \sum_i^n a_i \mu(E_i)$$

In other words, we simply took the measure of the domain of χ_{E_i} for each i. As usual, we take the convention that $0 \cdot \infty = 0$. We will also allow $\int f = \infty$ to be a valid result. If there is no confusion, we will write $\int f$ instead of $\int f d\mu$. It should also be emphasized that it is the measure of the *domain* and the coefficient from the *codomain* that we are taking.

If we want to integrate over just some $A \subseteq X$ where $A \in \mathcal{M}$, then $\phi|_A$ is also a simple function (simply take $\chi_{E \cap A}$ for each characteristic function). We usually denote this by

$$\int_A f d\mu = \int f \chi_A \, d\mu$$

Because of this notation, we can write $\int f d\mu$ as

$$\int f d\mu = \int_X f d\mu$$

Proposition 2.2.1: properties of Integrals of Simple Functions

Let ϕ and ψ be simple functions in L^+. Then

1. if $c \geq 0$, $\int c\phi d\mu = c \int \phi d\mu$
2. $\int (\phi + \psi) d\mu = \int \phi d\mu + \int \psi d\mu$
3. if $\phi \leq \psi$ then $\int \phi \leq \int \psi$
4. The map $A \mapsto \int_A \phi d\mu$ is a measure on $\mathcal{M}$ (TBD on the ϕ being where it is)

***Proof* :**
Do these as midterm exercises

With this definition, we can define integrable function more generally:

Definition 2.2.3: Integral of Measurable Function

Let $f \in L^+$. Then define

$$\int f d\mu := \sup \left\{ \int \phi d\mu \;\middle|\; \phi \text{ is a simple function, } 0 \leq \phi \leq f \right\}$$

It's important to see why the integral of a simple function will give the same result as the "simple integral" (i.e. integral over simple functions from earlier), in particular since the set over which we are containing the supremum will contain the simple function in question, and so the maximum will be achieved by itself (and proposition 2.13c?). Furthermore, results 1 and 3 from proposition 2.2.1 holds for the definition of integration just given. The other results, in particular linearity, will be established soon.

Given the supremum definition of the integral, it could be rather hard to find the integral of a function in L^+. We essentially never check this, and treat the following theorem as the de-facto way of finding the integral:

Theorem 2.2.1: Monotone Convergence Theorem (MCT)

If $\{f_n\}_{i=1}^{\infty} \subseteq L^+$ is a sequence such that $f_i \leq f_{i+1}$ for all $i \in \mathbb{N}$ and $f = \lim_{n\to\infty} f_n$ $(= \sup_n f_n)$, then

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

or, since $f = \lim_{n\to\infty} f_n$:

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

Note that this is *not* the theorem saying we can move limits through the integral in general (we need more build-up to sate such a result), but rather it is saying that if we have an non-decreasing sequence, then the limit seems like it can "move through" the integral, but in reality we are just equating two results (which I guess is what always happens, but somehow clarifying this in my mind feels useful)

Proof **:**
We'll show $\leq$ and $\geq$. For $\leq$, notice that $\{\int f_n\}$ is an increasing sequence, and so converges (possibly to ∞). Furthermore, $f_n \leq f$ for all n, and so by proposition 2.2.1 updated for general integral functions, we have:

$$\int f_n \leq \int f$$

and so

$$\lim_{n\to\infty} \int f_n \leq f$$

Conversely, we'll show that if we "shrink" the left hand side by a factor of $\alpha \in (0,1)$ so that we get $\geq$, then since this will be true for all α, we will get the result we want.

Fix some $\alpha \in (0,1)$, and choose a simple function where $0 \leq \phi \leq f$ (such a function exists by theorem 2.1.1). Let

$$E_n = \left\{x \in \overline{\mathbb{R}} \;\middle|\; f_n(x) \geq \alpha\phi(x)\right\}$$

Then $\{E_n\}_{n=1}^{\infty}$ is an increasing sequence (via inclusion) of measurable sets who's union if X (since $\lim_{n\to\infty} f_n = f$ and the f_n's are increasing). Then

$$\int f_n \geq \int_{E_n} f_n \geq \alpha \int_{E_n} \phi$$

Since the map $E_n \mapsto \int_{E_n} \phi$ is measurable by proposition 2.2.1 and the E_n's are continuous from above, we have

$$\int_{E_n} \phi = \int \phi$$

and so

$$\lim_{n\to\infty} \int f_n \geq \alpha \int \phi$$

Since this is true for all $\alpha < 1$, it is true for $\alpha = 1$, and so $\int f_n \geq \int \phi$. Taking the supremum over all ϕ, we get that

$$\lim_{n\to\infty} \int f_n \geq \int f$$

And since we showed the $\leq$ direction, we in fact have:

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

as we sought to show

Thus, to find the value of $\int f$, it suffices to find some sequence of simple functions $\{\phi_n\}$ that converges to f, which always exists since every measurable function has a standard representation (theorem 2.1.1).

An immediate consequence of the Monotone Convergence Theorem is that linearity also applies to integrals:

Proposition 2.2.2: Linearity of Integrals

Let $\{f_i\}_{i=1}^N \subseteq L^+$ where $N \in \mathbb{N} \cup \{\infty\}$ and let $f = \sum_i^N f_i$. Then

$$\int f = \sum_i^N \int f_n$$

Proof **:**
Let's first consider the case where $n = 2$ so that we have f_1 and f_2. Let $\{\phi_i\}$ and $\{\psi_i\}$ be a sequence for the standard representation of f_1 and f_2. Then clearly $\{\phi_i + \psi_i\}$ is a sequence for a standard representation of $f_1 + f_2$. Then by the Monotone Convergence Theorem

$$\begin{aligned}
\int f_1 + f_2 &= \lim_{i\to\infty} \int \phi_i + \psi_i \\
&\overset{\text{MCT}}{=} \int \lim_{i\to\infty} \phi_i + \psi_i \\
&= \int \lim_{i\to\infty} \phi_i + \lim_{i\to\infty} \psi_i \\
&= \int \lim_{i\to\infty} \phi_i + \lim_{i\to\infty} \psi_i \\
&= \int \lim_{i\to\infty} \phi_i + \int \lim_{i\to\infty} \psi_i \\
&\overset{\text{MCT}}{=} \lim_{i\to\infty} \int \phi_i + \int \lim_{i\to\infty} \psi_i \\
&= \int f_1 + \int f_2
\end{aligned}$$

By induction, it holds that

$$\int \sum_i^n f_i = \sum_i^n \int f_i$$

Letting $n \to \infty$, we can once again apply the Monotone Convergence Theorem and get

$$\begin{aligned}
\int \sum_i^\infty f_i &= \int \lim_{n\to\infty} \sum_i^n f_i \\
&\overset{\text{MCT}}{=} \lim_{n\to\infty} \int \sum_i^n f_i \\
&= \lim_{n\to\infty} \sum_i^n \int f_i \\
&= \sum_i^\infty \int f_n
\end{aligned}$$

completing the proof

The next result is a generalization from the same result with Riemann integrals, except we replace the condition of "finitely many points" with "null-set amount of points"

Proposition 2.2.3: Zero Integral

Let $f \in L^+$. Then $\int f = 0$ if and only if $f = 0$ a.e.

***Proof* :**
Starting with simple functions, this is evident in both directions. If ϕ was our simple function and $\int \phi = 0$, then $\sum_i^n a_i \mu(E_i) = 0$, so either each $a_i = 0$ or each $\mu(E_i) = 0$ which immediately implies ϕ is zero a.e. (in fact, it is only measure zero). Conversely, if each $\mu(E_i) = 0$, then clearly $\int \phi = 0$.

Moving onto to a measurable function f, the $\Leftarrow$ comes almost immediately. If $f = 0$ a.e., then for every simple function $\phi \leq f$, then $\phi = 0$ a.e., which we have established means $\int \phi = 0$. Then by definition of $\int f$:

$$\int f = \sup_{\phi \leq f} \int \phi = 0$$

For the $\Rightarrow$ direction, assume that $\int f = 0$, and for the sake of contradiction let's say $f \neq 0$ a.e. Let

$$\bigcup_i^\infty E_n = \bigcup_i^\infty \left\{ x \;\middle|\; f(x) > \frac{1}{n} \right\} = \{x \mid f(x) > 0\} = E$$

Since $f \neq 0$ a.e., it must be that $\mu(E) > 0$, and so by construction $\mu(E_n) > 0$. But then

$$f \geq n^{-1} \chi_{E_n} \Leftrightarrow \int f \geq n^{-1} \mu(E_n) > 0$$

so the supremum of values in f contain simple functions of nonzero measure, showing that $\int f > 0$ – a contradiction to our original assumption.

This theorem let's us upgrade the Monotone Conversely Theorem by needing the convergence to happen up to a measure zero set:

Corollary 2.2.1: MCT a.e.

Let $\{f_n\}_{i=1}^\infty \subseteq L^+$ and $f \in {}^+$. Then if $f_1 \leq f_2 \leq \cdots \leq f$ and $f_n \to f$ a.e., then

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

***Proof* :**
Let $f_n(x) \to f(x)$ for all $x \in E$ such that $\mu(E^c) = 0$. Then $f - f|_E = 0$ a.e. which implies $f_n - f_n|_E = 0$ a.e., so by the Monotone Conversely Theorem

$$\int f = \int f|_E \overset{\text{MCT}}{=} \lim_{n \to \infty} \int f_n|_E = \lim_{n \to \infty} \int f_n$$

as we sought to show.

We next address the fact that the sequence $\{f_i\}$ must be increasing (at least a.e.) for the MCT. If it were not, then it need not be the case that the limit passes through the integral!

Example 2.3: not increasing $\not\Rightarrow$ Limit commutes with $\int$

Let $X = R$ and μ be the Lebesgue measure. Consider the two following examples:

$$\{n\chi_{(0,1/n)}\}_{n=1}^{\infty}$$

then it is clear that $n\chi_{(0,1/n)} \to 0$. However, notice that $\int n\chi_{(0,1/n)} = 1$ for each element in the sequence. Thus:

$$0 = \int \lim_{n\to\infty} n\chi_{(0,1/n)} \neq \lim_{n\to\infty} \int n\chi_{(0,1/n)} = 1$$

Even if the function is bounded, the problem remains: take $\chi_{(n,n+1)}$ which also converges pointwise to 0 but the sequence of the integrals of the functions is a sequence of constants 1 and so converges to 1 which is clearly not 0.

Notice that in both examples, there is some "escaping to infinity" going on. We address this when talking about the Dominated Convergence Theorem in the next section.

However, not all hope is lost; there is weakening to only the $\liminf$ and the equality to $\leq$ which works in the general case:

Theorem 2.2.2: Fatou's Lemma

If $\{f_n\} \subseteq L^+$ is *any* sequence, then

$$\int \liminf_n f_n \leq \liminf_n \int f_n$$

This is interesting if we interpret it from a sort of "physics" perspective: somehow, we have "lost our mass" in the process of taking our limit. What Fatou's lemma is giving us is a measure of how we might bound this loss of mass. There In fact, there are essentially 3 ways we can "lose mass', the first two begin the counter-examples above, the last one being a function that oscillates faster and faster (this last one requires a function that is also defined in the negative, and so we will get back to this).

Proof **:**

This comes directly from properties of the $\liminf$. Let $k \geq 1$. Then

$$\inf_{i\geq k} \leq f_i \leq f_j \quad \forall\, j \geq k$$

Then by theorem 2.2.1

$$\int \inf_{i\geq k} \leq \int f_i \quad \forall j \geq k$$

and since this holds for all $j \geq k$, we get:

$$\int \inf_{i\geq k} \leq \inf_{j\geq k} \int f_i$$

Since the $\liminf$ defines an increasing sequence of function, by letting $k \to \infty$ and applying the

Monotone Conversely Theorem:

$$\int \liminf f_n = \lim_{k\to\infty} \int \inf_{n\geq k} f_n \leq \liminf \int f_n$$

as we sought to show

Like before, the result holds if we loosen the condition so that $f_n \to f$ a.e.

Corollary 2.2.2: Fatou's Lemma a.e.

let $\{f_n\} \subseteq L^+$ be any sequence and $f \in L^+$. If $f_n \to f$ a.e. Then

$$\int f \leq \liminf_n \int f_n$$

Proof **:**
Using proposition 2.2.1, make $f_n \to f$ everywhere arguing that it is equivalent to do so, and then the rest is Fatou's Lemma.

Finally, there is this result which I think is for some future proof:

Proposition 2.2.4

Let $f \in L^+$. If $\int f < \infty$, then $\{x \mid f(x) = \infty\}$ is a null set and $\{x \mid f(x) > 0\}$is σ-finite

Proof **:**
was left as an exercise to the reader in the book

2.3 Integration of Real and Complex Functions

We'll again let $(X, \mathcal{M}, \mu)$ be the fixed measure space for this section, except the codomain now will be either $\mathbb{R}$ or $\mathbb{C}$.

For the case of $\mathbb{R}$, we can quickly extend our results from last section by splitting f into f^+ and f^- (which are both measurable if f is by corollary 2.1.3), and if at least one of f^+ or f^- is finite, then

$$\int f = \int f^+ - \int f^-$$

Notice that if they are both minus we have $\infty - \infty$ which is in general not well-defined. Thus, we will limit our attention to functions where $\int |f| < \infty$ since $|f| = f^+ + f^-$:

For the case of $\mathbb{C}$, using the triangle inequality we get that:

$$|f| = |\Re f + \Im f| \leq |\Re f| + |\Im f| \leq f|f|$$

and so both $\Re f$ and $\Im f$ and so we get non-trivial results if each value is finite. In that case, we define

$$\int f := \int \Re f + i \int \Im f$$

We give a name over $\mathbb{C}$ (and with in the special case $\mathbb{R}$) that satisfy this finiteness condition:

Definition 2.3.1: Integrable Function

Let $f : X \to \mathbb{C}$ be measurable. We say f is *integrable* if

$$\int |f| < \infty$$

Let L^1 represent the set of all integrable functions.

If the codomain is $\mathbb{R}$, then we get that $\int |f^+| < \infty$ and $\int |f^-| < \infty$, as we claimed for our extension of the integral to $\mathbb{R}$. Similarly for $\mathbb{C}$. Furthermore, for any $E \in \mathcal{M}$, we will say that f is *integrable on* E if $\int_E |f| < \infty$.

The set of L^1 functions are well-behaved under addition and scalar multiplication in the sense that they form a vector-space:

Proposition 2.3.1: Vector Space on L^1

Let L^1 be the set of integrable functions. Then with $+$ and scalar multiplication, L^1 forms a vector-space (over $\mathbb{R}$ or $\mathbb{C}$ respective to the codomain). Furthermore, $\int$ is a linear functional on L^1 (i.e. $\int : L^1 \to \mathbb{R}$ is a linear functional)

Proof **:**
the proof that $cf + dg$ is bounded comes from the triangle inequality and that $|cf| = |c||f|$. The fact that $cf \in L^1$ for any $c \in \mathbb{R}$ comes from proposition 2.2.1 and $f + g \in L^1$ comes from manipulating the definition using the linearity of the integral.

The fact that $\int$ is a linear functional (i.e. $\int f + g = \int f + \int g$ and $\int cf = c \int g$) comes straight from our previous propositions.

Next, we would like to establish some properties of integrable functions, including a useful inequality, and as in the previous statement, two integrals are equal if $f = g$ a.e. Furthermore, the set of all "interesting" points (points where $f(x) = 0$) have a very nice finiteness property:

Proposition 2.3.2: Properties of Integrable Functions

Let $f, g \in L^1$. Then:

1. $\left|\int f\right| \leq \int |f|$
2. $\{x \mid f(x) \neq 0\}$ is σ-finite
3. for all $E \in \mathcal{M}$, $\int_E f = \int_E g$ if and only if $\int |f - g| = 0$ if and only if $f = g$ a.e.

Proof **:**

1. The result is essentially immediate in the $\mathbb{R}$ case. In the $\mathbb{C}$ case, I'll do soon.
2. Notice that $\{x \mid f(x) \neq 0\} = \{x \mid f^+(x) > 0\} \cup \{x \mid f^-(x) < 0\}$, which we've shown in

proposition 2.2.4 to be σ-finite

3. The fact that $\int |f-g| = 0$ if and only if $f = g$ a.e. comes from proposition 2.2.3, (since $|f-g| \in L^+$) so we will prove $\int_E f = \int_E g$ if and only if $\int |f-g| = 0$. Let's say $\int |f-g| = 0$. Then to show that $\int_E f = \int_E g$, it is equivalent to show that
$$\left| \int_E f - \int_E g \right| = 0$$
by the epsilon principle (i.e. $a = b$ if and only if $|a-b| < \epsilon$ for all ϵ). Using proposition 2.3.2, we get that:
$$\left| \int_E f - \int_E g \right| \leq \int \chi_E |f-g| \leq \int |f-g| = 0$$
completing the $\Rightarrow$ direction.

For the other direction, let's argue contra-positively and assume $\int |f-g| \neq 0$ a.e. Let
$$u = \Re(f-g) \qquad v = \Im(f-g)$$
so that u^+, u^-, v^+, v^- are the corresponding functions. Since $f \neq g$ a.e., at least one of these must have a set with positive measure. Without loss of generality, let's say $E = \{x \mid u^+(x) > 0\}$ and $\mu(E) > 0$. Then
$$\Re\left(\int_E f - g\right) = \Re\left(\int_E f - \int_e g\right) = \int_E u^+ + \int_E u^- = \mu(E) + 0 > 0$$
since $u^- = 0$ on E. Thus, the measure will be positive.

As a consequence of this, we usually think of integrable functions *up to a zero set*. If we wanted to, if f is *only* defined on a measurable set E whose compliment E^c has measure zero, then we can extend f to X by letting it be 0 on E^c, and the integral along with all the properties of f in terms of integrability remain the same. This also means that if we are working with an $\overline{\mathbb{R}}$-value function that is finite a.e., then it is equivalent to treat it as a $\mathbb{R}$-value function.

Proposition 2.3.2 also allows us to re-define L^1 in terms of functions that are equivalent up to a measure zero set. That is, we form an equivalence relation on L^1 if and only if $f = g$ a.e., and label the collection of equivalence classes as L^1. We abuse notation and write "$f \in L^1$" to mean we are selecting a function that is integrable (or an a.e.-defined integrable function). This is because we will usually be working with functions instead of the equivalence classes, and so this is a common abuse of notation.

The reason to introduce this new definition of L^1 is so that we can define a new metric on L^1:

Definition 2.3.2: L^1 metric

Let L^1 be the set of equivalence classes as defined above. Then
$$\rho(f,g) = \int |f-g|$$
is a metric on L^1

Symmetry and triangle inequality come immediately without even needing to use the fact that these are equivalence classes. Where the new construction of L^1 comes important is to say $\int |f - g| = 0$ if and only if $f = g$. Since this is true a.e., then we need these two to be in an equivalence class. Since we defined a metric, we can define a notion of convergence, in particular, f_n converges to f in L^1 if and only if $\int |f - f_n| \to 0$, meaning $f_n \to f$ almost everywhere.

We will soon explore more properties of this metric space, in particular, we would want this metric space to be complete so that the limit of a sequence of L^1 function's is L^1 (so far, we only have the [point-wise] limit of measurable function is measurable). To do so, we need more tools to analyze swapping limits and integrals. The next result we'll cover allows us to generalize the Monotone Conversely Theorem to the general case of $f_n \to f$ a.e., with the extra condition that each $|f_n|$ is bounded by a single $|g|$ with finite area (in this way, we are getting rid of the "escaping to infinity" cases in example 2.3)

Theorem 2.3.1: Dominated Convergence Theorem (DCT)

Let $\{f_n\}_{i=1}^{\infty} \subseteq L^1$ be a sequence of integrable functions such that

1. $f_n \to f$ a.e.
2. there exists a nonnegative $g \in L^1$ such that for all $n \in \mathbb{N}$, $|f_n| \leq g$ a.e.

Then

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

or equivalently:

$$\int \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int f_n$$

Proof **:**

By proposition 2.1.5 and 2.1.6 , f is measurable (up to a null set, in which case we can appropriately redefine f on that null-set). Since $|f_n| < g$ for all $n \in \mathbb{N}$, $f < g$, and so $f \in L^1$.

Since all L^1 functions are broken down to real parts, it suffices to show the result for real value f_n and f. Then by assumption, we have $g + f_n \geq 0$ and $g - f_n \geq 0$. Thus, by Fatou's Lemma:

$$\int g + \int f \leq \liminf \int (g + f_n) = \int g + \liminf \int f_n$$

$$\int g - \int f \leq \liminf \int (g - f_n) = \int g - \limsup \int f_n$$

Therefore, $\limsup \int f_n \leq \int f \leq \liminf \int f_n$, and so the two value are actually equal and so

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

as we sought to show

Notice that the result f was in L^1, meaning we have just shown L^1 is complete! Dominated convergence theorem allow's to define a form of countable linearity for L^1 functions:

Theorem 2.3.2: Linearity for Integrable Functions

Let $\{f_i\}_{i=1}^{\infty} \subseteq L^1$ be a sequence of functions, and $f = \sum_{i=1}^{\infty} f_i$. If $\sum_{i=1}^{\infty} |f_i| < \infty$, then $\sum_{i=1}^{\infty} f_i$ converges a.e. to a function in L^1, and

$$\int_1^{\infty} \sum_{i=1}^{\infty} f_i = \sum_{i=1}^{\infty} \int_1^{\infty} f_i$$

Proof **:**
First, since $\sum_{i=1}^{\infty} |f_i| < \infty$, by proposition 2.2.2,

$$\int_1^{\infty} \sum_{i=1} \infty |f_i| = \sum_{i=1} \infty \int_1^{\infty} |f_i|$$

Thus, $g = |\sum_{i=1}^{\infty} f_i|$ is a function in L^1. The function g is finite a.e. by propositioN 2.3.2[2], so by the dominated convergence theorem, any partial sum of functions will commute, and so taking the limit, we get:

$$\int_1^{\infty} \sum_{i=1} \infty f_i = \sum_{i=1} \infty \int_1^{\infty} f_i$$

as we sought to show

The dominated convergence theorem also gives us a dense set in L^1 to work with, namely the simple functions:

Theorem 2.3.3: Simple Functions Dense in L^1

Let $f \in L^1$. Then for all *epsilon* > 0, there exists an integrable simple function $\phi = \sum a_i \chi_{E_i}$ such that

$$\int |f - \phi| d\mu < \epsilon$$

Furthermore, if μ is a Lebesgue-Stieljes measure on $\mathbb{R}$, the sets E_i can be taken to be finite union of open intervals, and better, there is a continuous function g that vanishes outside a bounded interval (i.e. $g = 0$ outside the bounded interval) such

$$\int |f - g| d\mu < \epsilon$$

Proof **:**
p. 55 in the book. Also proven in fuller generality that simple functions are dense in L^p spaces, which includes L^1.

The next fact we want to establish is a quick result to relate differentiable functions and integral of functions:

Theorem 2.3.4: Relating Derivative and Integral

Let $(X, \mathcal{M}, \mu)$ be a measure space, and suppose $f : X \times [a, b] \to \mathbb{C}$ for $-\infty < a < b < \infty$ and that $f(\cdot, t) : X \to \mathbb{C}$ is integrable for every $t \in [a, b]$, so we can define $F(t) = \int_X f(x, t)\, d\mu$.

1. Suppose that there exists a $g \in L^1$ such that $|f(x, t)| \le g(x)$ for all x, t. Then if $\lim_{t\to t_0} f(x, t) = f(x, t_0)$ for every x, $\lim_{t\to t_0} F(t) = F(t_0)$; that is, if $f(x, \cdot)$ is continuous at every x, so is F.

2. Suppose $\frac{\partial f}{\partial t}$ exists and there is a $g \in L^1(\mu)$ such that $\left|\frac{\partial f}{\partial t}(x, t)\right| \le g(x)$ for all x, t. Then F is differentiable and
$$F'(x) = \int \frac{\partial f}{\partial t}(x, t)\, d\mu(x)$$

The way I like to think about is that there is a system X that at different times t acts/behaves differently. This difference in how the system acts at any given time t is captured by $f(x, t)$. The function F capture the "total value" of this action/behavior (note that f is non-negative too, so all behavior has a "non-negative" output).

Now, if the behavior varies continuously, then the total output varies continuously, and if the behavior is smooth smooth enough so that it is differentiable (at any slice), then F is also differentiable, and it is in fact even easy to compute the value of F'.

Proof **:**

1. Since $|f(x, t)|$ is dominated by g, simply apply the Dominated Convergence Theorem:
$$\lim_{t\to t_0} F(t) = \lim_{t\to t_0} \int f(x, t)\, d\mu(x) = \int \lim_{t\to t_0} f(x, t)\, d\mu(x) = \int f(x, t_0)\, d\mu(x) = F(t_0)$$
where $\{t_n\}$ is any sequence converging to t (since any $f(x, t)$ satisfies the bound, no more restrictions are needed on the choice of sequence)

2. Since $\frac{\partial f}{\partial t}$ exists, the limit of the derivative exists, and so define
$$\frac{\partial f}{\partial t} = \lim_{t_n\to t_0} h_n(x) \qquad h_n(x) = \frac{f(x, t_n) - f(x, t_0)}{t_n - t_0}$$
$\{t_n\}$ is any sequence converging to t. Since h_n difference and division of measurable functions, by proposition 2.1.3 it is measurable, and by the Dominated Convergence Theorem $\frac{\partial f}{\partial t}$ is measurable. Furthermore, by the Mean Value Theorem:
$$|h_n(x)| \le \sup_{t\in[a,b]} \left|\frac{\partial f}{\partial t}(x, t)\right| \le g(x)$$

and so by the Dominated Convergence Theorem again:

$$\begin{aligned}
F'(t_0) &= \lim_{t_n \to t} \left(\frac{F(t_n) - F(t_0)}{t_n - t_0} \right) \\
&= \lim_{t_n \to t} \left(\int h_n(x)\, d\mu(x) \right) \\
&\overset{\text{D.C.T.}}{=} \int \lim_{t_n \to t} h_n(x)\, d\mu(x) \\
&= \int \frac{\partial f}{\partial t}\, d\mu(x)
\end{aligned}$$

completing the proof

2.3.1 Relating Lebesgue and Riemann Integral

Let's consider $f \in L^1(m)$ to be a Lebesgue integrable function on $\mathbb{R}$. As a preliminary to this book, Riemann integration of real-valued functions have been studied. So far, we have done nothing to show that the definitions of the integral will produce the same result. The surely must, since the Riemann integral correctly captures the area under surfaces. In this section, we show how the Lebesgue integral is in fact an extension of the Riemann integral, meaning if we restrict to the set of all Riemann integrable functions (which will all be Lebesgue integral), then the Lebesgue and Riemann integral give the same result.

As a quick reminder of Riemann integration, choosing to follow Darboux' characterization, let $[a, b]$ be a closed interval (or more generally a compact subset of $\mathbb{R}^n$) Then for any finite partition $P = \{t_0, ..., t_n\}$ where $a = t_0 < t_1 < \cdots < t_n = b$, define:

$$U_P(f) = \sum_{i=1}^{n} M_i(t_i - t_{i-1}) \qquad L_P(f) = \sum_{i=1}^{n} m_i(t_i - t_{i-1})$$

where M_i and m_i are the supremum and infimum of f on $[t_{i-1}, t_i]$. Then take:

$$\overline{I}_a^b(f) = \inf_P U_P(f) \qquad \underline{I}_a^b = \sup_P L_p(f)$$

over all possible partitions P. If $\overline{I}_a^b(f) = \underline{I}_a^b$, then their common value is denoted as $\int_a^b f(x)\, dx$ and f is called Riemann integrable on $[a, b]$.

We'll now compare this to the Lebesgue integral:

Theorem 2.3.5: Lebesgue-Vitali Theorem

Let f be a bounded real-valued function on $[a, b]$.

1. If f is Riemann integrable, then f is Lebesgue measurable (and hence integrable on $[a, b]$ since f is bounded) and

$$\int_a^b f(x)\,dx = \int_{[a,b]} f\,dm$$

2. f is Riemann integrable if and only if $\{x \in [a, b] \mid f \text{ is discontinuous at } x\}$ has Lebesgue measure 0.

Note that $\chi_{\mathbb{Q}}$ on $[0, 1]$ is discontinuous everywhere, with discontinuities having Lebesgue measure 1, was shown to not be Riemann integrable (this was in fact the example of a non Riemann-integrable function presented in the preliminary chapter). However, it is certainly Lebesgue integrable!

Proof **:**

1. Let f be Riemann integrable. Let M_i and m_i are defined as the supremum and infimum as defined above, and for each partition P, let:

$$G_P = \sum_{i=1}^{n} M_i \chi_{(t_{i-1},t_i]} \qquad g_P = \sum_{i=1}^{n} m_i \chi_{(t_{i-1},t_i]}$$

so $U_p(f) = \int G_P\,dm$ and $L_p(f) = \int g_P\,dm$. Now, there exists a sequence of partitions P_n such that $\max_i(t_i - t_{i-1})$ tends to zero and P_{n+1} is a refinement of P_n (meaning g_{P_n} increase in value while G_{P_n} decreases) such that $U_{P_n}(f)$ and $L_{P_n}(f)$ converge to $\int_a^b f(x)\,dx$. Let $G = \lim G_{P_n}$ and $g = \lim g_{P_n}$. Since $g \leq f \leq G$, by the dominated convergence theorem:

$$\int G\,dm = \int g\,dm = \int_a^b f(x)\,dx$$

Therefore, $\int G - g = 0$ and so $G = g$ a.e. by proposition 2.2.3, and so $G = f$ a.e.

Finally, since G is measurable (being the point-wise limit of measurable functions)and m is complete, f is measurable and

$$\int_{[} a, b] f\,dm = \int G\,dm = \int_a^b f(x)\,dx$$

2. This was exercise 23 in the textbook

The big advantage we have with this is that we can now use the computational techniques that were developed for Riemann integrals to compute particular Lebesgue integrals! In most cases in analysis, functions are locally Riemann integrable, and so we might what does the generality give us. The main thing it gives us is *completeness.* So though computing Lebesgue integrals might no be so easy (or if they are not an elementary Riemann integrable function), they are easier to work with in a theoretical framework, and help us define concepts like L^p space (which are important for Fourier

Analysis).

(A quick word on the Gamma function)

2.4 Modes of Convergence

So far, we have accumulated the following forms of convergence of functions:

1. uniform convergence $f_n \rightrightarrows f$
2. point-wise convergence $f_n \to f$
3. point-wise convergence a.e. $f_n \to f$ a.e.
4. L^1 convergence $f_n \to f$ if and only if $\int |f_n - f| \to 0$

The first 3 are weaker than then the next: uniform convergence $\Rightarrow$ point-wise convergence $\Rightarrow$a.e. convergence, but the converse is false in each case. However, L^1 convergence. The examples to keep in mind to compare are:

1. $f_n = n^{-1}\chi_{[0,n]}$
2. $f_n = \chi_{[n,n+1]}$
3. $f_n = n\chi_{[0,1/n]}$
4. $f_1 = \chi_{[0,1]}$, $f_2 = \chi_{[0,1/2]}$, $f_3 = \chi_{[1/2,1]}$, $f_4 = \chi_{[0,1/4]}$, $f_5 = \chi_{[1/4,2/4]}$, and so on

We now introduce another way of thinking about convergence: instead of saying that the $\int f_n$ approaches that of f (i.e., evaluate the integral and look at the limiting behavior of the integral of the function), we will want there to exist an N such that $n \geq N$, the measure of the difference between f_n and f are epsilon away from the function we are converging to f. This is not the same thing as L^1 convergence, since this (1), (3) and (4) all converge to 0 in this type of convergence. Before any further analysis let's define our terms:

Definition 2.4.1: Cauchy in Measure

Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ a sequence of measurable functions. Then $\{f_n\}$ said to be *cauchy in measure* if:

$$\forall \epsilon > 0,\ \exists N \text{ such that } \forall n, m \geq N,\ \mu(\{x \mid |f_m(x) - f_n(x)| \geq \epsilon\}) < \epsilon$$

Equivalently, we can write the definition as $\forall \epsilon > 0$,

$$\mu(\{x \mid |f_n(x), f_m(x)| \geq \epsilon\}) \to 0 \text{ as } n, m \to \infty$$

Definition 2.4.2: Convergence in Measure

Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ a sequence of measurable functions. Then $\{f_n\}$ said to be *convergent in measure* if:

$$\forall \epsilon > 0,\ \exists N \text{ such that } \forall n \geq N,\ \mu(\{x \mid |f(x) - f_n(x)| \geq \epsilon\}) < \epsilon$$

With these definition, notice that (2) is in fact *not* Cauchy in measure: the difference between the two functions.

Proposition 2.4.1: L^1 convergence Then Measure Convergence

Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f_n \to f$ in L^1. Then $f_n \to f$ in measure

Proof **:**
Let ϵ be given, and define
$$E_{\epsilon,n} = \mu(\{x \mid |f_n(x) - f(x)| \geq \epsilon\})$$
Then:
$$\int |f - f_n| \geq \int_{E_{\epsilon,n}} |f - f_n| \geq \epsilon\mu(E_{\epsilon,n})$$
Thus:
$$\mu(E_{n,\epsilon}) \leq \epsilon^{-1} \int |f - f_n| \to 0$$
showing that it converges in measure.

Note that convergence in measure does not imply convergence in L^1 as we saw in example (1) and (3).

With this established, we can go onto estalbish why we care about convergence in measure: it gives us another way of finding point-wise convergent functions by establishing a condition weaker than L^1 convergence:

Theorem 2.4.1: Cauchy in Measure Then Pointwise a.e.

Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ Cauchy in measure. Then there exists a f such that $f_n \to f$ in measure, and furthermore there exists a subsequence $\{f_{n_j}\}$ such that $f_{n_j} \to f$ a.e.
If $f_n \to g$ in measure, then $f = g$ a.e.

Proof **:**
p. 61 Folland. Ommited for now since professor did not cover

Corollary 2.4.1: L^1 convergence and Finding Subsequence

Let $f_n \to f$ in L^1. Then there exists a subsequence $\{f_{n_j}\}$ such that $f_{n_j} \to f$ a.e.

Proof **:**
By proposition 2.4.1, $f_n \to f$ in measure, and by theorem 2.4.1, there exists a subsequence that converges to f a.e.

As we saw, $f_n \to f$ a.e. does not imply $f_n \to f$ in measure, as we saw in example (2). This is due to the "escape to infinity" that is happening, which is why we care about the dominated convergence theorem to help us stop such "sneaky infinity problem". However, if μ was a finite measure, then

$f_n \to f$ a.e. does imply $f_n \to f$ in measure. In fact, since the area is bounded, we can some even stronger convergent conditions:

Theorem 2.4.2: Ergoff's Theorem

Let $(X, \mathcal{M}, \mu)$ be a measure space and suppose that $\mu(X) < \infty$. If $\{f_i\}$ are measurable complex-valued functions such that $f_n \to f$ a.e., then for every $\epsilon > 0$, there exists a $E \subseteq X$ such that $\mu(E) < \epsilon$ and $f_n \to f$ uniformly on $\mu(E^c)$

***Proof* :**
p. 62 Folland

(finish this another time)

2.5 Product Measure

As we defined in definition 1.1.4, If $\mathcal{M}$ and $\mathcal{N}$ are σ-algebras, then we can define a new σ-algebra $\mathcal{M} \otimes \mathcal{N}$. If $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are measure spaces, we will show that we can define a new measure on $\mathcal{M} \otimes \mathcal{N}$ using μ and ν. We will use the tools we have done in constructing measure to construct this measure, and show that that there is a natural measure $\mu \times \nu$ with respect to the outer-measure given our space is σ-finite (this is just the uniqueness clause of theorem 1.3.2.

First, we'll give a name to sets in $\mathcal{M} \otimes \mathcal{N}$ that will be easier to measure:

Definition 2.5.1: [Measurable] Rectangle

Let $A \in \mathcal{M}$ and $B \in \mathcal{N}$. Then $A \times B \in \mathcal{M} \otimes \mathcal{N}$ is called a *measurable rectangle* or just *rectangle* for short

Notice that if $A \times B$ and $C \times D$ are rectangles, they form an elementary family:

$$(A \times B) \cup (C \times D) = (A \cap C) \times (B \cap D) \qquad (A \times B)^c = (X \times B^c) \cup (A^c \times B)$$

and so by proposition 1.1.6, the collection of finite disjoint unions of rectangles forms an algebra $\mathcal{A}$, and they generate $\mathcal{M} \otimes \mathcal{N}$.

Furthermore, not all sets in $\mathcal{M} \otimes \mathcal{N}$ are measurable rectangles: if we take $(\mathbb{R}, \mathcal{L}(\mathbb{R}), m)$ and consider $\mathcal{L}(\mathbb{R}) \otimes \mathcal{L}(\mathbb{R})$, then the set of points $\Delta = \{(x, x) \mid x \in \mathbb{R}\}$ is in $\mathcal{L}(\mathbb{R}) \otimes \mathcal{L}(\mathbb{R})$ (being the countable intersection of ever smaller squares on the line), however it is very far from being the product of two sets in $\mathcal{L}(\mathbb{R})$.

We'll now define how to integrate simple functions on $X \times Y$ with the associated measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$: Let $A \times B$ be a rectangle that is the (finite or countable) disjoint union of rectangle $A_i \times B_i$. Then notice that:

$$\chi_{A \times B}(x, y) = \chi_A(x)\chi_B(y)$$

and so:

$$\chi_{A \times B}(x, y) = \sum_i \chi_{A_i \times B_i}(x, y) = \sum_i \chi_{A_i}(x)\chi_{B_i}(y)$$

If we integrate with respect to x, treating $\chi_{B_i}(y)$ as some constant, then by theorem 2.2.2

$$\int \chi_A(x)\chi_B(y)d\mu = \mu(A)\chi_B(y)d\mu = \mu(A)\chi_B(y) = \sum_i \mu(A_i)\chi_{B_i}(y)$$

It might be confusing which variable is associated to which measure, in which case it is common to write:

$$\int \chi_A(x)\chi_B(y)d\mu(x)$$

to emphasize the function with the x variable is the one being measured by μ. Integrating with respect to y yields:

$$\int \mu(A)\chi_B(y)d\nu = \mu(A)\nu(B) = \sum_i \mu(A_i)\nu(B_i)$$

Thus, we can define $\pi : \mathcal{A} \to \mathbb{R}$ where if $E \in \mathcal{A}$ is a finite disjoint union of rectangles $A_i \times B_i$, then:

$$\pi(E) = \sum_i \mu(A_i)\nu(B_i)$$

with the usual convention of $0 \times \infty = 0$ (i.e., if one of the dimensions is "0", then we are making the value 0). Notice that π is independent of representation, since any finite disjoint union of rectangles have a common refinement. Thus π is a pre-measure, and so by theorm 1.3.2, π generates an outer measure on $X \times Y$ whose restriction to $\mathcal{M} \otimes \mathcal{N}$ is a measure where restricted to $\mathcal{A}$ is π (or conversely, whose measure extends π). Such a measure is given a name:

Definition 2.5.2: Product Measure

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{M}, \nu)$ be measure spaces, and define π as in the previous construction. Then the induced outer-measure is called the *product measure* and is usually denoted $\mu \times \nu$.

Notice further that if μ and ν are σ-finite, so $\mu(X) = \sum_i^\infty \mu(A_i)$ and $\nu(Y) = \sum_i^\infty \nu(B_i)$, then $(\mu \times \nu)(X \times Y) = \sum_{i,j} \mu(A_i)\nu(B_j)$, and since all terms are finite, $\mu \times \nu$ is also σ-finite, in which case by the same theorem as invoked in defining $\mu \times \nu$, the product measure is the unique measure such that $\mu \times \nu(A \times B) = \mu(A)\nu(B)$.

Using induction, we can extend this construction to any finite product of measures, that is, for measure spaces $(X_i, \mathcal{M}_i, \mu_i)$, we can define $\mu_1 \times \cdots \mu_n$ on $\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \cdots \otimes \mathcal{M}_n$ such that

$$\mu_1 \times \mu_2 \times \cdots \mu_n(A_1 \times A_2 \times \cdots \times A_n) = \prod_i^n \mu_i(A_n)$$

and if every μ_i is σ-finite, then the result product measure is the unique extension from the premeasure on the rectangles of $\prod_i^n \mathcal{M}_i$.

Next, (should I prove associativity, it is exercise 45)

With what we've established so far, we can already measure sets using the outer-measure definition. However, this is very tedious, and ultimately comes down to some measure properties of the respective component measure. There is an easier way to use these component measure to get the resulting measure: notice in our construction of π that we integrated with respect to one variable first, and then the next. We will ostensibly generalizing this trick so that we can measure spaces, or more specifically integrate functions on product spaces, by taking by fixing all variables except one, integrating on it, and then moving the result out (since it's a constant) and integrate the next variable, until we integrate over all variables.

Since we can generalise our results by induction, let's take $n = 2$ and consider the measure space $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. The following definition tries to capture the idea of fixing everything except one variable:

Definition 2.5.3: n-section

Let $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$ be a product measure space, and consider $E \subseteq X \times Y$. Then for any $x \in X$, $y \in Y$, define the *x-section* E_x and *y-section* E^y to be:

$$E_x = \{y \in Y \mid (x, y) \in E\} \qquad E^y = \{x \in X \mid (x, y) \in E\}$$

If f is a function on $X \times Y$, we define the *x-section* f_x and *y-section* f^y to be:

$$f_x(y) = f(x, y) \qquad f^y(x) = f(x, y)$$

The most simple example of the section of a function would be:

$$(\chi_E)x = \chi_{E_x} \qquad (\chi_E)^y = \chi_{E^y}$$

Proposition 2.5.1: Measure of Section

1. Let $E \in \mathcal{M} \otimes \mathcal{N}$. Then $E_x \in \mathcal{N}$ and $E^y \in \mathcal{M}$ for all $x \in X$ and $y \in Y$ (i.e., all sets in a produce σ-algebra can be broken down into x-sections and y-section)
2. If f is $\mathcal{M} \otimes \mathcal{N}$-measurable, then f_x is $\mathcal{N}$-measurable and f^y is $\mathcal{M}$-measurable

Proof **:**

1. The proof of this fact comes down to taking advantage of the fact that the pre-measure of rectangles generates $\mathcal{M} \otimes \mathcal{N}$. Let $\mathcal{R}$ be the collection of all subsets of $E \subseteq X \times Y$ such that $E_x \in \mathcal{N}$ and $E^y \in \mathcal{M}$ for all $x \in X$ and $y \in Y$. This clearly contains the empty set, and the set must contain all rectangles, since if $A \neq 0$ and $x \in A$, then $(A \times B)_x = B$ (if $x \notin A$, then $(A \times B)_x = \emptyset$). Now, since

$$\left(\bigcup_i^\infty E_i\right)_x = \bigcup_i^\infty (E_i)_x \qquad (E^c)_x = (E_x)^c$$

by some basic set manipulation, we see that $\mathcal{R}$ is a σ-algebra. Since it contains the generating set of $\mathcal{M} \otimes \mathcal{N}$, we have that $\mathcal{R} \supseteq \mathcal{M} \otimes \mathcal{N}$. Thus, all the sets in $\mathcal{M} \otimes \mathcal{N}$ can be decompose in this way (note that we have not proved equality: we have not proved that if every slice is measurable then E is measurable)

2. By some set-theory manipulation, and part (1):

$$f_x^{-1}(B) = (f^{-1}(B))_x \qquad f^y{-1}(B) = (f^{-1}(B))^y$$

and so f_x and f^y must be $\mathcal{N}$ and $\mathcal{M}$ measurable (respectively)

With this, we are almost ready to start proving we can integrate $A \times B \in \mathcal{M} \otimes \mathcal{N}$ with respect to 1 measure at a time. Need to prove a technical result for the next theorem which gives (again) another way of constructing σ-algebra:

Definition 2.5.4: Monotone Class

Let X be a set and $P(X)$ the powerset of X. Then the subset $\mathcal{C} \subseteq P(X)$ is called a *monotone class* if

1. it is cosed under countable unions of increasing sets, that is, if $E_1 \subseteq E_2 \subseteq \cdots$ where $\{E_i\} \subseteq \mathcal{C}$, then $\cup_i E_i \in \mathcal{C}$
2. it is cosed under intersection unions of decreasing sets, that is, if $E_1 \supseteq E_2 \supseteq \cdots$ where $\{E_i\} \subseteq \mathcal{C}$, then $\cap_i E_i \in \mathcal{C}$

Clearly, every σ-algebra is a monotone class (we've even used this property in a couple of proofs). Furthermore, the intersection of any family of monotone class is again a monotone class, and so we can define the unique smallest monotone class that contains the set $\mathcal{E} \subseteq P(X)$, where $\mathcal{E}$ is called the monotone class generated set, and the monotone class is called the monotone class generated by $\mathcal{E}$, and is dentoed $\mathcal{C}(\mathcal{E})$

Lemma 2.5.1: Monotone Class Lemma

Let X be a set and $\mathcal{A} \subseteq X$ be an algebra on some subsets of X. Then the monotone class $\mathcal{C}$ generated by $\mathcal{A}$ is equal to the σ-algebra $\mathcal{M}$ generated by $\mathcal{A}$, that is:

$$\mathcal{C}(\mathcal{A}) = \mathcal{M}(\mathcal{A})$$

Proof **:**
Since $\mathcal{M}$ is itself a monotone class, we automatically have $\mathcal{E} \subseteq \mathcal{M}$, so we'll prove $\mathcal{E} \supseteq \mathcal{M}$.

This is just a lot of set-theory manipulation, I'll leave it for another time

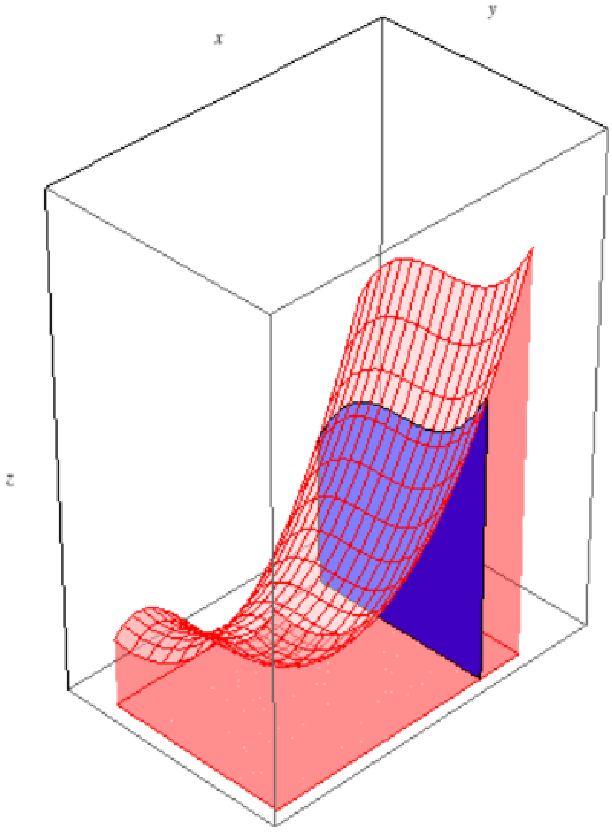
We now are able to start proving the main results:

Theorem 2.5.1: Fubini-Tonelli for Characteristic Functions

Let $(X, \mathcal{M}, \mu)$ and $(N, \mathcal{N}, \nu)$ be measure spaces with μ, ν being σ-finite. Then if $E \in \mathcal{M} \otimes \mathcal{N}$, the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable on X and Y respectively, and

$$(\mu \times \nu)(E) = \int \nu(E_x) d\mu = \int \mu(E^y) d\nu$$

If you are visual, you can think of $\mu(E_x)$ as measuring the area under the slice (inside the appropriate measure so that it doesn't have measure 0), and see that are saying there exists a function that maps x to every slice of the graph:



In particular, it maps x the measure of that slice, so we get a new function which at every slice captures the effect of the X component. After that, we integrate the function to get the effect from the Y component.

Furthermore, it will turn out that we can measure the slices with respect to X, or with respect to Y, and get the same result, hence the second statement of this theorem! After proving the theorem, we'll go over why the assumption in the statement of the theorem are important.

***Proof* :**
The way we will prove it is by taking advantage that rectangles already satisfy the property of the theorem, and since rectangles are a generating set for $\mathcal{M} \otimes \mathcal{N}$, we will be able to extend this result any measurable set.

We'll first prove it for when μ and ν are finite, and then generalize for when they are σ-finite. The way we'll do it is by starting with a subset of $\mathcal{M} \otimes \mathcal{N}$ for which we know the conclusion holds, and then generalize the result using the MCT and DCT.

Let $\mathcal{C}$ be the subset of $\mathcal{M} \otimes \mathcal{N}$ for which the result of the theorem is true, that is, the two maps are measurable and we can integrate one way or another. Then if $A \in \mathcal{M}$ and $B \in \mathcal{M}$, we have that $A \times B \in \mathcal{C}$. To see this, first compute the values $\nu(E_x)$ and $\mu(E^y)$:

$$\nu(E_x) = \chi_A(x)\nu(B) \qquad \mu(E^y) = \mu(A)\chi_B(y)$$

Then the functions is clearly measurable, being a bump function, and:

$$\int \nu(E_x)d\mu = \int \chi_A(x)\nu(B)d\mu = \mu(A)\nu(B) = \int mu(A)\chi_B(y)d\nu = \int \mu(E^y)d\nu$$

Showing the two integrals are equivalent, and since $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$, all equality hold, and so $E = A \times B \in \mathcal{C}$. By additivity of the integral, it follows that all finite disjoint unions of rectangles are also in $\mathcal{C}$. Thus, by lemma 2.5.1, it suffices to show that $\mathcal{C}$ is a monotone class, for then it is a σ-algebra containing the generating set of $\mathcal{M} \otimes \mathcal{N}$, meaning $\mathcal{C} \supseteq \mathcal{M} \otimes \mathcal{N}$, and so all measurable sets can be broken down in this way.

First, let $\{E_n\}_{n=1}^{\infty} \subseteq \mathcal{C}$ be an increasing subset and let $E = \bigcup_n^{\infty} E_n$. We want to show that $E \in \mathcal{C}$. Then since $E_n \in \mathcal{C}$ for all $n \in \mathbb{N}$, the sequence of y-section functions f_n, $f_n(y) = \mu((E_n)^y)$ are measurable and increase pointwise to $f(y) = \mu(E^y)$. By corrollary 2.1.4, f is measurable. Thus, we can find the value of $\int \mu(E^y)d\nu$ by applying the Monotone Convergence Theorem:

$$\begin{aligned} \int \mu(E^y)d\nu &= \int \lim_{n\to\infty} \mu((E_n)^y)\nu \\ &= \lim_{n\to\infty} \int \mu((E_n)^y)\nu \\ &= \lim_{n\to\infty} (\mu \times \nu)(E_n) && E_n \in \mathcal{C} \\ &= (\mu \times \nu)(E) && \cup_i E_i = E \end{aligned}$$

Doing the same trick, we can see that:

$$\int \nu(E_x)d\mu = (\mu \times \nu)(E)$$

which combined together shows us that:

$$\int \mu(E^y)d\nu = \int \nu(E_x)d\mu = (\mu \times \nu)(E)$$

and so $E \in \mathcal{C}$.

Next, let $\{E_n\}_{n=1}^{\infty} \subseteq \mathcal{C}$ be a decreasing subset and let $\bigcap_n E_n = E$. Then sine $E_1 \in \mathcal{C}$, the function f_1, $f_1(y) = \mu((E_1)^y)$ is measurable. In fact, it's in L^1, since:

$$\mu((E_1)^Y) \leq \mu(X) < \infty \qquad \nu(Y) < \infty$$

and so the resulting integral must have finite measure. Furthermore, this function will dominate the sequence of functions f_n, and so we can apply the Dominated Convergence Theorem to do the same limit trick we have just done and conclude that $E \in \mathcal{C}$. Therefore, $\mathcal{C}$ is a monotone class, hence a σ-algebra containing the generating set of $\mathcal{M} \otimes \mathcal{N}$, and the result is true for all measurable sets.

Now, suppose that μ and ν are σ-finite. Then we can write $X \times Y$ to be the union of an increasing sequence of rectangles $\{X_i \times Y_i\}$ of finite measure. Thus, for any $E \in \mathcal{M} \otimes \mathcal{N}$, take $E \cap (X_i \times Y_i)$, then this set now has finite measure, and so what we have just proved applies giving us:

$$(\mu \times \nu(E \cap (X_i \times Y_i)) = \int \chi_{X_i}(x)\nu(E_x \cap Y_i)d\mu = \int \mu(E^y \cap X_i)\chi_{Y_i}(y)d\nu$$

and so, a final application of the Monotone Convergence Theorem gives us our desired result

The hypothesis that $\mu \times \nu$ is σ-finite is necessary. Without it (see exercise 46 in Folland).

Notice that the functions within the integral are in fact the values of characteristic functions! Hence by the linearity of the integral, we have the result for simple functions. Thus, with the statement proven for simple functions on $\mathcal{M} \otimes \mathcal{N}$, we move on to proving it for integrable functions in general:

Theorem 2.5.2: Fubini-Tonelli Theorem

Let $(X, \mathcal{M}, \mu)$ and $(N, \mathcal{N}, \nu)$ be σ-finite measures. Then:

1. (Tonelli) If $f \in L^+(X \times Y)$, then the functions $g(x) = \int f_x d\nu$ and $h(y) \int f^y d\mu$ are in $L^+(X)$ and $L^+(Y)$ respectively, and
$$\int \left[\int f(x,y) d\nu(y) \right] d\mu(x) = \int f(x,y) d(\mu \times \nu) = \int \left[\int f(x,y) d\mu(x) \right] \nu(y)$$
2. (Fubinni) If $f \in L^1(\mu \times \nu)$, then $f_x \in L^1(\nu)$ for a.e. $x \in X$ and $f^y \in L^1(\mu)$ for a.e. $y \in Y$. Furthermore, the a.e. defined function $g(x) = \int f_x d\nu$ and $h(x) = \int f^y d\mu$

The resulting integral is called the *iterated integral*

Proof **:**

1. This essentially comes down to using MCT and the standard representation of measurable functions. By theorem 2.5.1, we have the simple functions satisfy the theorem. To extend it to integrable functions, take $f \in L^+(X \times Y)$, and let $\{f_n\}$ be some simple representation for it, that is, a sequence of simple functions converging upwards to f. From $\{f_n\}$, we have a corresponding $\{g_n\}$ and $\{h_n\}$ converging upwards to g and h: $g_n =$. By theorem 2.5.1, these functions are measurable, and so h and g are measurable. With this, we can put together all our information to form the following chain of equalities:
$$\begin{aligned} \int \left[\int f_x \, d\nu \right] d\mu = \int g d\mu = & \int \lim g_n \, d\mu \\ & = \lim \int g_n \, d\mu \\ & = \lim \int f_n \, d(\mu \times \nu) && \text{theorem } 2.5.1 \\ & = \int f \, d(\mu \times \nu) && \text{MCT} \\ & = \lim \int f_n \, d(\mu \times \nu) \\ & = \lim \int h_n \, d\nu && \text{theorem } 2.5.1 \\ \int \left[\int f^y \, d\mu \right] d\nu = \int h d\nu = & \int \lim h_n \, d\nu \end{aligned}$$
Thus establishing the equality of Tonelli's Theorem. Furthermore, since $\int f \, d(\mu \times \nu) < \infty$, then $g < \infty$ a.e. and $h < \infty$ a.e., so $\int f_x < \infty$ and $\int f^y < \infty$, and therefore $f_x \in L^1(\nu)$ and $f^y \in L^1(\mu)$, which some of the preliminary results for Fubini's Theorem.

2. Let $f \in L^1(\mu \times \nu)$, so $f^+ \in L^+(\mu \times \nu)$ and $f^- \in L^+(\mu \times \nu)$ (or $\Re f$ and $\operatorname{sgn} f$). Thus, the result applies to each component function, and so by linearity applies to f

Often, the parenthesis are dropped, and we combine the integrals into a double integral:

$$\int \left[\int f(x,y)\, d\mu(x) \right] d\nu(y) = \int \int f(x,y)\, d\mu(x)\, d\nu(y) = \int \int f\, d\mu\, d\nu$$

Also, you might wonder if instead of assuming $f \in L^+(X \times Y)$ or $f \in L^1(X \times Y)$, we can assume that each $f_x \in L^+(\nu), f^y \in L^+(\mu)$ (L^1 respectively) and that the iterated integrals $\int \int f\, d\mu\, d\nu$ and $\int \int f\, d\nu\, d\mu$ existed, then $f \in L^+(X \times Y)$ (L^1 respectively), that is, is the converse true? The answer is actually no:

Example 2.4: converse not true
exercise 48

One practical application of Fubini-Tonelli's Theorem is to solve integrals by integrating over the simpler variable first. Usually, one first uses Tonelli's theorme to evaluate the integral $\int f\, d(\mu \times \nu)$ as an iterated integral $\int \int f\, d\mu\, d\nu$ and show that the result is finite, at which point Fubini's Theorem can be invoked to get $\int \int f\, d\mu\, d\nu = \int \int f\, d\nu\, d\mu$.

To close off the discussion on product measures, notice that if μ and ν are complete measure, $\mu \times \nu$ is almost never a complete measure. Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be measure spaces, take $A \in \mathcal{M}$ is a non-empty set with $\mu(A) = 0$, and suppose $\mathcal{N} \neq P(Y)$ (For example, choose $(\mathbb{R}, \mathcal{L}, m)$ for both measure sapces). If $E \in P(Y) \setminus \mathcal{N}$, then $A \times E \notin \mathcal{M} \times \mathcal{N}$ since not all slices of $A \times E$ (in particular, the slice E) are in the respective measure spaces (proposition 2.5.1). However, $A \times E \subseteq A \times Y$, and $(\mu \times \nu)(A \times Y) = 0$ (since the measure is finite, any out-measure is equal to the usual outer-measure on these sets, which is $\mu(A)\nu(Y) = 0 \cdot n = 0$, since we take on the convention that $0 \cdot \infty = 0$). Thus, $\mu \times \nu$ is not complete. If we want, we can apply theorem 1.2.1 to complete the measure, howver we can no longer directly apply Fubini-Tonelli's Theorem. (Put Why Here!). We thus have to slightly amend Fubini-Tonelli's theorem for the completion:

Theorem 2.5.3: Fubini-Tonelli's Theorem for Complete Measures

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be complete σ-finite measures and let $(X \times Y, \mathcal{L}, \lambda)$ be the completion of $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. If f is $\mathcal{L}$-measurable is either:

1. non-negative, that is $f \geq 0$
2. $f \in L^1(\lambda)$

Then f_x is $\mathcal{N}$-measurable for a.e. x and f^y is $\mathcal{M}$-measurable for a.e. y, and if (2) applies as well, then f_x and f^y are integrable for a.e. x and y. Furthermore, $x \mapsto \int f_x\, d\nu$ and $y \mapsto \int f^y\, d\mu$ are measurable (and in the case of (2), also integrable) and

$$\int f\, d\lambda = \int \int f(x,y)\, d\mu(x)\, d\nu(y) = \int \int f(x,y)\, d\nu(y)\, d\mu(x)$$

Proof **:**
homework! Uses Fubini-Tonelli.

2.6 The n-dimensional Lebesgue Integral

In this section, we focus on $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ with the completed product measure $m \times m \times \cdots \times m = m^n$ on $\mathcal{L} \otimes \mathcal{L} \otimes \cdots \otimes \mathcal{L} = \mathcal{L}^n$. The domain $\mathcal{L}^n$ of m^n is called the set of *Lebesgue measurable sets in* $\mathbb{R}^n$: $(\mathbb{R}^n, \mathcal{L}^n, m^n)$ (sometimes, we will restrict m^n to have domain $(\mathcal{B}_{\mathbb{R}})^n = \mathcal{B}_{\mathbb{R}^n}$). If it is unambiguous, we will usually write $(\mathbb{R}^n, \mathcal{L}^n, m)$, dropping the power on the m, and for the $n = 1$ case, we will often write $\int f(x)\, dx := \int f\, dm$.

The following 3 theorems are simply generalizations of previous properties of the Lebesgue measure. In the following, let $E = \prod_{i=1}^n E_i$ where $E_i \subseteq \mathbb{R}$. We will call each E_i the *sides* of E

Theorem 2.6.1: Simplifying outer-measure

Suppose $E \in \mathcal{L}^n$

1. $m(E) = \inf\{m(U) \mid E \supseteq U,\ U \text{ is open}\} = \sup\{m(K) \mid K \subseteq E,\ K \text{ is compact}\}$
2. $E = A_1 \cup N_1 = A_2 \setminus N_2$ where $A_1 \in F_\sigma$, $A_2 \in G_\delta$, and $m(N_1) = m(N_2) = 0$
3. If $m(E) < \infty$, for all $\epsilon > 0$, there is a finite disjoint collection of rectangles $\{R_i\}_{i=1}^n$ whose sides are intervals such that $m(E \Delta \cup_i^n R_i) < \epsilon$

Proof **:**
p. 70 in folland

Theorem 2.6.2: completeness of $L^1(m)$

Let $f \in L^1(m)$ and let $\epsilon > 0$. Then there exists a simple function $\phi = \sum_{i=1}^n a_i \chi_R$ where each R_i is a product of intervals such that

$$\int |f - \phi| < \epsilon$$

and there is a continuous function that vanishes outside a bounded set such that

$$\int |f - g| < \epsilon$$

Proof **:**
Just like in theorem 2.3.3, approximate f by a simple function, then apply proposition 2.6.1(3) to approximate ϕ appropriately. Finally, approximate ϕ by continuous functions by applying the natural generalization of the argument in theorem 2.3.3

Theorem 2.6.3: Translation Invariance in $\mathbb{R}^n$

Let $(\mathbb{R}^n, \mathcal{L}^n, m)$ be a measure space. Then m is translation invariant, that is, for any $a \in \mathbb{R}^n$, if we define $\tau_a : \mathbb{R}^n \to \mathbb{R}^n$, $\tau_a(x) = a + x$, then

1. If $E \in \mathcal{L}^n$, then $\tau_a(E) \in \mathcal{L}^n$ and $m(E) = m(\tau_a(E))$
2. If $f : \mathbb{R}^n \to \mathbb{C}$ is Lebesgue measurable, then so is $f \circ \tau_a$. Furthermore, if either $f \geq 0$ or $f \in L^1(m)$, then $\int (f \circ \tau_a)\, dm = \int f\, dm$

Notice that we did not mention scaling. This is because scaling is not abstracted to linear transformations which requires more work to show how the scaling of a linear transformation affects the value, and so we will prove that later in this section.

Proof **:**
Folland p.71

(here, some on cube approximation and Jordan content)

I'll for now simply state the effect of linear transformations and come back to this:

Theorem 2.6.4: Linear Transformation on Integrability

Suppose $T \in GL_n(\mathbb{R})$. Then:

1. If f is a Lebesgue measurable function on $\mathbb{R}^n$, so is $f \circ T$. If $f \geq 0$ or $f \in L^1(m)$, then:

$$\int f\, dm = |\det(T)| \int f \circ T\, dm$$

2. If $E \in \mathcal{L}^n$, then so is $T(E)$, and $m(T(E)) = |\det(T)| m(E)$

Proof **:**
here

There is also the statement on how diffeomorphisms affect the integral.

Chapter 3

Signed Measure and Differentiation

here

Chapter 4

L^p space

here

4.1 Basics

So far, we have limited our attention to measurable functions f such that $\int |f| < \infty$ and have labelled them L^1. Often, what we are looking are function that in fact converge *faster*, that is, that

$$\int |f|^p < \infty$$

for some $p \geq 1$, including ∞ (which we will define shortly). We'll let L^p be the equivalence class of the set of functions such that $|f|^p$ is integrable. Lots of definition require that f shrink at a certain speed in order for them to be well-defined (Fourier series being the canonical example, see ref:HERE), which is why we are analysing this space. With this added restriction, it becomes harder to prove certain properties of L^p such as whether L^p is a vector-space or is complete. Thus, what we will do here by establishing the basics is showing that L^p is indeed a vector space, and in fact a complete vector space (i.e. a Banach space), and go over some relations between L^p spaces.

We will start by defining L^p

Definition 4.1.1: L^p

Let f be measurable. Define the "*p-norm*" to be:

$$\|f\|_p = \left[\int |f|^p \, d\mu\right]^{1/p}$$

Note that we have yet to prove it's a norm. From this norm, define:

$$L^p(X, \mathcal{M}, \mu) := \left\{f : X \to \mathbb{C} \;\middle|\; f \text{ is measurable and } \|f\|_p < \infty\right\}$$

As with L^1, we often abbreviate to $L^p(\mu)$, $L^p(X)$, or just L^p. Just like L^1, two functions will be equal in L^p (i.e. belong in the same equivalence class) if they are equal almost everywhere. Furthermore, we claimed that $\|\cdot\|_p$ is a norm. That $\|f\|_p = 0$ if and only if $f = 0$ a.e. comes almost immediately, since

$$\left(\int |f|^p\right)^{1/p} = 0 \quad \text{iff} \quad \int |f|^p = 0 \quad \text{iff} \quad \int |f| = 0 \quad \text{iff} \quad f = 0 \text{ a.e.}$$

For $\|cf\|_p = |c| \, \|f\|_p$, we take advantage of our normalisation:

$$\|cf\|_p = \left(\int |cf|^p\right)^{1/p} = \left(\int |c|^p |f|^p\right)^{1/p} = (|c|^p)^{1/p} \left(\int |f|^p\right)^{1/p} = |c| \, \|f\|_p$$

However, the triangle inequality is not so trivial. We will in fact take good part of this section to show that the triangle inequality is satisfied. For now, we take a moment to explore an example of an L^p space to gain an intuition:

Example 4.1: L^p and counting measure

1. Let $f : [0, \infty) \to \mathbb{R}$ $f(x) = \frac{1}{x^{1/p}}$. Then $f \in L^q$ for $q > p$. This comes down to the fact that:

$$x > x^{1/p} \;\Leftrightarrow\; \frac{1}{x} < \frac{1}{x^{1/p}}$$

   and by the fact that $\int \left|\frac{1}{x^n}\right| < \infty$ if $n > 1$. We have therefore found a class of functions that fits into all possible $p \in [1, \infty)$. In fact, we can find functions that fit into any connected subinterval of $[1, \infty)$! Take

   Here, q4b of hw

   So if $f \in L^p$, it does not imply that $f \in L^q$ for $p > q$ or $p < q$! Later on, we will show some of the possible relations between L^p spaces.

2. A particular case that will be common for analyzing sequences is when μ is the counting measure on A. In that case, we usually denote $L^p(\mu)$ by $l^p(A)$. If $A = \mathbb{N}$, we often simply abbreviate notation to l^p. If we write $f : \mathbb{N} \to \mathbb{R}$, $f(k) = a_k$, then the sequence f (i.e. $\{a_k\}$ is in l^p if:

$$\sum_{k=1}^{\infty} |a_k|^p < \infty$$

   Thus, L^1 is the set of convergence series.

Notice that we are "normalizing" each integral by a factor of $1/p$. This normalization is done to make sure that the constant can be brought out of the norm. It might seem tempted to say that:

$$\left(\int |f|^p\right)^{1/p} \text{“=”} \int |f|$$

However, this is most certainly not true (simply take x and see for yourself). As mentioned before, this makes it non-trivial to show that $\|\cdot\|_p$ is a norm on L^p. However, it is easy to see that L^p is a vector-space. If $f, g \in L^p$, then:

$$|f + g|^p \leq [2 \max(|f|, |g|)]^p \leq 2^p(|f|^p + |g|^p)$$

which is finite by assumption, and so $f + g \in L^p$. It is also immediate that $cf \in L^p$. Thus, we may freely add and multiply by L^p functions without worrying about leaving L^p. Because of this, it makes it meaningful to ask whether

$$\|f_1 + f_2\|_p \leq \|f_1\|_p + \|f_2\|_p$$

What we will now do is show that $\|\cdot\|_p$ is does indeed satisfy the triangle inequality or $p \geq 1$. If $p < 1$, then the triangle inequality in fact fails.

Example 4.2: $p < 1$ then no Triangle Inequality

Take $a, b > 0$ and $0 < p < 1$. Then we have that:

$$t^{p-1} > (a+t)^{p-1}$$

Integrating both sides from 0 to b, we get:

$$\begin{aligned} \int_0^b t^{p-1} &> \int_0^b (a+t)^{p-1} \\ \frac{b^p}{p} &> \frac{(a+b)^p}{p} \\ b^p &> (a+b)^p \\ a^p + b^p &> (a+b)^p \end{aligned}$$

Thus, if E and F are disjoint union with positive finite measures in X such that $\mu(E)^{1/p} = a$ and $\mu(E)^{1/p} = b$, taking advantage of the fact that $\chi_E \chi_F = 0$ (i.e. the zero function) since $E \cap F = \emptyset$, then:

$$\|\chi_E + \chi_F\|_p = (a^p + b^p)^{1/p} > a + b = \|\chi_E\|_p + \|\chi_F\|_p$$

Giving us that the triangle inequality does not hold.

Thus, we stick to showing the case where $p \geq 1$ First, we will state a simple inequality fact similar Jensen's inequality. This is in fact the key idea behind the entire proof:

Lemma 4.1.1: Young's Inequality

If $a, b \geq 0$ and $0 < \lambda < 1$, then:

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda) b$$

with equality if and only if $a = b$. In particular, notice that if $\frac{1}{p} + \frac{1}{q} = 1$, then if $\lambda = p^{-1}$, $(1-\lambda) = q^{-1}$ and so:

$$a^{p^{-1}} b^{q^{-1}} \leq p^{-1} a + q^{-1} b$$

***Proof* :**

This proof is trying to capture what this image is doing in formal language:

`https://www.desmos.com/calculator/jdznjzqcju`

If $a = b$, equality clearly holds. If $a = 0$ or $b = 0$, inequality clearly holds. If $a, b \neq 0$, then dividing by b and letting $t = a/b$, we get

$$a^\lambda b^{-\lambda} \leq \lambda(a/b) + (1-\lambda) \quad \Leftrightarrow \quad t^\lambda \leq \lambda t + (1-\lambda) \quad \Leftrightarrow \quad t^\lambda - \lambda t \leq (1-\lambda) \tag{4.1}$$

From here, we use some calculus: we see that $t^\lambda - \lambda t$ is strictly increasing for $0 < t < 1$ (since $a, b \geq 0$, $t \geq 0$) and strictly decreasing for $t > 1$, so the maximal value occurs at $t = 1$, and in fact we get that the maximal value is $1 - \lambda$:

$$1^\lambda - \lambda 1 = 1 - \lambda \leq 1 - \lambda$$

and so the inequalities in equation (4.1) hold! Equality happens when $t = a/b = 1$ which implies $a = b$, completing the proof.

Usually, Young's inequality is stated as

$$a^\alpha b^\beta \leq \alpha a + \beta b \qquad 0 \leq \alpha, \beta \leq 1, \quad \alpha + \beta = 1$$

which is exactly what we have, since if we fix some α, then it must be that $\beta = 1 - \alpha$. Another formulation of Young's inequality is:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

with $p^{-1} + q^{-1} = 1$. This proof is a "Jensen equality" like proof that more clearly shows the role of concavity and so I'll dedicate a moment to prove this one too:

***Proof* :**
If $a = 0$ or $b = 0$, the result is immediate, so assume $a, b > 0$. Then notice that since log is monotone:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad \Leftrightarrow \quad \log(ab) \leq \log\left(\frac{a^p}{p} + \frac{b^q}{q}\right)$$

which is great, since we know log is concave and so Jensen's inequality applies. Since $p^{-1} + q^{-1} = 1$, we get $t = 1/p$ and $(1-t) = 1/q$, and so by Jensen's inequality:

$$\log(ta^p + (1-t)q) \geq t\log(a^p) + (1-t)\log(b^q) = \frac{p}{p}\log(a) + \frac{q}{q}\log(b) = \log(ab)$$

completing the proof.

The second interpretation of $p^{-1} + q^{-1} = 1$ in Young's inequality is a nice way to write the straight line (or sub-straight line) equation without the $1 - \lambda$ term, but just two numbers. It furthermore let's us relate two numbers that can be much larger. For example, if I choose $p = 100$, then it must be that $q = 100/99$ so that

$$p^{-1} + q^{-1} = 1/100 + 99/100 = 1$$

which, if we consider p and q to be the exponents associated to L^p and L^q, we see that this gives us a way to relate different L^p space!

To see this, we use this convexity property given by Young's inequality in the "build-up" lemma towards the triangle inequality. This lemma is important later on in functional analysis, but for our purposes is important for the triangle inequality:

Theorem 4.1.1: Hölder's Inequality

Let $1 \leq p, q \leq \infty$ satisfy:

$$\frac{1}{p} + \frac{1}{q} = 1$$

or equivalently $q = p/(p-1)$, where we will interpret $\frac{1}{\infty}$ as 0. Then if f and g are measurable functions on X then:

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

In particular, if $f \in L^p$ and $g \in L^q$ then $fg \in L^1$. Equality holds if and only if $\alpha|f|^p = \beta|g|^q$ a.e. for some constants α, β with $\alpha\beta \neq 0$

Notice that when $p = q = 2$ (if $p = 2$, it must be that $q = 2$ to satisfy the equation), then this inequality is called the *Cauchy-Schwartz inequality*. Notice too that this is the only time when $p = q$; this will become important when we analyze the $p = 2$ case later.

(I also want to include a quick word on Cauchy-Schwartz inequality since I didn't know how important it is: (commented))

Proof **:**
If $\|f\|_p = 0$ or $\|g\|_p = 0$, then $f = 0$ or $g = 0$ a.e., and so equality holds. Similarly if $\|f\|_p = \infty$ or $\|g\|_p = \infty$, inequality immediately holds. Furthermore, notice that if we find a particular f and g for which the result holds, then the result holds for af and bg since

$$|ab| \|fg\|_1 \leq |ab| \|f\|_p \|g\|_q$$

So, it suffices to show the result when $\|f\|_p = \|g\|_q = 1$, and equality holds of $|f|^p = |g|^q = 1$. To this end, we use the previous lemma and set $a = |f(x)|^p$ and $b = |g(x)|^q$ and $\lambda = p^{-1}$ (so that $1 - \lambda = 1 - p^{-1} = q^{-1}$). Then by the previous lemma we have:

$$|f(x)|^{p \cdot p^{-1}} |g(x)|^{q \cdot q^{-1}} \leq p^{-1}|f(x)|^p + q^{-1}|g(x)|^q \quad \Leftrightarrow \quad |f(x)g(x)| \leq p^{-1}|f(x)|^p + q^{-1}|g(x)|^q \tag{4.2}$$

Now, integrating both sides, we get:

$$\|fg\|_1 \leq p^{-1} \int |f|^p + q^{-1} \int |f|^q = p^{-1} + q^{-1} = 1 = \|f\|_p \|g\|_q$$

where equality holds when equality holds in equation (4.2) a.e., which holds only when $|f|^p = |g|^q$ a.e., as we sought to show

When p and q satisfy the equality in the equation, we call them *conjugate exponents*. As mentioned before, the conjugate exponents is simply another way or writing the equation of a straight line to apply Young's inequality. From this, we can take advantage of this convexity property yet again to prove the triangle inequality of $\|\cdot\|_p$, which has a different name since for the special case of the p-norm due to the fact it is used on other places and needs to be referenced:

Theorem 4.1.2: Minkowski's Inequality

If $1 \leq p < \infty$ and $f, g \in L^p$. Then:

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

Before proceeding with the proof, a quick word a common notation convention:

$$\|f\|_p^p = \left(\int |f|^p\right)^{p/p} = \int |f|^p$$

In as sense, this is not a convention, but just $\left(\|f\|_p\right)^p$. However, I have gotten thrown off by this notation, so I felt a remark was in order.

***Proof* :**
If $p = 1$, then the result immediately follows by the regular triangle inequality and linearity of the integral. If $f + g = 0$ a.e., then the result also follows, so assume that $f + g \neq 0$ a.e., which using the triangle inequality:

$$|f+g|^p \leq (|f| + |g|)|f+g|^{p-1} = |f||f+g|^{p-1} + |g||f+g|^{p-1}$$

So, we will apply Hölder's inequality using the fact that $q = \frac{p}{p-1}$ and integrating:

$$\begin{aligned}
\|f+g\|_p^p &= \int |f+g|^p d\mu \\
&= \int |f+g| \cdot |f+g|^{p-1} d\mu \\
&\leq \int (|f| + |g|)|f+g|^{p-1} d\mu \\
&= \int |f||f+g|^{p-1} d\mu + \int |g||f+g|^{p-1} d\mu \\
&= \left\| |f|(|f+g|^{p-1}) \right\|_1 + \left\| |g|(|f+g|^{p-1}) \right\|_1 \\
&\leq \||f|\|_p \|f+g\|_q^{p-1} + \|g\|_p \|f+g\|_q^{p-1} && \text{Hölder's Inequality} \\
&= \left(\left(\int |f|^p d\mu\right)^{\frac{1}{p}} + \left(\int |g|^p d\mu\right)^{\frac{1}{p}}\right)\left(\int |f+g|^{(p-1)\left(\frac{p}{p-1}\right)} d\mu\right)^{\frac{p-1}{p}} && \text{recall } q = \frac{p}{p-1} \\
&= \left(\left(\int |f|^p d\mu\right)^{\frac{1}{p}} + \left(\int |g|^p d\mu\right)^{\frac{1}{p}}\right)\left(\int |f+g|^p d\mu\right)^{\frac{p-1}{p}} \\
&= (\|f\|_p + \|g\|_p)\frac{\||f+g|^p\|_p}{\|f+g\|_p}
\end{aligned}$$

Thus, multiplying both sides by $\frac{\||f+g|^p\|_p}{\|f+g\|_p}$, we get

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

As we sought to show

Thus, we get that $\|\cdot\|_p$ is indeed a norm! Thus L^p is a normed vector space. In fact, just like L^1, it is a complete normed vector space, i.e., a Banach space. We first prove a lemma to simply our task. To understand the lemma, notice that every norm defines a metric:

$$\rho(x, y) = \|x - y\|$$

and so defines a topology, in particular a normal space, that has the notion of sequences well-defined. Next, if $\{x_n\}$ is a sequence in V, it's series is said to converge if there exists an $x \in V$ such that $\sum_{n=1}^{\infty} x_n = x$, and is said to be absolutely convergence if $\sum_{n=1}^{\infty} \|x_n\| < \infty$

Lemma 4.1.2: Cauchy and Absolute Convegence

Let V be a normed vector space. Then V is complete if and only if every absolutely convergent series in V converges.

***Proof* :**
Let's first say that X is complete and that $\sum_{n=1}^{\infty} \|x\| < \infty$. To show it converges, we take advantage of complete and show there is a Cauchy sequence. In particular, let $S_n = \sum_{k=1}^{n} x_k$. Then since the series is absolutely convergent, there for all $n > m$,

$$\|S_n - S_m\| \leq \sum_{m+1}^{n} \|x_k\| \to 0 \quad \text{as } m, n \to \infty$$

so the sequence $\{S_n\}$ is Cauchy, and so converges.

Conversely, let's say an absolutely convergent series converges, and choose any Cauchy sequence $\{x_k\}$. Since it's a Cauchy sequence, choose a subsequence such that

$$\left\|x_{k_{i+1}} - x_{k_i}\right\| < 2^{-i}$$

Take $y_1 = x_{k_1}$ and $y_i = x_{i+1} - x_i$ for $i > 1$ so that $\sum_{i=1}^{n} y_i = x_{k_n}$. Then notice that $\{y_i\}$ is in fact absolutely convergent:

$$\sum_{i=1}^{\infty} \|y_i\| \leq \|y_1\| + \sum_{i=1}^{\infty} 2^{-i} = \|y_1\| + 1 < \infty$$

So $\sum_{i=1}^{\infty} y_i = \lim x_{k_n}$ exists. Since the subsequence of every Cauchy sequence approaches the same "point" (just like convergent sequences), we see that $\{x_n\}$ also converges to the same point, showing that Cauchy sequences are in fact convergent sequences, as we sought to show.

Theorem 4.1.3: L^p is a Banach Space

For $1 \leq p < \infty$, L^p is a complete normed vector-space, i.e., a Banach space. Moreover, if $f - n \to f$ in L^p, then there exists a subsequence that converges pointwise to f a.e.

***Proof* :**
Let $1 \leq p < \infty$. By the previous lemma, it suffices to show that every absolutely convergent series

converges:

$$\sum_{k=1}^{\infty} \|f\|_p = B < \infty \quad \Rightarrow \quad \exists f \text{ s.t. } \sum_{k=1}^{\infty} f_k = f \in L^p$$

Let $G_n = \sum_{k=1}^{n} |f_k|$ be the partial sums of the series and $G = \sum_{k=1}^{\infty} |f_k|$ be the series. Then by Minkowski's inequality, we get:

$$\|G_n\|_p \leq \sum_{k=1}^{n} \|f_k\|_p \leq B \quad \forall n \in \mathbb{N}$$

and so $\|G\|_p = \left(\int |G|^p\right)^{1/p} < \infty$ and so $\left(\int |G|^p\right) < \infty$. Since $G_n \leq G_{n+1}$, and $G_n : X \to [0, \infty)$, by the Monotonce Convergence Theorem:

$$\int \lim_{n\to\infty} G_n^p = \lim_{n\to\infty} \int G_n^p \leq B^p < \infty$$

and so $G \in L^p$. By proposition 2.3.2, $G(x) < \infty$ a.e., so the series $\sum_{k=1}^{\infty} f_k$ converges a.e. If we let $F = \sum_{k=1}^{\infty} f_k$, we have that $|F| \leq G$ a.e., and so $F \in L^p$. From this, we can see that:

$$\left\| F - \sum_{k=1}^{\infty} f_k \right\|_p^p = \int \left| F - \sum_{k=1}^{\infty} f_k \right|^p \to 0$$

Thus, the series $\sum_{k=1}^{\infty} f$ converges (to F) in the p-norm, and so by lemma 4.1.2, L^p is complete, as we sought to show.

This is the proof the prof was doing in class, but I got stuck so found a different proof (the above proof) (Commented it out for now)

This also finally showed that L^1 is a complete space under $\|\cdot\|$! Furthermore, we can now extend theorem 2.3.3 to show that simple function are dense in every L^p!

Theorem 4.1.4: Simple Functions Dense in L^p

For $1 \leq p < \infty$, the set of simple functions $f = \sum_{i=1}^{n} a_i \chi_{E_i}$ where $\mu(E_i) < \infty$ for all i is dense in L^p

***Proof* :**
First of all, the set of simple functions is in L^p for every $1 \leq p < \infty$. Next, choosing any $f \in L^p$, choose some standard representation $\{f_n\}$ where $f_n \to f$ a.e. Then sine $|f_n| \leq |f|$, we get that $|f_n| \in L^p$. Furthermore

$$|f - f_n|^p \leq |2f|^p = 2^p |f|^p \in L^1$$

and so by the dominated convergence theorem, we get:

$$\int |f - f_n|^p \to 0 \quad \Rightarrow \quad \left(\int |f - f_n|^p\right)^{1/p} = \|f - f_n\|_p \to 0$$

Moreover, if $f_n = \sum_{k=1}^{\infty} a_k \chi_{E_k}$ was a simple function where all E_i are pairwise disjoint and a_i are nonzero, it must be that $\mu(E_i) < \infty$ since

$$\sum_{k=1}^{\infty} |a_k|^p \mu(E_k) = \int |f|^p < \infty$$

completing the proof

Rmk: $1 \le p, q < \infty$, each L^p is dense in L^q.

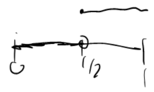
L^∞, $f(x) \equiv 1$ counter-example for $p = \infty$.

continuous functions vanish at ∞.

Lemma: We have $C_0(\mathbb{R})$ is dense in $L^p(\mathbb{R})$ $1 \le p < \infty$,

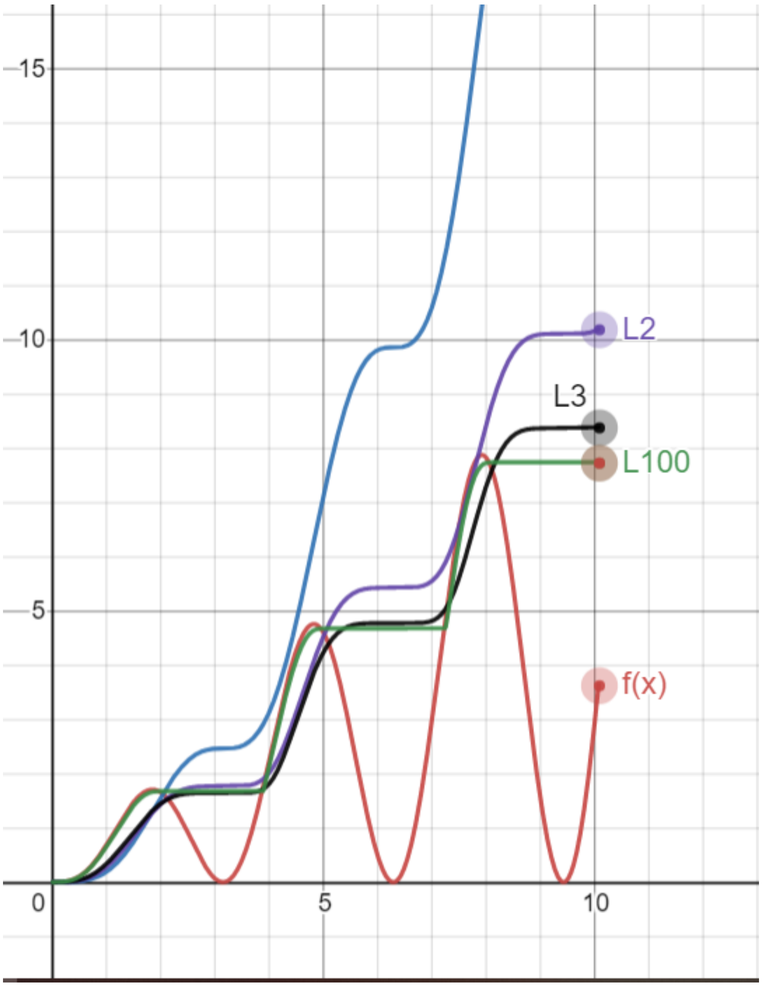
($C([a,b])$ dense in $L^p([a,b])$.)

False for $p = \infty$



← can't be uniform limit of cont. func.

So far, we have defined L^p spaces where $0 < p < \infty$ and showed that for $1 \geq p < \infty$, L^p is a Banach space. However, if you take a look at a graph of $\|f\|_p$ as $p \to \infty$, you'll find that there is a striking pattern that comes back many times:



Notice that as $p \to \infty$, the resulting $\|f\|_p$ value got closer to a max function. In fact, that is indeed the case! We might be tempted to then say that $\|f\|_\infty = \sup f$. However, the problem with this is that we might have a measure 0 set that diverges, while the supremum of the rest of the function is attained by a value $k < \infty$. The idea of $\sup f$ on all but a zero measure set will be called the *essential supremum*, and all functions that have a finite essential supremum will be denoted as L^∞. In this way, L^∞ is simply a natural extension of our definition of L^p:

Definition 4.1.2: L^∞

Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f : X \to [0, \infty]$ be a measurable function. Let

$$S = \left\{ r \in \mathbb{R} \;\middle|\; \mu(f^{-1}((r, \infty])) = 0 \right\}$$

Then the *essential supremum* of f, denoted $\|f\|_\infty$ or $\operatorname{ess\,sup} f$ is :

$$\|f\|_\infty := \begin{cases} \infty & \text{if } S = \emptyset \\ \inf S & \text{if } S \neq \emptyset \end{cases}$$

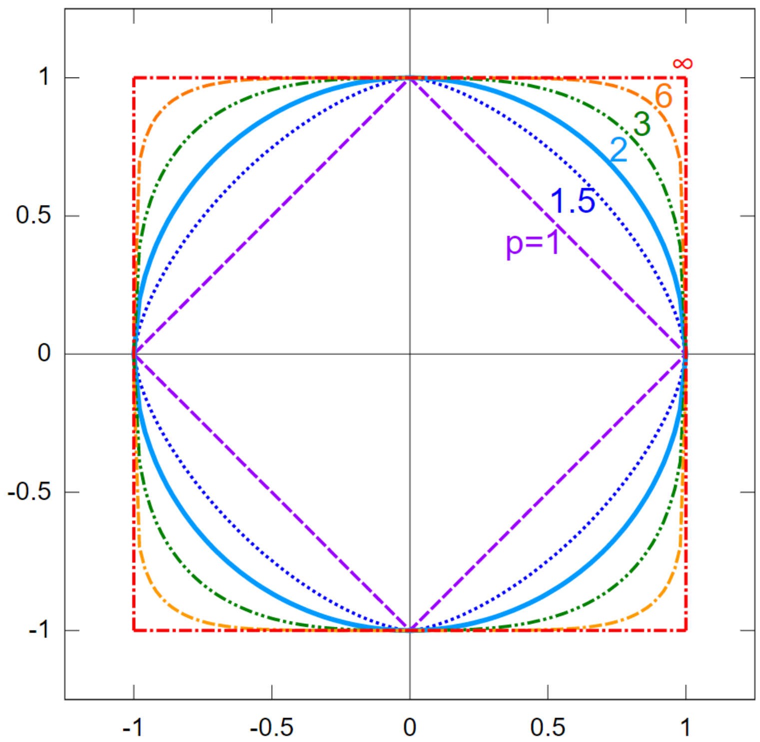
or, equivalently:

$$\|f\|_\infty = \inf_{\substack{E \in \mathcal{M} \\ \mu(X \setminus E) = 0}} \left(\sup_{x \in E} |f(x)| \right)$$

If f is a complex measurable, function, then $\|f\|_\infty = \||f|\|_\infty$

Essentially, $\|f\|_\infty$ is the least upper bound of f, ignoring measure zero sets worth of points. Using this definition, we can define L^∞ to be the equivalence classes of functions f such that $\|f\|_\infty < \infty$.

(Make this rigorous) Another quick way of visualizing these norms is by taking the set of all points $\mathbb{R}^2$ such that $\|(x, y)\|_p = 1$ A quick visual way of seeing the relation between L^p and L^∞ is as follows: if we take $X = R^2$ and take $S_p = \left\{ x \in \mathbb{R}^2 \;\middle|\; \|x\|_p = 1 \right\}$, we'd get



(some remarks Folland does on p. 184)

Proposition 4.1.1: Properties of L^∞

Let $(X, \mathcal{M}, \mu)$ be a measure and L^∞ the set of measurable function such that $\|f\|_\infty < \infty$. Then:

1. If f, g are measurable functions on X, then
$$\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$$
If $f \in L^1$ and $g \in L^\infty$, then $\|fg\|_1 = \|f\|_1 \|g\|_\infty$ if and only if $|g(x)| = \|g\|_\infty$ a.e. on the set where $f(x) \neq 0$
2. $\|\cdot\|_\infty$ is a norm on L^∞
3. $\|f_n - f\|_\infty \to 0$ if and only if there exists $E \in \mathcal{M}$ such that $\mu(E^c) = 0$ and $f_n \to f$ uniformally on E
4. L^∞ is a Banach space
5. The simple functions are dense in L^∞

As a consequence of the well-definedness of L^∞, $1^{-1} + \infty^{-1} = 1$ is indeed a well-defined notion (since $\infty^{-1} = 1/\infty = 0$. Thus, 1 and ∞ are conjugate exponents!

Proof **:**

1. We'll be using the $\|f\|_= \inf_{\substack{E\in\mathcal{M}\\ \mu(X\setminus E)=0}} (\sup_{x\in E} |f(x)|)$ definition of the essential supremum to take advantage of the supremum within the definition. Our goal is to show that
$$\int |fg| = \|fg\|_1 \leq \|f\|_1 \|g\|_\infty$$
First, we will limit our attention of $\int |fg|$ to some $E \in \mathcal{M}$ such that $\mu(X \setminus E) = \mu(E^c) = 0$ (i.e., we are ignoring finite-measure points). Then:
$$\int |fg| = \int_E |fg| + \int_{E^c} |fg| = \int_E |fg|$$
Next, we take advantage the fact that we are working with an infimum to see that:
$$\|f\|_\infty \leq \sup_{x\in A} |f(x)| < \|f\|_\infty + \epsilon$$
Therefore, we can take:
$$\int_E |fg| \leq \int_E |\sup_{xinE} |f(x)| \cdot g| = \int_E \left|\sup_{x\in E} |f(x)|\right| \cdot |g| = \left|\sup_{x\in E} |f(x)|\right| \int_E |g|$$
where the last equality comes from the fact that the supremum is a constant. Finally, by the property of the infimum, we have:
$$\left|\sup_{x\in E} |f(x)|\right| \int_E |g| \leq (\|f\|_\infty + \epsilon) \int |g|$$

Notice that this inequality does not depend on ϵ, and so:

$$\|fg\|_1 = \int_E |fg| \leq \|f\|_\infty \|g\|_1$$

as we sought to show

4.1.1 Relating L^p spaces

In this section, we will show how we can relate L^p spaces with each other. Fundamentally, the way we get any order with L^p spaces is through Hölder's inequality. We take full advantage of it in the following proofs. The end-goal is to show that if $p < q < r$ where we allow $r = \infty$, then:

$$L^p \cap L^r \subseteq L^q \subseteq L^p + L^r$$

which at first glance might seem strange since it feels like the relation should be $L^p \subseteq L^q$ for $p < q$ (or the contravariant). However, as we saw in example 4.1, due to the varying types of growths a function can have (in particular, it can shrink too slowly to have a finite measure, or it explodes at some point, or a combination of both), the relationship is not so simple. We thus explore these relation's in this section.

Proposition 4.1.2: L^p subset

Let $0 < p < q < r \leq \infty$. Then

$$L^q \subseteq L^p + L^r$$

In particular, each $f \in L^q$ is the sum of a function in L^p and a function in L^q

***Proof* :**
very short. p.185

Proposition 4.1.3: L^p superset

Let $0 < p < q < r \leq \infty$. Then

$$L^p \cap L^r \subseteq L^q$$

In particular, if if we find that f is bounded in growth from above and bellow by p and q, $\|f\|_p < \infty$ and $\|f\|_r < \infty$, then f is bounded in growth by q.

***Proof* :**
Let $\|f\|_r < \infty$ and $\|f\|_q < \infty$. We need to show that $\|f\|_p < \infty$. Notice that it is equivalent to show that $\|f\|_r^n < \infty$ and $\|f\|_q^m < \infty$ then $\|f\|_p^\ell < \infty$ for $n, m, \ell \in \mathbb{R}_{\geq 0}$ (since if it's finite, then all powers will be finite). We'll prove it in the case where $r = \infty$ and $r < \infty$. In the infinite case, we simply take advantage of bounding properties of the essential supremum, and in the other case we use Hölder's inequality.

Starting with the $r = \infty$ case, notice that

$$|f|^q = |f|^{q-p+p} = |f|^q |f|^{q-p} \leq \|f\|_\infty^{q-p} |f|^p$$

and so, taking the integral and the qth root, we get:

$$\begin{aligned}
\|f\|_q &= \left(\int |f|^q\right)^{1/q} \\
&\leq \left(\int \|f\|_\infty^{q-p}|f|^p\right)^{1/q} \\
&= \|f\|_\infty^{\frac{q-p}{q}} \left(\int |f|^p\right)^{1/q} \\
&= \|f\|_\infty^{\frac{q-p}{q}} \left(\int |f|^p\right)^{p/pq} \\
&= \|f\|_\infty^{1-\frac{p}{q}} \|f\|_p^{p/q} < \infty
\end{aligned}$$

The fact that the last inequality is less than infinity comes from the that $\|f\|_p < \infty$ and $\|f\|_\infty < \infty$, so any power is less than infinity, and so $\|f\|_q < \infty$. Furthermore, notice that the sum of the power's is 1; this fact will come back for another intuition of $\|f\|_\infty$.

For the $r < \infty$ case the key is to find a way to integrate p and r using conjugate exponents. Since

$$p < q < r$$

then

$$\frac{1}{p} > \frac{1}{q} > \frac{1}{r}$$

therefore, define the continuous function $\lambda \mapsto \frac{1-\lambda}{p} + \frac{\lambda}{r}$. Since $[0, 1]$ is open, and $1/p > 1/r$, there exists a λ_0 such that

$$\frac{1-\lambda_0}{p} + \frac{\lambda_0}{r} = \frac{1}{q}$$

manipulating the equation around to get conjugate exponents, notice that:

$$\frac{(1-\lambda_0)q}{p} + \frac{\lambda_0 q}{r} = 1 \quad \Leftrightarrow \quad \frac{1}{\frac{p}{(1-\lambda_0)q}} + \frac{1}{\frac{r}{\lambda_0 q}} = 1$$

Thus, we have that $\frac{p}{(1-\lambda_0)q}$ and $\frac{r}{\lambda_0 p}$. With conjugate exponents established, we proceed with trying to bound $\|f\|_q$. As we commented, it is equivalent to bound $\|f\|_q^m$, so let $m = q$ (this will be a common trick to apply Hölder's inequality). Also, since $1 - \lambda_0 + \lambda_0 = 1$, we have that

$q = \lambda_0 q + (1-\lambda_0)q$, giving us the ability to split up our integral:

$$\begin{aligned}
\|f\|_q^q &= \int |f|^q \\
&= \int |f|^{\lambda_0 q + (1-\lambda_0)q} \\
&= \int |f|^{\lambda_0 q}|f|^{(1-\lambda_0)q} \\
&= \left\| |f|^{\lambda_0 q}|f|^{(1-\lambda_0)q} \right\|_1 \\
&\leq \left\| |f|^{\lambda_0 q} \right\|_{\frac{p}{(1-\lambda_0)q}} \left\| |f|^{(1-\lambda_0)q} \right\|_{\frac{r}{\lambda_0 q}} && \text{Hölder's inequality} \\
&= \left(\int |f|^p \right)^{\frac{(1-\lambda_0)q}{p}} \left(\int |f|^r \right)^{\frac{\lambda_0 q}{r}} \\
&= \|f\|_p^{(1-\lambda_0)q} \|f\|_r^{\lambda_0 q} < \infty
\end{aligned}$$

Since the last two terms are finite, any power of them are also finite, and so we have bounded $\|f\|_q^q < \infty$, and so $\|f\|_q < \infty$, as we sought to show.

Proposition 4.1.4: ℓ^p Covariant Inclusion

Let A. If $0 < p < q \leq \infty$, then $\ell^p(A) \subseteq \ell^q(A)$

Proof **:**
They key part is to take advantage of the properties of the counting measure. Since the only set of measure 0 in the counting measure is the empty-set, we have that the essential supremum is actually the supremum!

Let's say $f \in \ell^p$, so that $\sum_{a\in A} |f(a)|^p = \|f\|_p < \infty$. We'll start by showing that $\|f\|_\infty < \infty$. Since the $\|f\|_\infty$ is in fact the supremum, we have that:

$$\|f\|_\infty^p = \sup_{a\in A} |f(a)|^p \leq \sum_{a\in A} |f(a)|^p = \|f\|_p < \infty$$

Therefore, $\|f\|_\infty < \infty$.

Proposition 4.1.5: L^p Contravariant Inclusion

Let $(X, \mathcal{M}, \mu)$ be a measure space where $\mu(X) < \infty$. If $0 < p < q \leq \infty$, then $L^p \supseteq L^q$

Proof **:**
We need to show that if $f \in L^q$ then $f \in L^p$. Since $f \in L^q$, $\|f\|_q < \infty$, so $\|f\|_q^n < \infty$ for all $n \in \mathbb{N}$. If we can show that $\|f\|_p^m \leq \|f\|_q^n$ for some $m, n \in \mathbb{N}$, then the proof is complete.

We'll apply Hölder's inequality in the caes of $q = \infty$ and $q < \infty$. If $q = \infty$, then:

$$\|f\|_p^p = \int |f|^p = \int |f|^p \cdot 1 \overset{\text{H.}}{\leq} \|f\|_\infty^p \int 1 = \|f\|_\infty^p \mu(X) < \infty$$

since $\int 1 = \mu(X)$. If $q < \infty$, then we do a similar argument. First, since $p < q$, notice that:

$$\frac{1}{\frac{q}{p}} + \frac{1}{\frac{q}{q-p}} = \frac{p}{q} + \frac{q-p}{q} = \frac{q}{q} = 1$$

and so $\frac{q}{p}$ and $\frac{q}{p-q}$ are conjugate exponents. Therefore, by Hölder's inequality:

$$\begin{aligned}
\|f\|_p^p &= \int |f|^p \\
&= \int |f|^p \cdot 1 \\
&= \||f|^p \cdot 1\|_1 \\
&\leq \||f|^p\|_{\frac{q}{p}} \|1\|_{\frac{q}{p-q}} && \text{Hölder's Inequality} \\
&= \left(\int |f|^{p\frac{q}{p}}\right)^{\frac{p}{q}} \left(\int 1\right)^{\frac{q-p}{p}} \\
&= \|f\|_q^p \, \mu(X)^{\frac{q-p}{p}} < \infty
\end{aligned}$$

therefore, $\|f\|_p^p < \infty$, so $\|f\|_p < \infty$, showing that $f \in L^p$, as we sought to show.

Proposition 4.1.6: Relating L^p and L^∞

Let $f \in L_1$ and $f \in L^\infty$, then:

$$\lim_{p \to \infty} \|f\|_p = \|f\|_\infty$$

Proof **:**
here

(write this somewhere)

1. L^1 is natural for integration
2. L^∞ natural for pointwise questions
3. L^p interpolates in between (L^2 particularly important, will let us get an inner product!)