Real Analysis Oral Exam study notes

Notes transcribed by Mihai Nica

Abstract. These are some study notes that I made while studying for my oral exams on the topic of Real Analysis, mostly covering the theory of integration. I took these notes from parts of the textbooks by Richard Bass [1] and a few other sources which are indicated throughout. Please be extremely caution with these notes: they are rough notes and were originally only for me to help me study and are not complete or guaranteed to be free of errors. I have made them available to help other students on their oral exams.

Contents

Families of sets

These are notes from Chapter 2 of [**1**].

1.1. Algebras and $\sigma -$ algebras

Fix a universal set X to work in.

DEFINITION. (2.1) An **algebra** is a collection $\mathcal{A}$ of subsets of X that:
i) $\emptyset, X \in \mathcal{A}$
ii) $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
iii) closed under finite unions and intersections
A $\sigma-$**algebra** has:
iv) closed under *countable* unions and intersections.

LEMMA. *(2.7.) Any arbitrary intersection of $\sigma-$algebras is a sigma algebra*

DEFINITION. We define the $\sigma-$algebra **generated** by a subset $\mathcal{C}$ to be the smallest $\sigma -$ algebra cotaining $\mathcal{C}$: i.e. the intersection of all $\sigma -$ algebras containing $\mathcal{C}$. This is denoted by $\sigma(\mathcal{C})$.

DEFINITION. The Borel $\sigma -$ algebra is the $\sigma -$ algebra that is generated by any of the following: 1) open intervals 2) closed intervals 3) half open intervals 4) semi-infinite open intervals $(a, \infty]$

PROOF. Its easy to check by doing some intersections unions that these all generate the same thing. $\square$

1.2. Monotone Class Theorem

DEFINITION. (2.9.) A **monotone class** is a collection of subsets of X that is closed under *increasing*, and *decreasing* sets. I.e. $A_n \uparrow A \implies A \in \mathcal{M}$ and $A_n \downarrow A \implies A \in \mathcal{M}$.

THEOREM. *(2.10) Suppose $\mathcal{A}_0$is an algebra and, $\mathcal{A} = \sigma(\mathcal{A}_0)$ is the smallest $\sigma -$ algebra containing $\mathcal{A}_0$ and $\mathcal{M}$ is the smallest monotone class containing $\mathcal{A}_0$.*
Then $\mathcal{M} = \mathcal{A}$.
(Mihai's version) In other words: for an algebra $\mathcal{A}_0$, the smallest monotone class containing $\mathcal{A}_0$ is in fact, a $\sigma -$ algebra.

PROOF. We must show $\mathcal{A} \subset \mathcal{M}$ as the other inclusion is clear. To do this we have to show that the monotone class is closed under intersections and unions, for then $\mathcal{M}$ will be an algebra containgn $\mathcal{A}_0$ and hence $\mathcal{A} = \sigma(\mathcal{A}_0) \subset \mathcal{M}$.
The idea of the proof is the usual trick with sigma algebras and so on: Define the set you are interested in and show it contains what you want.
<u>Claim 1:</u> Closed under complements

Let $\mathcal{N} = \{A \in \mathcal{M} \ : \ A^c \in \mathcal{M}\}$. Since A_0 is an algebra and closed under complements, $A_0 \subset \mathcal{N}$. We claim now that $\mathcal{N}$ is a monotone class, for if $A_n \uparrow A$ with $A_n \in \mathcal{N}$ then by def'n $A_n^c \in \mathcal{M}$ and we have $A_n^c \downarrow A^c$ shows $A^c \in \mathcal{M}$ since $\mathcal{M}$ is a monotone class. Similar for $A_n \downarrow A$. Hence $\mathcal{N}$ is a monotone calss containg A_0 and consequently $\mathcal{M} \subset \mathcal{N}$ i.e. $\mathcal{M} = \mathcal{N}$.

Claim 2: Closed under intersections with $\mathcal{A}_0$

Let $\mathcal{N} = \{A \in \mathcal{M} \ : \ A \cap B \in \mathcal{M} \ \forall B \in \mathcal{A}_0\}$. Again $\mathcal{N}$ contains $\mathcal{A}_0$ since $\mathcal{A}_0$ is an algebra, and $\mathcal{N}$ is a monotne class since if $A_n \uparrow A$ then $A_n \cap B \uparrow A \cap B$ for every $B \in A_0$, so the monotone-class-ness of $\mathcal{M}$ gives us $A \cap B \in \mathcal{M}$ and this holds for every $B \in \mathcal{A}_0$. (same idea for $A_n \downarrow A$). As before, we get $\mathcal{N} = \mathcal{M}$.

Claim 3: Closed under intersections

Let $\mathcal{N} = \{A \in \mathcal{M} \ : \ A \cap B \in \mathcal{M} \ \forall B \in \mathcal{M}\}$. This contain $\mathcal{A}_0$ by the previous claim! It is a monotone class by the same argument as in claim 2. Hence, we have $\mathcal{M} = \mathcal{N}$.

This shows that $\mathcal{M}$ is in fact a $\sigma -$ algebra (finite unions/intersections + monotone limits yields countable unions/intersections by splitting up any countable union/intersection up into a monotne limit of finite unions/intersections) $\square$

REMARK. The sneaky thing in the proof is that you can't jump to closed under intersections right away, because you don't know that $\mathcal{A}_0$ is in there a priori. The little extra step is to first prove that $\mathcal{A}_0$ is closed under intersections from $\mathcal{M}$ and then move on to the whole space.

1.3. Pi-Lambda Theorem

I'm going to include a little thinger on the Pi-Lambda theorem, as this comes up occasionally in Probability theory.

DEFINITION. A $\pi-$**system** is a collection of subsets which is closed under finite intersections.

DEFINITION. A $\lambda-$**system** $\mathcal{C}$ is a collection of subsets which has:
i) $X \in \mathcal{C}$
ii) $A, B \in \mathcal{C}$ and $A \subset B \implies B - A \in \mathcal{C}$
iii) $A_n \in \mathcal{C}$, and $A_n \uparrow A \implies A \in \mathcal{C}$

PROPOSITION. *(Alternative definition of a $\lambda-$system)*
a) $X \in \mathcal{C}$
b) $A \in \mathcal{C} \implies A^c \in \mathcal{C}$
c) If $A_n \in \mathcal{C}$ *are disjoint then* $\dot{\bigcup}_{n=1}^{\infty} A_n \in \mathcal{C}$

PROOF. Its easy to see that i), ii) and iii) imply a,b,c. We prove the converse statements one at a time assuming a), b) and c):

ii): For any A, B with $A \cap B = \emptyset$, we have that:

$$\begin{aligned} A \cup B &\in \mathcal{C} \\ \implies A^c \cap B^c &\in \mathcal{C} \end{aligned}$$

But $A^c \cap B^c = A^c - B$ so this shows that whenver $A \cap B = \emptyset$, we have $A^c - B \in \mathcal{C}$. Noticing that $A \cap B = \emptyset \iff A \subset B^c$ and relabeling $B^c = C$ gives the result ii)

iii) Take set differences $B_n = A_n - A_{n-1}$which are disjoint. $\square$

THEOREM. *If* $\mathcal{C}$ *is both a* $\pi-$*system and a* $\lambda-$*system, then* $\mathcal{C}$ *is a* $\sigma-$*algebra*

PROOF. For any arbitarty sequence $A_n \in \mathcal{C}$, we can create $B_n \in \mathcal{C}$ which are disjoint with $\cup_{k=1}^{n} B_k = \cup_{k=1}^{n} A_k$ by doing intersections (ok since $\mathcal{C}$ is a $\pi-$system) and complements (ok since $\mathcal{C}$ is a λ-system). Then, since $\mathcal{C}$ is a $\lambda-$system, we have that $\cup_{k=1}^{\infty} B_k \in \mathcal{C}$ and so $\cup_{k=1}^{\infty} A_k = \cup_{k=1}^{\infty} B_k \in \mathcal{C}$ too! □

THEOREM. *(7.4.) (Dynkin's $\pi - \lambda$ theorem) Suppose $\mathcal{P}$is a π-system and, $\mathcal{A} = \sigma(\mathcal{P})$ is the smallest $\sigma -$ algebra containing $\mathcal{P}$ and $\mathcal{L}$ is the smallest monotone class containing $\mathcal{P}$.*

Then $\mathcal{L} = \mathcal{A}$.

(Mihai's version) In other words: for a $\pi-$system $\mathcal{P}_0$, the smallest $\lambda-$system containing $\mathcal{P}_0$ is in fact, a $\sigma -$ algebra.)

REMARK. This is very similar in feel to the monotone class theorem, it has the same "two step" trick to it.

PROOF. We will show that $\mathcal{L}$ is a σ-algebra. By the above theorem, it suffices to show that is is a π-system, so we will do this easier task instead.

Claim 1: For a fixed $C \in \mathcal{L}$, define $\mathcal{L}^C := \{D \in \mathcal{L} : C \cap D \in \mathcal{L}\}$. Then $\mathcal{L}^C$ is a λ-system.

Pf: Use the first def'n of $\lambda-$system (because this one is easier to check) and the result follows wihtout too much difficutly because $\mathcal{L}$ itself is a $\lambda-$system.

Now, choose any $C \in \mathcal{P}$. Since $\mathcal{P}$ is a $\pi-$system, it is closed under intersections and we know at least that $\mathcal{P} \subset \mathcal{L}^C$ since $\mathcal{L}$ contains $\mathcal{P}$. Now, since $\mathcal{L}^C$ is a $\lambda-$system containing $\mathcal{P}$, we have that $\mathcal{L} = \mathcal{L}^C$ since $\mathcal{L}$ is the smallest such λ-system.

Since $\mathcal{L}^C = \mathcal{L}$ for every $C \in \mathcal{P}$, this is saying that $C \cap D \in \mathcal{L}$ for every $C \in \mathcal{P}$ and every $D \in \mathcal{L}$.

Now for any $E \in \mathcal{L}$, we claim that $\mathcal{L}^E$ contain $\mathcal{P}$. Indeed, to check this we would need $E \cap C \in \mathcal{L}$ for every $C \in \mathcal{P}$, but this is exactly the above property. Hence $\mathcal{L}^E = \mathcal{L}$.

Since this holds for all $E \in \mathcal{L}$, we have that $\mathcal{L}$ is in fact a $\pi-$system. □

REMARK. There should be a nicer way to make the proofs of the monotone class theorem and the $\pi - \lambda$ theorem look more similar, but I'm not going to do that now.

Measures

These are notes from Chapter 3 of [**1**].

Definition. (3.1) A **measure** on a set X and an $\sigma-$algebra$\mathcal{A}$ is a a function $\mu:\mathcal{A}\to[0,\infty]$ such that:

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

ii) Countably additive for disjoint sets:

$$\mu(\dot{\bigcup}_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}\mu(A_i)\text{ for disjoint }A_i's$$

Proposition. *(3.5.) (Basic Properties of Measures)*

i) $A\subset B\implies\mu(A)\leq\mu(B)$

ii) $A=\cup_{i=1}^{\infty}A_i\implies\mu(A)\leq\sum_{i=1}^{\infty}\mu(A_i)$

iii) $A_i\uparrow A\implies\lim_{n\to\infty}\mu(A_i)=\mu(A)$

iv) If $\mu(A_1)<\infty$, *then* $,A_i\downarrow A\implies\lim_{n\to\infty}\mu(A_i)=\mu(A)$

Proof. These all follow by doing constructions involving complements and set differences and so on in such a way as to reduce the sets in question as unions of disjoint sets, for which we know that measures play nice.

Most of them use the "Treat Disjointly Trick" (See remark below)

i) Holds since $\mu(A-B)\geq0$

ii) Make $B_i\subset A_i$ by taking set differences so that $\cup_{i=1}^{n}B_i=\cup_{i=1}^{n}A_i$ for each n and the B_i's are disjoint. Then $\mu(A)=\mu(\cup B_i)=\sum\mu(B_i)\leq\sum\mu(A_i)$

iii) $\mu(A_n)$ is a montone increasing function and is bounded above by $\mu(A)$, so it is convergent (to possibly ∞if $\mu(A)=\infty$). Hence, by the monotone convergence theorem for real numbers, this sequence has a limit. To see that the limit is actually $\mu(A)$, take $B_i=A_{i+1}-A_i$ so the B_i's are disjoint and have $\cup_{i=1}^{n}B_n=\cup_{i=1}^{n}A_i$ and apply the countable additivity to get the result. Explicitly:

$$\begin{aligned}\mu(A)&=\mu(\cup_{n=1}^{\infty}A_n)\\&=\mu(\cup_{n=1}^{\infty}B_n)\\&=\sum_{n=1}^{\infty}\mu(B_n)\\&=\lim_{n\to\infty}\sum_{i=1}^{n}\mu(B_i)\\&=\lim_{n\to\infty}\mu\left(\cup_{i=1}^{n}B_i\right)\\&=\lim_{n\to\infty}\mu\left(\cup_{i=1}^{n}A_i\right)\end{aligned}$$

iv) Take complements of the result in iii) (or, more generally if you arent in a fintis measures space$B_k=A_1-A_k$). The hypothesis that $\mu(A_1)<\infty$ is needed to avoid a "$\infty - \infty$" problem. Counterexample to keep in mind: $A_n = (n, \infty)$ which is decreasing to the empty set but always has infinte Lebesgue measure. □

REMARK. A common theme that comes up for this type of thing: The proof is easy for disjoint sets. The arbitary case can be reduced to the disjoint case by making sets B_i which are disjoint and have $\cup_{i=1}^{n} B_i = \cup_{i=1}^{n} A_i$ for every n. I will label this trick as [Treat Disjointly]

DEFINITION. (3.7.) Definition of finite and σ-finite measure spaces. A **null set** is a set which is a subset of a 0 measure measurable set. (i.e. null sets need not be measurable) A **complete measurable space** is one where all null sets are measurable. The completion of $\mathcal{A}$ is the smallest sigma algebra containing all the null sets.

Construction of Measures

These are notes from Chapter 4 of [**1**].

Recall the diffuculty with measures is that if you try to define them on all subsets of a space X, you get problems. The solution is to carefully choose the right $\sigma-$algebra to define the measure on. As a consequence, "measurable sets" will be nice enough to work with! This is outlined here.

3.4. Outer Measures

Definition. (4.1.) An **outer measure** is a function μ^* defined on a collection of subsets of X satisfying:
i) $\mu^*(\emptyset) = 0$
ii) $A \subset B \implies \mu^*(A) \subset \mu^*(B)$
iii) $\mu^*(\cup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \mu^*(A_i)$

Remark. There is no notion of "measurable" or σ-algebra or whatever for an outer measure. It has to be defined for every subset. It is called an "outer measure" because of the following common way to construct an outer measure by approximation with sets "from the outside". (Sometimes you'll see the construction here as the definition of an outer measure, and the properties we demand in the definition here would be something to be proven)

Proposition. *(4.2.) Suppoce $\mathcal{C}$ is a collection of subsets of X containg both $\emptyset$ and X. Suppse that $\ell : \mathcal{C} \to [0, \infty]$ with $\ell(\emptyset) = 0$. Define:*

$$\mu^*(E) = \inf \left\{ \sum_{i=1}^{\infty} \ell(C_i) : C_i \in \mathcal{C} \forall i \text{ and } E \subset \cup_{i=1}^{\infty} C_i \right\}$$

Then $\mu^*$is an outer measure.

Proof. The points i) and ii) from the defintion are very easy. To prove iii) consider as follows. Given any collection of subsets $A_1, A_2, \ldots$ and any $\epsilon > 0$, find a collection of subsets $C_{ij} \in \mathcal{C}$ so that for each fixed i the$C_{ij}'s$ cover A_i and$\mu^*(A_i) \ge \sum \ell(C_{ij}) + \epsilon/2^i$ for each i. (This is by definition of "inf"). Notice that the collection C_{ij} now covers all of $\cup A_i$ and so we have that:

$$\begin{aligned} \mu^*(\cup A_i) &\le \sum_{ij} \ell(C_{ij}) \\ &\le \ldots \le \sum_{i=1}^{\infty} \mu^*(A_i) + \epsilon \end{aligned}$$

Since this holds for any ϵ, the desired inequality must be true. $\square$

EXAMPLE. Lebesgue outer measure is what you get when you put in $\mathcal{C}$ =half open intervals and $\ell(a,b] = b - a$. The Lebesgue-Stieltjes outer measure is what you get if you again choose $\mathcal{C}$ =half open intervals and you have a non-decreasing and right continuous function α.

DEFINITION. (4.5) A set $A \subset X$ is **measurable with respect to an outer measure** μ^* (or more simply μ^*-measurable....I like the first phrasing better because it reminds me that outer measures are defined on all subsets of a space) if:

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \ \forall E \subset X$$

REMARK. By the definition of an outer measure, it is ALWAYS TRUE for any set A that:

$$\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c) \ \forall E \subset X$$

So the property of being measurable is really about the other inequality. This is most often how one proves certain sets are measurable.

One way to rememember the direction of the sign is to think that "$E \cap A$ and $E \cap A^c$ together form a cover of E. Hence, since $\mu^*(E)$ is the inf over the measur of all covers, $\mu^*(E) \leq$the sum". Of coruse, this is a mnemonic device only since stricly speaking the measure μ^* is defined only as the inf over covers formed by the special sets $\mathcal{C}$.

THEOREM. *(4.6.) Let $\mathcal{A} = \{measurble\ w.r.t.\ \mu^*\ sets\}$ .Then $\mathcal{A}$ is a $\sigma-$algebra. If we define $\mu := \mu^*|_{\mathcal{A}}$, then μ is a measure on $\mathcal{A}$. Finally, $\mathcal{A}$ contains all the sets of outer measure* 0.

REMARK. One of the main themes to keep in mind for this proof is that we are trying hard to get the inequalities of the form $\mu^*(E) \geq \ldots$, because this is the non-trivial inequality for being measurable. Where will these inequalities come from? The only way is to use the trivial side of the inequality in a clever way: you must manipulate $\mu^*(E) = \ldots = \mu^*(U \cap V) + \mu^*(U \cap V^c)$ and then remark $\mu^*(U \cap V) + \mu^*(U \cap V^c) \geq \mu^*(U)$ since this always holds. (i.e. use property iii) from the definition of outer measure.)

PROOF. Claim 1: $\mathcal{A}$ is closed under complements.

Pf: Follows by the symmetry of the definition of measurable w.r.t. μ^*.

Claim 2: $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$

Pf: For any E, use the measurablity property of A at the set E and use the measurability property of the set B once at the set $E \cap A$ and once at the set $E \cap A^c$:

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= (\mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \\ &\quad (\mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c))) \end{aligned}$$

Now notice that the first three terms are $E \cap A \cap B,\ E \cap A^c \cap B,\ E \cap A \cap B^c$. Since $A \cup B = (A \cap B) \cup (A \cap B^c) \cup (B \cap A^c)$, so consequently the first three sets here cover $E \cap (A \cup B)$. Hence the sum of their measures is $\geq \mu^*(E \cap (A \cup B))$ and we get (alspo use $A^c \cap B^c = (A \cup B)^c$by DeMorgan):

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

So indeed, $A \cup B$ satisfy the non-trivial direction of the measurable inequality, and are hence measurable.

Claim 3: $A_n \in \mathcal{A} \implies \cup_{n=1}^{\infty} A_n \in \mathcal{A}$

Pf: We firstly remark that is suffices to check this for disjoint sets A_n, because by using the property of Claim 2, we can rewrite any countable union as a countable union of disjoint sets modulo some finite union operations.

For disjoint sets A_n let $B_n = \cup_{i=1}^{n} A_i$ and $B = \lim_{n\to\infty} B_n$. For $E \subset X$ have:

$$\begin{aligned}\mu^*(E \cap B_n) &= \mu^*(E \cap B_n \cap A_n) + \mu^*(E \cap B_n \cap A_n^c) \\ &= \mu^*(E \cap A_n) + \mu^*(E \cap B_{n-1})\end{aligned}$$

Moving the $\mu^*(E \cap B_{n-1})$ to the other side, and then using this as a telescoping sum, we get that:

$$\begin{aligned}\mu^*(E \cap B_n) - \mu^*(E \cap B_1) &= \sum_{i=1}^{n} \mu^*(E \cap A_i) \\ \implies \mu(E \cap B_n) &\geq \sum_{i=1}^{n} \mu^*(E \cap A_i)\end{aligned}$$

Hence:

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

Taking $n \to \infty$, (ok since the sum is monotone increasing) we get:

$$\begin{aligned}\mu^*(E) &\geq \sum_{i=1}^{\infty} \mu^*(E \cap A_i) + \mu^*(E \cap B^c) \\ &\geq \mu^*\left(\cup_{i=1}^{\infty} E \cap A_i\right) + \mu^*(E \cap B^c) \\ &= \mu^*(E \cap B) + \mu^*(E \cap B^c) \\ &\geq \mu^*(E)\end{aligned}$$

So indeed B is μ^*-measurable!

These claims together show that $\mathcal{A}$ is a sigma algebra.

Claim 4: $\mu := \mu^*|_{\mathcal{A}}$ is a measure.

Pf: We have only to show that it is countable additive. Following the discussion above, we had that $\mu^*(E) = \sum_{i=1}^{\infty} \mu^*(E \cap A_i) + \mu^*(E \cap B^c)$ for any set E. Taking $E = B$ gives exactly the result for countable additivity.

Claim 5: If $\mu^*(A) = 0$ then A is μ^*measurable.

Pf: This follows from the monotone property of the outer measure, for if $\mu^*(A) = 0$ then:

$$\begin{aligned}\mu^*(E \cap A) + \mu^*(E \cap A^c) &\leq \mu^*(A) + \mu^*(E \cap A^c) \\ &\leq \mu^*(A) + \mu^*(E) \\ &= \mu^*(E)\end{aligned}$$

Shows the non-trivial direction of the measurability inequality. □

3.5. Lebesgue-Stieltjes Measure

Let $\alpha(x)$ be an increasing right continuous function, let $\mathcal{C} = \{(a, b] : a, < b\}$ and define $\ell((a, b]) = \alpha(b) - \alpha(a)$. Then define the outer measure:

$$m^*(E) = \inf\left\{\sum \ell(A_i) : A_i \text{ cover } E\right\}$$

By Prop 4.2 this is an outer measure, and by theorem 4.6 this defines a measure on the collection of m^*-measurable sets.

The convenience of using semi open intervals is due to the following fact:

LEMMA. *If K and L are disjoint semi open intervals, (of the form $(a,b]$) and $K \cup L$ is also a semi open interval then:*

$$\ell(K) + \ell(L) = \ell(K \cup L)$$

PROOF. Both are equal to $\alpha(c) - \alpha(a) = (\alpha(c) - \alpha(b)) + (\alpha(b) - \alpha(a))$ where a, b, c denote the endpoints of the intevrals in question. □

What kinds of sets are m^*-measurable? Here is a useful fact!

PROPOSITION. *(4.7.) Every set in the Borel $\sigma-$algebra on $\mathbb{R}$ is m^*-measurable.*

PROOF. Since the collection of m^*-measurable sets is a $\sigma-$algebra (proven in Prop 4.6) it suffices to show that every interval J of the form $(c,d]$ is m^*-measurable. Let E be any set with $m^*(E) < \infty$. We need to show that:

$$m^*(E) \geq m^*(E \cap J) + m^*(E \cap J^c)$$

Choose $I_1, I_2, \ldots$ of the form $(a_i, b_i]$ that cover J and so that $m^*(E) \geq \sum_i [\alpha(b_i) - \alpha(a_i)] - \epsilon$. Now, since $E \subset \cup I_i$ we have that:

$$\begin{aligned} m^*(E \cap J) &\leq \sum_i m^*(I_i \cap J) \\ m^*(E \cap J^c) &\leq \sum_i m^*(I_i \cap J^c) \end{aligned}$$

So we get in sum that:

$$m^*(E \cap J) + m^*(E \cap J^c) \leq \sum_i m^*(I_i \cap J) + m^*(I_i \cap J^c)$$

Now we get to the convenience of using semi-open intervals! Since each I_i and J are semi open intervals, $I_i \cap J$ is an interval too! $I_i \cap J^c$ is the union of zero, one or two such intervals. This is exactly the setup for us to apply the Lemma before. Applying this zero, one or two times we have:

$$m^*(I_i \cap J) + m^*(I_i \cap J^c) = m^*(I_i)$$

Thus our inequality above is:

$$m^*(E \cap J) + m^*(E \cap J^c) \leq \sum_i m^*(I_i) \leq m^*(E) + \epsilon$$

Since this holds for any $\epsilon > 0$, we have the desired inequality. □

The next thing to verify is that m^* and ℓ agree on semi-open sets. This is not at all surprising, but it needs to be verified.

LEMMA. *(4.8.) Let $J_k = (a_k, b_k)$ be a collection of finte open intervals which covera finite closed interval $[C, D]$ then:*

$$\sum_{k=1}^{n} [\alpha(b_k) - \alpha(a_k)] \geq \alpha(D) - \alpha(C)$$

PROOF. By looking at a subset of the intervals, we may suppose WOLOG that the intervals are "in order" so to speak so that:

$$a_1 \leq C \leq b_1 \text{ and } , a_n \leq D \leq b_n \text{ and } a_k < b_{k-1} < b_k$$

Then just write out both sums and see by comparison to a telescoping sum that the inequality holds. □

PROPOSITION. *(4.9) If e and f are finite and $I = (e, f]$ then $m^*(I) = \ell(I)$*

PROOF. (This is the one that uses compactness of closed intervals. The idea is to convert any cover by semi-open intervals to a cover using open intervals (this gives an ϵ of error by the continuity of α) and then use compactness to get down to finitely many intervals of interest. The apply the previous lemma.)

Clearly $m^*(I) \leq \ell(I)$ by the inf definition of m^* since I is a cover for itself. For the other inequalty, suppose that $I \subset \cup A_i$ where each $A_i = (c_i, d_i]$ is a an interval. By the right continuity of the function α, choose $C \in (e, f)$ so that $\alpha(C) < \alpha(e) + \epsilon/2$. Let $D = f$. Choose $d_i^{'} > d$ so that $\alpha(d_i^{'}) < \alpha(d_i) + \epsilon/2^{i+1}$ and let $B_i = (c_i, d_i')$

(Each (c_i, d_i') aproximates the semi open interval $(c_i, d_i]$ and the whole thing is done so that the total error from the point of view of the α function is no more than ϵ.)

Now use the compactness of $[C, D]$ to find a finite set of the B_i's that cover all of $[C, D]$ still. By the previous lemma we get that:

$$\ell(I) \leq \alpha(D) - \alpha(C) + \epsilon/2 \leq \sum \ell(A_i) + \epsilon/2^{i+1} + \epsilon/2$$

and since ϵ is arbitary the result follows. □

Now that all the major points have been hit, we will drop the $*$ and refer to m as the Lebesgue-Stieltjes measure corresponding to α.

3.6. Examples and related results

3.6.1. Approximating Sets.

EXAMPLE. (4.10) For the Lebesgue measure its not hard to show that singletons are measure 0 and so the open, half-open, closed, half-closed intervals all have the same measure. All countable sets are Lebesgue measure 0 too.

EXAMPLE. (4.11) The middle thirds Cantor set is uncontable but still has Lebesgye measure 0. The Cantor ternary function is also breifly described in this example.

EXAMPLE. (4.13) By changing the fraction of removed set as you go, you can end up with a positive measure generalized Cantor set. This set is closed and contains no intervals, and every point is a limit point.

PROPOSITION. *(4.14) Let $A \subset [0, 1]$ be a Borel measurable set and let m be the Lebesgue measure. Then:*
i) Given any $\epsilon > 0$ there is an open set G so that $A \subset G$ and $m(G - A) < \epsilon$
ii) Given any $\epsilon > 0$ there is a closed set F so that $F \subset A$ and $m(A - F) < \epsilon$
iii) There is a G_δ set H so that $A \subset H$ and $m(H - A) = 0$
iv) There is an F_σ set F so that $F \subset A$ and $m(F - A) = 0$

REMARK. We used half open sets in our construction of the Riemann-Stietijes measure, so we need to adapt from our half open intervals to the open intervals that make up open sets. If one had used open intervals to construct the measure (which is totally legit btw) then this proposition would be straight from the inf definition of m^*; but some early results (mostly the proof that $m^* = \ell$ for intervals would be slightly harder)

PROOF. i) Approximate A be a union of semi-open intervals so that the error is no more than $\epsilon/2$ (this holds by the inf definition of m^*) Then approximate each semi-open interval by an open interval so that th error is no more than $\epsilon/2^{n+1}$.

ii) Take complements and use the result in i)

iii) Choose $\epsilon = \frac{1}{n}$ and apply the result in i) to get a sequence of open set G_n. WOLOG they are decreasing. Take their intersection to get the desired $G_\delta$set.

iv) Same procedure as in iii) applied to closed sets F_n generated from $\epsilon = \frac{1}{n}$ in part ii) □

3.7. Nonmeasurable Sets

THEOREM. *(4.15) [Existence of a non-measurable set] Let m^*be the outer measure defined in the usual way with the collection $\mathcal{C} = \{(a,b]\}$ and $\ell((a,b]) = b - a$. The m^* is not a measure on the collection of all subsets of $\mathbb{R}$.*

REMARK. Since we showed that m^*IS a measure when restricted to the m^*measurable sets, this is saying that there are some sets that are not m^*-measurable, and that these sets are badly behaived enough that they prohibit $m^*$from being a measure here.

PROOF. (Here we have the construction using shifting by rationals. There is also the "$x \sim y$ if $x - y = k\alpha \mod 1$ proof where α is irrational" that works in a someqhat similar way.)

Suppose m^* is a measure. Define $x \sim y$ if $x - y$ is a rational. This is an equivalence relation. Use the axiom of choice to get a representative from each equivalence class from the points in $[0,1]$. Let A be the set of these representatives. Clearly:

$$[0,1] \subset \cup_{q \in [-1,1] \cap \mathbb{Q}} A + q$$

because every point in $[0,1]$ belongs to some equivalence class, and is hence a rational shift away from some point in A by definition. Moreover, each set $A + q_1$ is disjoint from $A + q_2$ for $q_1 \neq q_2$ for if there was a point w in there intersection then $w - q_1 \in A$ and $w - q_2 \in A$ are both in the same equivalence class and this contradicts the choice of A. Hence:

$$\begin{aligned} 1 &\leq \sum_q m^*(A+q) \\ &= \sum_q m^*(A) \end{aligned}$$

($m^*(A) = m^*(A+q)$ comes from the fact that $\ell(I) = \ell(I+q)$ and the definition of m^*). On the other hand, $\cup_{q \in [-1,1] \cap \mathbb{Q}} A + q \subset [-1,2]$ since $A \subset [0,1]$. So we get:

$$1 \leq \sum_q m^*(A) \leq 3$$

There is no value for $m^*(A)$ that satisfies these inequalites! Either $m^*(A) = 0$ and the lower bound is violated, or $m^*(A) > 0$ and the upper bound is viloated. □

3.8. The Caratheodory Extension Theorem

DEFINITION. Recall the definition of an algebra. (It was like a $\sigma-$algebra but fintie unions instead of countable ones.) A **premeasure** on an algebra $\mathcal{A}_0$ is a function $\mu : \mathcal{A}_0 \to [0,\infty]$ satisfying:
i) $\mu(\emptyset) = 0$
ii) Countably additive for disjoint sets that happen to stay in the algebra:
If $A_1, \ldots$ are disjoint and it happens that $\cup_{i=1}^{\infty} A_i \in \mathcal{A}_0$, then

$$\mu(\dot{\bigcup}_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

The next theorem repackages the constructuion of the Lebesgue-Stieltjes measure in a slightly more general framework:

THEOREM. *(4.16) [CARATHEODORY EXTENSION THEOREM] Suppose that $\mathcal{A}_0$ is an algebra and that $\ell : \mathcal{A}_0 \to [0,\infty]$ is a premeasure on $\mathcal{A}_0$. Define:*

$$\mu^*(E) = \inf\left\{\sum_{i=1}^{\infty} \ell(A_i) : \text{ each } A_i \in \mathcal{A}_0 \text{ cover } E\right\}$$

Then:
1) μ^* is an outer measure
2) $\mu^*(A) = \ell(A)$ if $A \in \mathcal{A}_0$
3) Every set in $\mathcal{A}_0$ is μ^*-measurable
4) If $\ell$is $\sigma-$finite then there is a unique extension to $\sigma(\mathcal{A}_0)$

PROOF. (1) follows by proposition 4.2
(2) $\mu^*(E) \le \ell(E)$ for $E \in \mathcal{A}_0$ by the inf def'n of μ^* since E covers itself. The other inequality holes because if $E \subset \cup A_i$, then we can let $B_n = E \cap \big(A_n - \cup_{i=1}^{n-1} A_i\big)$ so that the $\cup^n B_k = \cup^n A_k$ and since the B_i's are pairwise disjoing and each in $\mathcal{A}_0$ and their union is E which also happens to be in $\mathcal{A}_0$, so by the premeasure property:

$$\ell(E) = \sum_{i=1}^{\infty} \ell(B_i) \le \sum_{i=1}^{\infty} \ell(A_i)$$

So clearly then by taking inf's we have that$\ell(E) \le \mu^*(E)$
(3) To see that every set $A \in \mathcal{A}_0$ is μ^*measurable, consider as follows. For any $E \subset X$, take $B_1, B_2, \ldots \in \mathcal{A}_0$ that cover E and $\sum \ell(B_i) \le \mu^*(E) + \epsilon$. Then, since $\ell(B_i) = \ell(B_i \cap A) + \ell(B_i \cap A^c)$ because ℓ is a premeasure and $A, B \in \mathcal{A}_0$. Have then:

$$\begin{aligned}\mu^*(E) + \epsilon &\ge \sum_{i=1}^{\infty} \ell(B_i) = \sum_{i=1}^{\infty} \ell(B_i \cap A) + \sum_{i=1}^{\infty} \ell(B_i \cap A^c) \\ &\ge \mu^*(E \cap A) + \mu^*(E \cap A^c)\end{aligned}$$

Since ϵ arbitrary, we get the desired non-trivial half of the measurability inequality. □

(4) The fact that the measure exists follows from our previous work: We know that $\mathcal{A} = \{\mu^* - \text{measurable sets}\}$is a $\sigma-$algebra, it contains $\mathcal{A}_0$ by (3), and we showed earlier that $\mu = \mu^*|_{\mathcal{A}}$ is a measure. Uniqueness is most easily done with the $\pi - \lambda$ theorem:

If μ, ν are two measures so that $\mu(A) = \nu(A) = \ell(A)$ for all $A \in \mathcal{A}_0$, let $\mathcal{P} = \{A : \mu(A) = \nu(A)\}$. This contains the algebra $\mathcal{A}_0$ by construction, and we verify that it is a $\lambda-$system. Hence, by the $\pi - \lambda$ theorem, it is in fact a σ-algebra! Since it contains $\mathcal{A}_0$, it must in fact be all of $\sigma(\mathcal{A}_0)$.

Measurable Functions

These are notes from Chapter 5 of [**1**].

4.9. Measurability

Fix a measurable space $(X, \mathcal{A})$

DEFINITION. (5.1.) A function $f : X \to \mathbb{R}$ is **measurable** or $\mathcal{A}$-**measurable** if $f^{-1}((a,\infty)) \in \mathcal{A}$ for all $a \in \mathbb{R}$

PROPOSITION. *(5.5.)The following are equivalent (i)* $f^{-1}((a,\infty)) \in \mathcal{A}$ *for all* $a \in \mathbb{R}$ *ii)* $f^{-1}([a,\infty)) \in \mathcal{A}$ *for all* $a \in \mathbb{R}$ *iii)* $f^{-1}((\infty,a)) \in \mathcal{A}$ *for all* $a \in \mathbb{R}$ *iv)* $f^{-1}((\infty,a]) \in \mathcal{A}$ *for all* $a \in \mathbb{R}$

PROOF. Use compleemnts and intersections/unions with a fudge factor of $\frac{1}{n}$ $\square$

PROPOSITION. *(5.6.) If* X *is a metric space and* $\mathcal{A}$ *contains all the open sets, then continuous functions are always measurable*

PROOF. The preimage through any continuous function of any open set is open, so $f^{-1}((a,\infty)) \in$ Open Sets $\subset \mathcal{A}$ $\square$

PROPOSITION. *(5.7.) If* f, g *are measurable then so are* $f+g, -f, cf, fg, \max(f,g)$ *and* $\min(f+g)$

PROOF. i) $\{f+g<a\} = \cup_{r\in\mathbb{Q}} \{f<r\} \cup \{g<a-r\}$
ii) $\{-f>a\} = \{f<-a\}$
iii) Assume WOLOG by ii that $c>0$ and then:$\{cf>a\} = \{f>a/c\}$
iv) $fg = \frac{1}{2}\left((f+g)^2 - f^2 - g^2\right)$
v) $\{\max(f,g)<a\} = \{f<a\} \cap \{g<a\}$
vi) $\{\min(f,g)>a\} = \{f>a\} \cap \{g>a\}$ $\square$

PROPOSITION. *(5.8.) If* f_n *are all measurable then so are* $\sup f_n, \inf f_n, \limsup f_n, \liminf f_n$

PROOF. $\{\sup f_n > a\} = \cap_n \{f_n > a\}$ and $\{\limsup f_n > a\} = \cup_n \cap_{k>n} \{f_k > a\} = \limsup \{f_n > a\}$. Similar for inf $\square$

PROPOSITION. *(5.10) If* f *is monotone, then* f *is Borel measurable.*

PROOF. Suppose WOLOG that f is increasing. Look at $x_0 = \sup\{y : f(y) \le a\}$ since f is increasing, for $x > x_0$ we have $f(x) \ge a$ by uisng the definiton of inf and the fact that f is increasing. Hence $(x_0,\infty) \subset \{f>a\}$. Similarly, $(-\infty, x_0) \subset \{f \le a\}$. So then $\{f>a\} = [x_0,\infty)$ or (x_0,∞) and in either case it is measurable. $\square$

PROPOSITION. *(5.11.) f is measurable if and only if $f^{-1}(A) \in \mathcal{A}$ for all Borel measurable A*

PROOF. Check that the set $\left\{B : f^{-1}(B) \in \mathcal{A}\right\}$ is a sigma algebra. It is equal to the Borel sigma algebra if and only if it contains the open intervals. □

4.9.1. Non-Borel set of Lebesgue measure 0.

EXAMPLE. (5.12) Let f be the Cantor ternary function and let F be its "inverse"

$$F(x) := \inf \{y : f(y) \geq x\}$$

This is an increasing function, so it is measurable. Morevoer, $F[0,1] = C$ the Cantor set. Since F is Borel measurable, F^{-1} maps Borel sets to Borel sets.

Now, take any non-measurable set A. Let $B = F(A)$. Since $B = F(A) \subset C$, B is a Lebesgue-measure-zero set. On the other hand, B cannot be Borel measureable, for it it were then we would have $A = F^{-1}(B)$ and that would mean that A is measurable, a contradiction.

4.10. Approximation of functions

DEFINITION. (5.13) Define the **characteristic function** χ_E to be the 1 on E and 0 off E. A **simple** function is a a finite linear combination of char functions.

PROPOSITION. *(5.14) Suppose f is a non-negative and measurable function. Then there exists a sequence of non-negative measurable simple functions s_n increasing to f*

PROOF. Perform the "diadic decompostion", $E_{n,i} = \left\{\frac{i-1}{2^n} \leq f < \frac{i}{2^n}\right\}$, and then "round down" to the diadic sets. (You also have to increase the "ceiling" with n as you go) □

4.11. Lusin's Theorem

THEOREM. *(5.15) Suppose $f : [0,1] \to \mathbb{R}$ is Borel measurable and m is the Lebesgue measure. For any $\epsilon > 0$ there exists a continuous function $g : [0,1] \to \mathbb{R}$ which is continuous and which has $\{f = g\}$ is a closed set with $m\{f \neq g\} < \epsilon$.*

PROOF. [Standard ladder proof]

First prove that it is true for characteristic functions. Find a closed set E and an open set G so that $E \subset A \subset G$ with measure $m(G - E) < \epsilon$ and put $\delta = \inf\{|x - y| : x \in E, y \in G^c\}$ and define $g(x) = (1 - d(x,E)/\delta)^+$. Then $g = f$ on $F = E \cap G^c$ (both have the value 1 on E and 0 on G^c) and $m([0,1] - F) < \epsilon$ by the choice of E, G .

For simple functions, $f = \sum_{i=1}^{M} a_i\chi_{A_i}$ approximate each $a_i\chi_{A_i}$ by a continuous function g_i on a closed set F_i so that $m\left([0,1] - F_i\right) < \epsilon/M$ and then put $F = \cap_i F_i$ will be the set we want.

For arbitrary non-negative bounded functions, do the diadic decomposition $f_n = \sum \frac{i}{2^n}\chi_{A_{i,n}}$ which is increasing to f. Notice that $h_n = f_{n+1} - f_n$ is a simple non-negative function which is bounded by 2^{-n} and we have that (by telescoping) $f_n = \sum_{k=1}^{n} h_k$. Now approximate each h_k by a continous function g_k so that they agree on a set F_k so that $m([0,1] - F) \leq \epsilon/2^k$. Then we will have that f_n and $\sum_{k=1}^{n} g_k$ will agree on a set $F = \cap_k F_k$ so that $m([0,1] - F) < \epsilon$. Finally, notice that since each $g_k \leq \frac{1}{2^k}$the infinte sum $\sum_{k=1}^{\infty} g_k$ is uniformly converging (by say M-test) and since each g_k is continuous, it converges to a continous function g. This function is equal to $\sum_{k=1}^{\infty} h_k$ on all of F. But by definition, $\sum_{k=1}^{\infty} h_k = \lim_{n\to\infty} f_n = f$ so this is exactly what we want.

Finally, for arbitarty functions, split up f into f^+ and f^- and do the dance for both pieces. □

EXAMPLE. (5.16) A good exmaple to see what this theorem is and is not saying is to look at $f = \chi_{[0,1]-\mathbb{Q}}$. Since the irrationals are dense in $[0,1]$, the function f takes values of both 1 and 0 in any interval, so you might be tempted to thing that its impossible for a continuous function to be equal to this on a set of positive measure.

The resolution is that, since $m(\mathbb{Q}) = 0$, we can find an closed set F that contains no rationals with $m(F) > 1-\epsilon$ (this can be done explicitly in this case: put $G = \cup_{q\in\mathcal{Q}\cap[0,1]} \left(q - \frac{\epsilon}{2^n}, q + \frac{\epsilon}{2^n}\right)$ and $F = [0,1] - G$ contains no rational numbers.) Put $g \equiv 1$ everywhere on $[0,1]$. Now since F contains no rationals, g and f agree on F (they are identically 1 there) and $m([0,1] - F) < \epsilon$.

So in this case, Lusin's theorem gives us a very useless approximation ot $\chi_{[0,1]-\mathbb{Q}}$, namely the constant function 1. This example nicely illustrates that the type of approximation "there exists a continuous function that agree except for a small measure set" is actually not so useful.

The Lebesgue Integral

These are notes from Chapter 6 of [**1**].

Definition. (6.1.) Let $X, \mathcal{A}, \mu$ be a measure space. If $s = \sum_{i=1}^{n} a_i \chi_{E_i}$ is a simple function, we define the **integral of a simple function** to be:

$$\int s \mathrm{d}\mu = \sum_{i=1}^{n} a_i \mu(E_i)$$

We define the integral of a **non-negative measurable function** to be:

$$\int f \mathrm{d}\mu = \sup \left\{ \int s \mathrm{d}\mu : \ s \leq f \text{ and } s \text{ is simple} \right\}$$

For arbitary functions we define $\int f \mathrm{d}\mu = \int f^{+} \mathrm{d}\mu - \int f^{-} \mathrm{d}\mu$

Definition. (6.2.) We say f is integrable when $\int |f| \, \mathrm{d}\mu < \infty$

Proposition. *(6.3.) Some basic conseqeunces of the definitions*
i) If $a \leq f(x) \leq b$ for all $x \in X$ then $a\mu(X) \leq \int f d\mu \leq b\mu(X)$
ii) If $f \leq g$ then $\int f d\mu \leq \int g d\mu$
iii) If f is integrable, then $\int cf d\mu = c \int f d\mu$ for all c
iv) If $\mu(A) = 0$ and f is integrable then $\int f \chi_A d\mu = 0$

Proof. These are pretty straightforward, most of them are proven by first proving it for simple functions, and then using the definition of the integral as the sup over simple functions to get the result (sometimes the trick you have to do is to do $\sup \left\{ \int s \mathrm{d}\mu : \ s \leq f \text{ and } s \text{ is simple} \right\} = \sup \left\{ \int s \mathrm{d}\mu : \ s \leq f \text{ and } s \text{ is simple and has a certain property} \right\}$ where the certain property helps you get the result. (For example in part i) the property is $s \geq a$) □

Proposition. *(6.4.) If f is integrable, then:*

$$\left| \int f \right| \leq \int |f|$$

Proof. For the real case its easy using $f \leq |f|$ and $-f \leq |f|$ □

Limit Theorems

These are notes from Chapter 7 of [**1**].

6.12. Monotone Convergence Theorem

THEOREM. *(7.1.) Suppose f_n is a sequence of* ***non-negative*** *measurable functions with $f_1 \leq f_2 \leq \ldots$ for all x and with:*

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

Then:

$$\int f_n \, d\mu \to \int f \, d\mu$$

|

PROOF. Since f_n is an increasing sequence of functions, by the result in the last chapter that says integration is monotone (i.e. $f \leq g \implies \int f \leq \int g$), we know that $\int f_n$ is an increasing sequence of numbers. Moreover, it is bounded above by $\int f$ (because $f_n \leq f$ and again because integration is monotone).

By the monotone convergence theorem for real number sthen, there is a limit $\lim_{n\to\infty} \int f_n \mathrm{d}\mu = L$. $L \leq \int f \mathrm{d}\mu$ since each term in the sequence has this property, so we desire only to show that $L \geq \int f \mathrm{d}\mu$. By the definition of the integral, it is sufficient to show that $L \geq \int s \mathrm{d}\mu$ for any simple function s with $s \leq f$.

To do this we give ourselves a multiplicative factor of $1 - \epsilon$ of room. Fix $\epsilon > 0$. For any simple function s with $f \leq s$ look at the set $A_n = \{f_n > (1-\epsilon)s\}$. Since $f_n \uparrow f$ and $s < f$ we know that $A_n \uparrow X$. Now notice that (from the basic properties of integration we proved) that:

$$\begin{aligned}
\int f_n &\geq \int f_n \chi_{An} \\
&\geq (1-\epsilon) \int s \chi_{A_n} \\
&= (1-\epsilon) \sum a_i \mu \left(E_i \cap A_n\right) \\
&\to (1-\epsilon) \sum a_i \mu(E_i) \\
&= (1-\epsilon) \int s \mathrm{d}\mu
\end{aligned}$$

Since this holds for any $\epsilon > 0$ we have $\int f \geq \int s$ and the result then follows by the defintion of the integral! □

6.13. Linearity of the Integral

THEOREM. *(7.4.) If f, g are non-negative and measurable, then:*

$$\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu$$

PROOF. (Ladder Proof)

We first prove this for non-negative simple functions, which is easy from the defintion.

Then we use the diadic decomposition to get $s_n \uparrow f$ and $r_n \uparrow g$ and then use the monotone convergence theorem to get the result for non-negative f,g.

For arbitarty f, g, split it into f^+ and f^- □

6.14. Fatou's Lemma

THEOREM. *(7.6.) Suppose that f_n are* ***non-negative*** *and measurable. Then:*

$$\int \liminf f_n \leq \liminf \int f_n$$

PROOF. Let $g_n = \inf_{k\geq n} f_k$ so that $g_n \uparrow \liminf f_n$. Clearly, $g_n \leq f_k$ for each $k \geq n$ so we have:

$$\int g_n \leq \inf_{k\geq n} \int f_k$$

Now take the limit of this inequality as $n \to \infty$. By the Monotone covnergence theorem, the LHS goes to $\int \liminf f_n$ (since $g_n \uparrow \liminf f_n$) while the RHS goes to $\liminf \int f_n$ just by the definition of the it □

REMARK. Let's compare this to Fatou's lemma for sets:

$$\mathbf{P}\left(\liminf A_n\right) \leq \liminf \mathbf{P}\left(A_n\right)$$

PROOF. Let $B_n = \cap_{k\geq n} A_k$ so that $B_n \uparrow \liminf A_n$ by the definition of the liminf of sets. Also clearly, $B_n \subset A_k$ for any $k \geq n$ so in particular:

$$\mathbf{P}\left(B_n\right) \leq \inf_{k\geq n} \mathbf{P}(A_k)$$

Take limit as $n \to \infty$now. By the continuity of probability, $B_k \uparrow \liminf A_n \implies \mathbf{P}(B_k) \uparrow \mathbf{P}\left(\liminf A_n\right)$ and on the other side tends to $\liminf \mathbf{P}(A_n)$ □

REMARK. So the proof is exactly the same, with the continuity of probability doing the work in the set case, and the monotone convergence theorem doing the work in the function case. (notice actually that $A_n \uparrow X \implies \mu(E_i \cap A_n) \to \mu(E_i)$ is also used explicitly in the MCT proof)

Once the set one is proven btw, I think you can do a standard ladder to get to the Monotone convergence theorem for arbitary functions.

6.15. Dominated Convergence Theorem

THEOREM. *(7.7.) Suppose f_n are measurable real-valued functions and $f_n \to f$ a.e.. Suppose there exists an integrable function g such that $|f_n| \leq g$ for all x. Then:*

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

PROOF. Use Fatous's lemma once on $f_n + g$ which is non-negative and once on $g - f_n$ which is also non-negative. You get a limsup/liminf sandwhich and get the desired equality. □

THEOREM. *(Exercise 7.2.) If f_n, g_n, f, g are integrable with $f_n \to f$ a.e., $g_n \to g$ a.e. and $|f_n| \leq g_n$ for each n and $\int g_n \to \int g$, then $\int f_n \to \int f$*

PROOF. Since $f_n - g_n$ and $g_n - f_n$ are non-negative we can apply Fatou to know that:

$$\begin{aligned} \liminf \int (f_n - g_n) &\geq \int (\liminf f_n - g_n) \\ &= \int f - g \\ &= \int f - \int g \end{aligned}$$

On the other hand, $\liminf \left(\int f_n - g_n \right) = \liminf (\int f_n) - \int g$ by the fact that $\int g_n \to \int g$

(Use the lemma: If $b_n \to b$ then $\liminf (a_n + b_n) \to \liminf(a_n) + b$ Pf: (Subsequences) Take a subsequence n_k so that $a_{n_k}$is convergent and . Then $b_{n_k} \to b$ since b_n is convergent. Have then $\lim (a_{n_k} + b_{n_k}) = \lim(a_{n_k}) + b \geq \liminf(a_n) + b$. Since every limit point of $a_n + b_n$ is $\geq \liminf(a_n) + b$ the liminf obeys this inequailty too. By taking a subsequence so that $\lim(a_{n_k}) = \liminf(a_n)$, we see the other inequality.)

From here the proof procedes the same as the ordinary Fatou lemma. □

Product Measures

These are notes from Chapter 5 of [**1**].

7.16. Product $\sigma-$algebras

Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measure σ-finte measure spacdes.

DEFINITION. A **measurable rectangle** is a set of the form $A \times B$ with $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Let $\mathcal{C}_0$ be the collection of finite unions of disjoint measurable rectangles, i.e. $\mathcal{C}_0 = \{\cup_{i=1}^n A_i \times B_i\}$. One can easily verify that $\mathcal{C}_0$ is an algebra. (Some people say the measurable rectangles are a semi algebra and then have a theory based on that, but its exactly the same as this)

We define the **product sigma algebra** by:

$$\mathcal{A} \times \mathcal{B} = \sigma(\mathcal{C}_0)$$

For $E \in \mathcal{A} \times \mathcal{B}$ we define its x-section and y-section by:

$$\begin{aligned} s_x(E) &:= \{y \in Y : (x, y) \in E\} \\ t_y(E) &:= \{x \in X : (x, y) \in E\} \end{aligned}$$

Given a function $f : X \times Y \to \mathbb{R}$ that is $\mathcal{A} \times \mathcal{B}$ measurable, we define for each x and y $S_x f : Y \to \mathbb{R}$ and $T_y f : X \to \mathbb{R}$ by:

$$\begin{aligned} S_x f(y) &:= f(x, y) \\ T_y f(x) &:= f(x, y) \end{aligned}$$

LEMMA. *(11.1) i) If $E \in \mathcal{A} \times \mathcal{B}$ then $s_x(E) \in \mathcal{B}$ for each x and $t_y(E) \in \mathcal{A}$ for each y*

ii) If f is $\mathcal{A} \times \mathcal{B}$ measurable, then $S_x f$ and $T_y f$ are $\mathcal{B}$ and $\mathcal{A}$ measurable respectively.

PROOF. (Sigma Albgebra proof)

Let $\mathcal{C}$ be the collection of sets for which $s_x(E) \in \mathcal{B}$. Check that its a sigma algebra and that it contains the rectangles.

For ii), do a ladder proof. The result holds for simple functions by i) and so on. □

PROPOSITION. *(11.2) For $E \in \mathcal{A} \times \mathcal{B}$, define $h(x) = \nu(s_x(E))$ and $k(y) = \mu(t_y(E))$ then:*

i) h, k are measurable

ii) We have:

$$\int h(x)\mu(dx) = \int k(y)\nu(dy)$$

In other words:

$$\int\left[\int S_x\chi_E(y)\nu(dy)\right]\mu(dx)=\int\left[\int T_y\chi_E(x)\mu(dx)\right]\nu(dy)$$

This is sometimes written more succinctly as:

$$\int\int\chi_E(x,y)\nu(dy)\mu(dx)=\int\int\chi_E(x,y)\mu(dx)\nu(dy)$$

PROOF. (This is where the monotone class theorem is finally used)

Let $\mathcal{C}$ be the collection of sets where this holds. We check that it contains the rectangles in $\mathcal{A}\times\mathcal{B}$ and also similarly for a finite union of rectangles (can assume WOLOG they are disjoint)

Now suppose $E_n\uparrow E$ and each $E_n\in\mathcal{C}$. The result will follow by the monotone convergence theorem.

If $E_n\downarrow E$ and each $E_n\in\mathcal{C}$, then the result will follow by using the dominated convergence theorem. (To make this one work you must assume that μ and ν are finite measures....for $\sigma-$finite it will still work by proving it for each finite piece first)

Hence $\mathcal{C}$ is a monotone class containing the algebra of finite unions of rectangles, therefore it actually contains the whole sigma algebra generated by these, which is all of $\mathcal{A}\times\mathcal{B}$. □

DEFINITION. The **product measure** $\mu\times\nu$ is defined as:

$$\mu\times\nu(E)=\int h(x)\mu(\mathrm{d}x)=\int k(y)\nu(\mathrm{d}y)$$

This is indeed a measure. One can check it is finitely additive, and then use a monotone convergence theroem to get that it is countable additive.

7.17. The Fubini Theorem

THEOREM. *(11.3.) If $f:X\times Y\to\mathbb{R}$ is measurable with respect to $\mathcal{A}\times\mathcal{B}$. If either:*

a) f is non-negative

b) $\int|f(x,y)|\,d(\mu\times\nu)(x,y)<\infty$

Then:

$$\begin{aligned}\int\int f(x,y)d(\mu\times\nu) &= \int\left(\int f(x,y)d\mu(x)\right)d\nu(y)\\ &= \int\left(\int f(x,y)d\nu(y)\right)d\mu(y)\end{aligned}$$

REMARK. Even if $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ are complete, the product $\mu\times\nu$ on $\mathcal{A}\times\mathcal{B}$ will not be complete. (Example a singleton cross with a non-measurable set is not measurable) For this reason

Signed Measures

These are notes from Chapter 12 of [**1**].

8.18. Positive and Negative sets

DEFINITION. (12.1) Let $\mathcal{A}$ be a sigma algebra. A signed measure is a function $\mu : \mathcal{A} \to (-\infty, \infty]$ such that $\mu(\emptyset) = 0$ and $\mu\left(\dot{\cup}_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$ whenever the A_i's are pairwise disjoint.

REMARK. The only difference between this and an ordinary measure is that an ordinary measure is required to be non-negative as well.

DEFINITION. (12.2.) Let μ be a signed measure, a set $A \in \mathcal{A}$ is called a **positive set** for μ if $\mu(B) \geq 0$ for all $B \subset A$ and $B \in \mathcal{A}$. Similarly a negative set A is one where $\mu(B) \leq 0$ for all $B \subset A$, $B \in \mathcal{A}$.

EXAMPLE. (12.3) If m is the Lebesgue measure then :

$$\mu(A) = \int_A f \mathrm{d}m$$

for some integrable f is a signed measure. The sets $P = \{f > 0\}$ and $N = \{f < 0\}$ are positive and negative sets for μ.

PROPOSITION. *(12.4.) Let μ be a signed measure which takes values in $(-\infty, \infty]$ let E be measureable with $\mu(E) < 0$. Then there exists a measurable subset F of E which cis measurable and which is a negative set with $\mu(F) < 0$.*

PROOF. (The idea is to start with E and cut out all the positive measure sets of size $\frac{1}{n_k}$. Any subset of the remaining will be negative or else it would have been cut out)

I'm going to skip the actual details here. □

8.19. Hahn Decomposition Theorem

DEFINITION. $A \triangle B := (A - B) \cup (B - A)$.

THEOREM. *(12.5.) (1) Let μ be a signed measure taking values in $(-\infty, \infty]$. There exists disjoint measurable sets E and F in $\mathcal{A}$ whose union is X and such that E is a negative set and F is a positive set.*

2) The decomposition is essentially unique in the sense that: If E' and F' are another pair of positive and negative sets whose union is X, then $E \triangle E' = F \triangle F'$ is a null set w.r.t. μ

3) If μ is not a positive measure then $\mu(E) \neq 0$. If $-\mu$ is not a positive measure then $\mu(F) \neq 0$.

PROOF. (1) Let $L = \inf \{\mu(A) : A \text{ is a negative set } \}$ Choose sets A_n so that $\mu(A_n) \to L$. Let $E = \cup_{n=1}^{\infty} A_n$. Let $B_n = A_n - (B_1 \cup \ldots \cup B_n)$ so that $\cup_{k=1}^{n} B_k = \cup_{k=1}^{n} A_k$ and the B_k's are pairwise disjoint. Since the A_n's are all negative sets, so are the B_n's. We claim that $E = \cup_n B_n = \cup_n A_n$ is a negative set. This follows by the continuity of measure, sicne for any $C \subset E$ we have::

$$\mu(C) = \lim_{n\to\infty} \mu\left(C \cap (\cup_{i=1}^{n} B_i)\right) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(C \cap B_i) \leq 0$$

We claim that $\mu(E) = L$ since $L \leq \mu(E) = \mu(A_n) + \mu(E - A_n) \leq \mu(A_n) \to L$ makes the desired sammich.

Now we claim that $F = E^c$ is a positive set. Otherwise, there is a subset $C \subset F$ of negative measure. But then C contains a negative set (prop 12.4.) and so $E \cup C$ makes a negative set with measure stricly less than L, contradicting L as the size of the "largest" negative set.

(2) Rewritting $E \triangle E'$ as $(E \cap F') \cup (E' \cap F)$ makes it clear that $E \triangle E' = F \triangle F'$. Any subset of this can be written as a subset of $E \cap F'$ union with a subset of $F \cap E'$. Any subset of either of these must simultaneuous have non-negative and non-positive measure, so it is indeed a null set.

(3) Skip this for now. □

DEFINITION. We say two (non-negative) measures μ and ν are **mutually singular** if there exist two disjoint subsets E and F in $\mathcal{A}$ whose union is all of X and with $\mu(E) = 0$ and $\mu(F) = 0$. This is often written $\mu \perp \nu$.

EXAMPLE. (12.6) The Lebesgue measure restricted to $(0, \frac{1}{2}]$ and the Lebesgue measure restrived to $[\frac{1}{2}, 1)$ are mutually singular. This works because the Lebsegue measure of $\{\frac{1}{2}\}$ is 0.

EXAMPLE. (12.7.) Look at the Cantor Ternary function f on $[0, 1]$ and let ν be the assoiciated Lebesgue-Stieltjes measure. Let μ be the Lebesgue measure. Then $\mu \perp \nu$ by the partitioning of the $[0, 1] = C \cup C^c$ where C is the Cantor set. On C^c, ν is measure zero because f is constant on every interval there.

8.20. Jordan Decomposition Theorem

THEOREM. *(12.8.) If μ is a signed measure on a measurable space $(X, \mathcal{A})$, then there exist positive measure μ^+ and μ^- such that $\mu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$. Moreover this decomposition is unique.*

PROOF. Use the Hanh decomposition to find the negative set E and positive set F that decompose X/ Let $\mu^+(A) = \mu(A \cap F)$ and $\mu^-(A) = -\mu(A \cap E)$. This gives the desired decomposition.

The uniqueness follows with some work from the "uniqueness up to null sets" result for the Hanh decomposition. □

DEFINITION. The measure:

$$|\mu| = \mu^+ + \mu^-$$

Is called the **total variation measure** of μ and $|\mu|(X)$ is called the total variation of μ.

The Radon-Nikodym Theorem

These are notes from Chapter 13 of [**1**].

9.21. Absolute Continuity

Definition. (13.1) A measure ν is said to be **absolutly continuous** with respect to a measure μ if $\mu(A)=0 \implies \nu(A)=0$. This is written $\nu \ll \mu$

Proposition. *(13.2) Let ν be a finite measure. Then ν is absolutely continuous with respect to μif and only if for all $\epsilon>0$, there exists $\delta>0$ so that $\mu(A)<\delta \implies \nu(A)<\epsilon$*

Proof. ($\Rightarrow$) Suppose by contradiction this is not the case. Then there exists ϵ_0 so that for every δ there is a bad set A_δ with $\mu(A_\delta)<\delta$ and $\nu(A_\delta)>\epsilon_0$. Take $\delta=2^{-n}$ to get a sequence of these sets and let $A=\limsup A_n$. Then since $\sum \mu(A_n)<\infty$ we know that $\mu(A)=0$ by Borel-Ceantelli but on the other hand $\nu(A)=\nu\left(\limsup A_n\right) \geq \limsup \nu(A_n)>\epsilon_0$ which contradicts absolute continuity.

($\Leftarrow$) If $\mu(A)=0$ then $\mu(A)<\delta$ for all δ and so using the hypothesis we can conclude that $\nu(A)<\epsilon$ for any ϵ, meaning that $\nu(A)=0$ too. $\square$

Remark. This proposition gives the connection with the notion of absolute continuity for a function. Recall that a function is said to be absolutely continuous if for all $\epsilon>0$, there exists $\delta>0$ so that any pairwise disjoint sequence of intervals (x_k,y_k) has $\sum_k |y_k-x_k|<\delta \implies \sum_k |f(y_k)-f(x_k)|<\epsilon$.

9.22. The main theorem

Lemma. *(13.3.) Let μ and ν be finite positive measure on a measurable space $(X,\mathcal{A})$. Either $\mu \perp \nu$ or else there exists $\epsilon>0$ and $G\in\mathcal{A}$ such that $\mu(G)>0$ and G is a positive set for $\nu-\epsilon\mu$. (i.e. $\nu(A)\leq\epsilon\mu(A)$ for all $A\subset G$)*

Proof. Conisder the Hahn decomposition for $\nu-\frac{1}{n}\mu$. Let $F=\cup_n F_n$ be the union over all the positive sets and $E=\cap_n E_n$ be the intersection of all the negative sets. Notice that $E^c=F$ by DeMorgan's laws since $E_n^c=F_n$ for each n.

Claim: $\nu(E)=0$

Pf: $\nu(E)\leq\nu(E_n)\leq\frac{1}{n}\mu(E_n)\leq\frac{1}{n}\mu(X)\to 0$

If $\mu(E^c)=\mu(F)=0$ then the sets E and F divide up X to show that $\mu\perp\nu$.\

Otherwise, $\mu(E^c)=\mu(F)>0$. In this case, $\mu(F_n)>0$ for some F_n. Choosing $G=F_n$ and $\epsilon=\frac{1}{n}$ gives the result. $\square$

Theorem. *(13.4) Suppose that $\mu$is a $\sigma-$finite positve measure on a measurable space $(X,\mathcal{A})$ and ν is a finite positive measure on $(X,\mathcal{A})$ such that $\nu\ll\mu$. Then* *there exists a $\mu-$integrable non-negative f which is measurable with respect to $\mathcal{A}$ such that:*

$$\nu(A) = \int_A f\,d\mu$$

Moreover the f is unique up to$\mu-$a.e.

Proof. (The idea is to look at the set of functions f so that $\int_A f\mathrm{d}\mu \leq \nu(A)$ for all $A \in \mathcal{A}$, and then to take the "largest" of these) The uniqueness part is easy since if $\int_A f\mathrm{d}\mu = \int_A g\mathrm{d}\mu$ for all $A \in \mathcal{A}$ then $\int_A (f-g)\mathrm{d}\mu = 0$ for all $A \in \mathcal{A}$ and then $f = g$ a.e. (for example look at the set $A = \{|f-g| > \epsilon\}$ must be measure 0 for all ϵ)

Let $\mathcal{F} = \left\{g \text{ measurable} : g \geq 0, \int_A g\mathrm{d}\mu \leq \nu(A) \text{ for all } A \in \mathcal{A}\right\}$

Let $L = \sup\left\{\int g\mathrm{d}\mu : g \in \mathcal{F}\right\}$ and take a sequence $g_n \in \mathcal{F}$ so that $\int g_n \to L$. Let $h_n = \max(g_1, \ldots, g_n)$. Check that this is still in $\mathcal{F}$. Let $f = \sup h_n = \lim h_n$ since the h_n's are increasing. By the monotone convergence theorem, $\int f\mathrm{d}\mu = \lim_n \int h_n\mathrm{d}\mu \leq \nu(A)$ since $h_n \in \mathcal{F}$ which shows that $f \in \mathcal{F}$ and also $\int f\mathrm{d}\mu \geq \int h_n\mathrm{d}\mu \geq \int g_n\mathrm{d}\mu \to L$shows that $\int f\mathrm{d}\mu = L$.

Now there is some technical work (in partticular using the previous lemmas) to show that f is the desired function and also the fact that $\nu \ll \mu$ to make it work. □

9.23. Lebesgue Decomposition Theorem

Theorem. *(13.5.) If μ is a sigma finite measure and ν is a finite measure, then there exists positive measure $\lambda$and ρ so that:*

i) $\nu = \lambda + \rho$

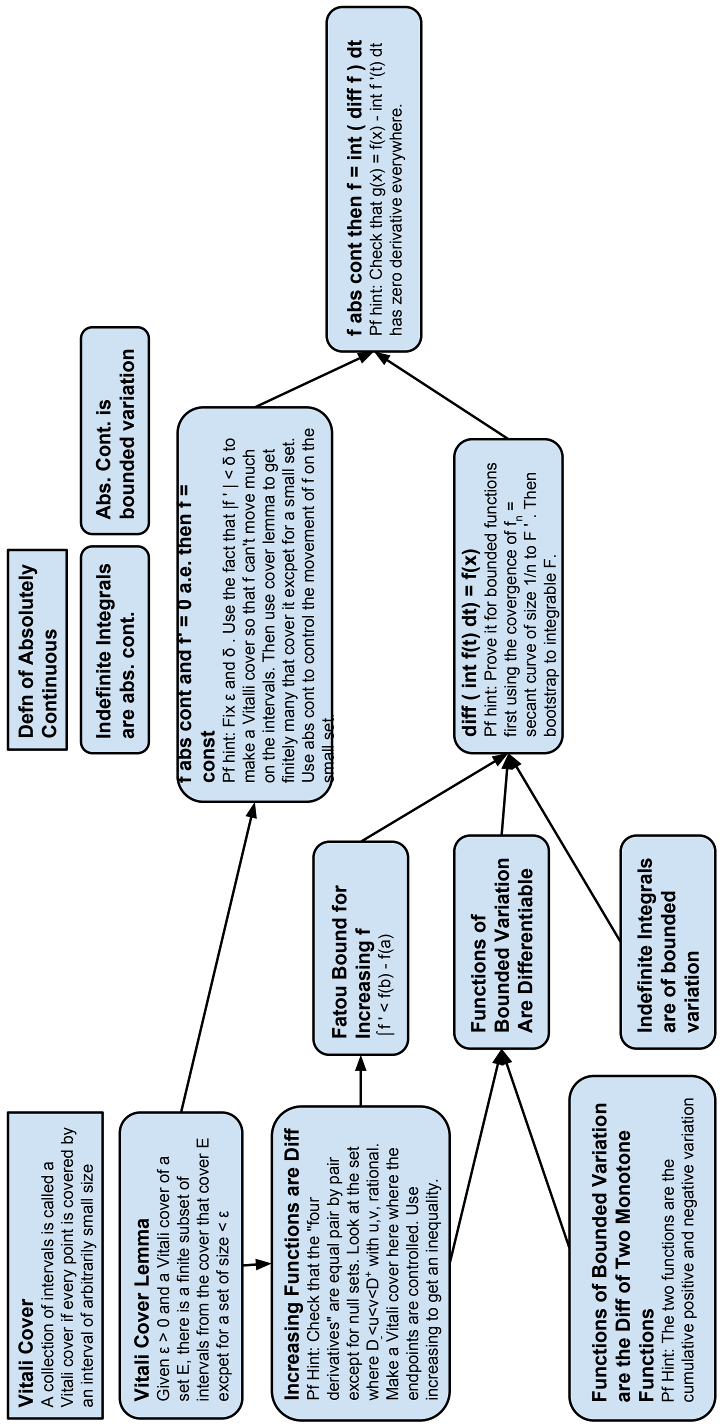
ii) $\rho \ll \mu$

iii) $\lambda \perp \mu$

Proof. Take f as in the proof of the R-N theorem and set $\rho(A) = \int_A f\mathrm{d}\mu$ and $\lambda = \nu - \rho$ and it will work out (technical details are omitted) □

Differentiation

These are notes from Chapter 5 of [**3**]. Here is a mind map of the proofs of this section:


Vitali Cover
A collection of intervals is called a Vitali cover if every point is covered by an interval of arbitrarily small size
Vitali Cover Lemma
Given ε > 0 and a Vitali cover of a set E, there is a finite subset of intervals from the cover that cover E excpet for a set of size < ε
Increasing Functions are Diff
Pf Hint: Check that the "four derivatives" are equal pair by pair except for null sets. Look at the set where D-<u<v<D+ with u,v, rational. Make a Vitali cover here where the endpoints are controlled. Use increasing to get an inequality.
Functions of Bounded Variation are the Diff of Two Monotone Functions
Pf Hint: The two functions are the cumulative positive and negative variation
Fatou Bound for Increasing f
∫f ' < f(b) - f(a)
Functions of Bounded Variation Are Differentiable
Indefinite Integrals are of bounded variation
Defn of Absolutely Continuous
Indefinite Integrals are abs. cont.
Abs. Cont. is bounded variation
f abs cont and f' = 0 a.e. then f = const
Pf hint: Fix ε and δ . Use the fact that |f ' | < δ to make a Vitalli cover so that f can't move much on the intervals. Then use cover lemma to get finitely many that cover it excpet for a small set. Use abs cont to control the movement of f on the small set.
diff (int f(t) dt) = f(x)
Pf hint: Prove it for bounded functions first using the covergence of fn = secant curve of size 1/n to F '. Then bootstrap to integrable F.
f abs cont then f = int (diff f) dt
Pf hint: Check that g(x) = f(x) - int f '(t) dt has zero derivative everywhere.

10.24. Differentiation of Monotone Functions

DEFINITION. Let $\mathcal{I}$ be a collection of intervals we say that $\mathcal{I}$ is a **Vitali covering** of a set E if for all $\epsilon > 0$ and any $x \in E$, $\exists I \in \mathcal{I}$ with $x \in I$ and $\ell(I) < \epsilon$. The intervals may be open, closed or half-opne but no degenerate intervals (i.e. singletons) are allowed.

REMARK. Another way to say this heurstically: "$\mathcal{I}$ is a Vitali covering if it covers every point in E with an arbitarily small interval"

LEMMA. *(Vitali) Let E be a set of finte outer measure and $\mathcal{I}$ a Vitali covering of E. Then given $\epsilon > 0$ there is a finite disjoint collection $\{I_1, \ldots, I_N\}$ of intervals in $\mathcal{I}$ such that:*

$$m^* \left[E - \bigcup_{n=1}^{N} I_n \right] < \epsilon$$

PROOF. (Pf Idea only)

Assume WOLOG all the intervals are closed and $O \supset E$ is an open set containg E and every interval is a subset of O.

The idea is to recusrvly construct a sequence $\{I_1, \ldots\}$ by adding in intervals one at a time, chosen so that they are disjoint and every new added inteval is close to being as large (in the sense of length) as it can be. (i.e. its length is at least $\frac{1}{2}$ the sup over the lengths of all candidate intervals) Have then $\sum_{k=1}^{\infty} \ell(I_k) \leq m(O)$. For any $\epsilon > 0$ take N so large so that $\sum_{k \geq N} \ell(I_k) < \epsilon$ then and then show that $\cup_{n=1}^{N} I_n$ is the set we are looking for. The fact that its a Vitalli covering, and that we chose the intervals to be close to as large as possible means we could not have missed a piece of E of size more than ϵ. □

DEFINITION. The **derivatives** (there are four of them) of a function f are defined as: (Can come from the left or right (denoted by+ or −) and can be the limsup or the liminf (denoted by either a superscript or a subscript)

$$\begin{aligned} D^+ f(x) &= \limsup_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \\ D^- f(x) &= \limsup_{h \to 0^-} \frac{f(x+h) - f(x)}{h} \\ D_+ f(x) &= \liminf_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \\ D_- f(x) &= \liminf_{h \to 0^-} \frac{f(x+h) - f(x)}{h} \end{aligned}$$

Clearly $D^+ \geq D_+$ and $D^- \geq D_-$. If all four of these are equal, then we say the function is differentiable and call its derivative f'. If $D^+ f = D_+ f$ we say that f has "right-hand" derivatevs and define $f'(x+)$ to be its value. Similary for $f'(x-)$

REMARK. The connection between the Vitali covering lemma and these derivatives is that at each point x, we can find (by definiton of limsup/liminf) we can find a points $h_n \to 0$ so that $f(x + h_n)$ is controlled. The intervals $[x, x + h_n]$ will be a nice vitalli covering for us.

PROPOSITION. *(2) If f is continuous on $[a, b]$ and one of its derivatives (for concreteness say D^+) is everywhere non-negative on (a, b), then f is non-decreasing on $[a, b]$ (i.e. $f(x) \leq f(y)$ for all $x \leq y$ in the interval)*

PROOF. ....I tried for a bit to come up with a proof by contradiction but coudlnt make it work. □

THEOREM. *Let f be an increasing real-valued function on the interval $[a,b]$. Then f is differentiable almost everywhere. The derivative f' is measurable and:*

$$\int_a^b f'(x)dx \leq f(b) - f(a)$$

REMARK. The Cantor ternary function is an example where $f' = 0$ a.e. and so $\int f' = 0 < f(b) - f(a)$

PROOF. The idea is to show that any two of the the four derivateves agree except on a set of measure 0. For concreteness we will handle $D^+f = D_-f$. To show that thesea re not equal except on a set of measure zero, it suffices to show by countability that $E_{uv} = \{D^+f > u > v > D_-f\}$ is a measure zero set for all u, v rational.

Let O be an open set so that $m(O) < m(E_{u,v}) + \epsilon$. For each $x \in E_{u,v}$ there is an arbitrarily small intervale $[x-h, x]$ contained in O such that (since $D_- < v$):

$$f(x) - f(x-h) < vh$$

This is a Vitalli covering! By the Vitali covering lemma, we choose a finite collection $\{I_1, \ldots, I_N\}$ of them whose interiors cover $E_{u,v}$ with measure at least $m(E_{u,v}) - \epsilon$. Denote $A = \cup I_k$. Then summing over these intervals, we have:

$$\begin{aligned} \sum_{n=1}^{N} [f(x_n) - f(x_n - h)] &\leq v \sum_{n=1}^{N} h_n \\ &\leq v\,(m(O) + \epsilon) \\ &\leq v\,(m(E_{u,v}) + \epsilon) \end{aligned}$$

Now, each point $y \in A$ is the left endpoint of some interval $(y, y+k)$so that $f(y+k) - f(y) > uk$. Supopose WOLOG that each of these is a subset of some I_k. Using this as a Vitalli covering lemma for THIS new vitalli covering $J_1, \ldots$ with each interval being a subset of some $I_i's$. Have (similar to before):

$$\begin{aligned} \sum_{i=1}^{M} f(y_i + k_i) - f(y_i) &> u \sum k_i \\ &> u(m(E_{u,v}) + \epsilon) \end{aligned}$$

Now, SINCE f is increasing, and every interval J_i is contained in some interval I_i we have for each $J_i$that:

$$\sum_{J_i \subset I_n} f(y_i + k_i) - f(y_i) \leq f(x_n) - f(x_n - h_n)$$

so summing these up we get:

$$\sum_{n=1}^{N} [f(x_n) - f(x_n - h)] \geq \sum_{i=1}^{M} f(y_i + k_i) - f(y_i)$$

But the LHS $\leq v\,(m(E_{u,v}) + \epsilon)$ and the RHS $> u(m(E_{u,v}) + \epsilon)$. Since this holds for every $\epsilon > 0$ we conclude that $vm(E_{u,v}) \geq um(E_{u,v})$. But since $u < v$ this can only hold if $m(E_{u,v}) = 0$, as desired.

This shows that:

$$g(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Is defined almost everywhere and that f is differentiable when g is finite.

To get the desired inequality on the integral of the derivative, we use Fatous lemma and the sequence $g_n(x) = \frac{f(x+h)-f(x)}{h}\Big|_{h=\frac{1}{n}}$ (Notice that g is non-negative since f is increasing, so we can indeed use Fatou's lemma!) which converges pointwise to g. Have:

$$\begin{aligned} \int_a^b g(x) &= \int_a^b \liminf g_n \\ &\leq \liminf \int_a^b g_n \\ &\leq \ldots \\ &= f(b) - f(a) \end{aligned}$$

□

10.25. Functions of Bounded Variation

Let f be a real-valued function defined on $[a, b]$ and define, for any subdivision $x_0 < x_1 < \ldots < x_n$:

$$\begin{aligned} p &= \sum_{i=1}^{k} [f(x_i) - f(x_{i-1})]^+ \\ n &= \sum_{i=1}^{k} [f(x_i) - f(x_{i-1})]^- \\ t &= n + p \\ &= \sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| \end{aligned}$$

Here $x^+ = |x| \, 1_{x>0}$ and $x^- = |x| \, 1_{x<0}$. Clearly, $f(b) - f(a) = p - n$. Set:

$$\begin{aligned} P &= \sup p \\ N &= \sup n \\ T &= \sup t \end{aligned}$$

These are called the positive, negative and total variations of f. We sometimes write $P_{a,b}$ to remind ourselves of the interval we are looking over.

DEFINITION. We say f is **bounded varaition** if $T < \infty$.

THEOREM. *If f is of bounded variation, then:*

$$T_{a,b} = P_{a,b} + N_{a,b}$$

and:

$$f(b) - f(a) = P_{a,b} - N_{a,b}$$

COROLLARY. *If f is of bounded variation, then we can write it as difference of two monotone functions*

PROOF. By the theorem, we can write:

$$f(x) = f(a) + P_{a,x} - N_{a,x}$$

$P_{a,x}$ and $N_{a,x}$ are seen to be monotone functions of x. □

REMARK. The converse is also true since any monotone function is of bounded varaition, and sums of bounded variation functions are bounded variation.

PROOF. For any subdivision we have:

$$p = n + f(b) - f(a)$$

and taking suprema over all possible subdivisions we get:

$$P = N + f(b) - f(a)$$

Now we can write:

$$t = p + n = 2p - (f(b) - f(a))$$

And taking superma gives:

$$T = 2P - (f(b) - f(a)) = P + N$$

□

COROLLARY. *Every function of bounded variation is differentiable.*

10.26. Differentiation of an Integral

We will show that the derivative of $F(x) = \int_a^x f(t)\mathrm{d}t$ is differentiable whenever f is integrable.

LEMMA. *If f is integrable on $[a, b]$ then the function F defined by:*

$$F(x) = \int_a^x f(t)\,dt$$

is a continuous function of bounded variation on $[a, b]$

PROOF. The fact that its continuous follows from the "uniform integrability" property of integrals; for any $\epsilon > 0$ there is a $\delta > 0$ so that the integral over a set of size $< \delta$ is less than ϵ.

To show that f is bounded variation, notice that:

$$\begin{aligned} \sum |F(x_i) - F(x_{i-1})| &= \sum \left| \int_{x_i}^{x_{i-1}} f(t)\mathrm{d}t \right| \\ &\leq \sum \int_{x_i}^{x_{i+1}} |f(t)|\,\mathrm{d}t \\ &= \int |f(t)|\,\mathrm{d}t \end{aligned}$$

So taking sup, we have an upper bounder for the total varaition. □

LEMMA. *If f is integrable on $[a, b]$ and:*

$$\int_a^x f(t)\,dt = 0$$

for all $x \in [a, b]$ then $f(t) = 0$ a.e. in $[a, b]$

PROOF. I think I have a creative proof using the monotone class theorem. Let:

$$\mathcal{M}=\left\{A:\int f\mathtt{1}_A=0\right\}$$

We see (with a bit of work) from the hypothesis that $\mathcal{M}$ contains the algebra of finite collections of intervals.

Now $\mathcal{M}$ is a monotone class, since if $A_n\uparrow A$ or $\downarrow A$ then $f\mathtt{1}_{A_n}\to f\mathtt{1}_A$ a.e. and is dominated by the integrable function $|f|$, so by LDCT $0=\int f\mathtt{1}_{A_n}\to\int f\mathtt{1}_A$ shows $A\in\mathcal{M}$. (I guess I could have done this with σ-algebra instead too)

Hence $\mathcal{M}$ is the Borel sets. This means f is 0, becuase for example the set $\left\{f>\frac{1}{n}\right\}$ must be measure 0 for each n. □

LEMMA. *(9) If f is <u>bounded</u>, $|f|\leq K$ and measureable on $[a,b]$ and:*

$$F(x)=\int_a^x f(t)\,dt+F(a)$$

Then $F'(x)=f(x)$ for almost all $x\in[a,b]$

PROOF. Since F is of bounded variation, we know it is differentiable. Set $f_n=\left.\frac{F(x+h)-F(x)}{h}\right|_{h=\frac{1}{n}}$ so that $f_n\to F'(x)$ a.e.. Notice also that $f_n=\frac{1}{h}\int_x^{x+h}f(t)\mathrm{d}t\leq K$ is bounded, so we are ok to use a bounded convergence theorem. We now check that:

$$\begin{aligned}\int F'\mathrm{d}x &= \lim_{n\to\infty}\int f_n\\ &= \lim_{n\to\infty}n\int F(x+\frac{1}{n})-F(x)\mathrm{d}x\\ &= F(b)-F(a)\\ &= \int_a^b f(x)\mathrm{d}x\end{aligned}$$

So then $F'-f$ integrates to zero on every interval, and consequently must vanish everywhere. □

THEOREM. *The above works if f is integrable instead of boudned:*

$$\frac{d}{dx}\int_a^x f(t)\,dt=f(x)$$

PROOF. Take $f_n\uparrow f$ with each $f_n=f\wedge n$ bounded by say n. We know that $\frac{d}{dx}\int_a^x f_n=f_n(x)$ a.e. since these are bounded. and so we have that:

$$\begin{aligned}F^{'}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\int f\\ &= \frac{\mathrm{d}}{\mathrm{d}x}\int f_n+\frac{\mathrm{d}}{\mathrm{d}x}\int(f-f_n)\\ &\geq f_n(x)\end{aligned}$$

Since $f-f_n\geq 0$. But then $\int\left(F^{'}-f_n\right)\geq 0$ for any n and consequently $\int F'-f\geq 0$ too (LDCT justifies us) For any interval. On the other hand, we showed (it was a Fatou's lemma in the end) that $\int F_n'\leq F(b)-F(a)=\int_a^b f$. So we have cobining these inequalites that $\int(F'-f)=0$ over any interval and so $F'=f$ a.e. □

10.27. Absolute Continuity

DEFINITION. A real valued function f is said to be absolutly continuous on $[a, b]$ if and only if, for all $\epsilon > 0$ there is a $\delta > 0$ such that:

$$\sum_{i=1}^{n} |x_i' - x_i| < \delta \implies \sum_{i=1}^{n} |f(x_i') - f(x_i)| < \epsilon$$

REMARK. Compare this with the "definition" of absolute continuity for a measure, $\nu \ll \mu$ if for all $\epsilon > 0$ there exists $\delta > 0$ so that:

$$\nu(A) < \delta \implies \mu(A) < \epsilon$$

(The more usual way to think about this is $\nu(A) = 0 \implies \mu(A) = 0$, which is equivalent.) This is how you could think of absolutly continus functions too heuristically: the change in the function over a tiny set is tiny.

THEOREM. *For f non-negative integrable, the function $g(x) = \int_a^x f(x)dx$ is absolutly continuous. More precisly, for all $\epsilon > 0$ there exists $\delta > 0$ so that $\mu(A) < \delta \implies \int_A f dx < \epsilon$*

PROOF. Can do this using the equivalence with absoltule continuity for measurs above, thinking of $\int_A f$ as a measure. Its not hard by bare hands either though: take $f_n = f \wedge n$ so that $f_n \uparrow f$. By LDCT we have:

$$\int f_n \to \int f$$

Hence, for any $\epsilon > 0$, there exits n large enough so that $\int f \mathbb{1}_{\{f>n\}} < \epsilon/2$. Now take $\delta = \frac{\epsilon/2}{n}$ and we find that for any set A with $\mu(A) < \delta$ that:

$$\begin{aligned} \int_A f &= \int_{A\cap\{f\leq n\}} f + \int_{A\cap\{f>n\}} f \\ &\leq n\mu(A) + \epsilon/2 \\ &< \epsilon/2 + \epsilon/2 \end{aligned}$$

□

LEMMA. *If f is absolutly continuous on $[a, b]$ then it is of bounded variation on $[a, b]$*

PROOF. Take $\epsilon = 1$ to get a δ in the definition of absolute continuity. Let $K = 1 + \frac{b-a}{\delta}$ be the largest number of intervals of size δ one could cram into a, b. The varaition on each interval of size δ is no more than 1 (by the definition of absolutle continiuity) and so the total variation on the whole $[a, b]$ can be no more than K, the total number of such intervals. □

COROLLARY. *If f is absolutly continuos then f has a derivative almost everywhere.*

LEMMA. *If f is abs. continuous on $[a, b]$ and $f'(x) = 0$ a.e. then f is constant.*

PROOF. (Uses vitalli covering lemma) Fix ϵ and η arbitrarily small. Get the δ from absolute continuity of f. The idea is as follows, for any interval (a, c) we use the fact that $|f'(x)| < \eta$ everywhere to make a Vitalli covering of $[a, c]$ with $|f(x + h) - f(x)| \leq \eta h$ on each interval $[x, x + h]$. Then by the Vitalli covering lemma, extract a finite set of intervals that cover all of $[a, c]$ except a set of size δ.

Beacuse $|f(x+h)-f(x)| \leq \eta h$ on the intervals, the total contribution to $|f(c)-f(a)|$ on the intervals is no more than $\eta(c-a)$.

Because the δ was chosen by the abs. cont. of f at ϵ, the total the total contribution to $|f(c)-f(a)|$ on the complement of the intervals (which is size no more than δ) is $< \epsilon$.

In sum $|f(c)-f(a)| < \epsilon + \eta(c-a)$ which can be made arbitarily small as we please. □

COROLLARY. *Every absolutly continuous function can be written as:*

$$f(x) = f(a) + \int_a^x f'(x)\,dx$$

PROOF. Let $g(x) = f(x) - \int_a^x f'(x)\mathrm{d}x$. Since indefinite integrals are abs continuous, g is the differnce of absolutely continuous functions and is hence abs. continuous. Moreover, $g' = f' - f' = 0$ a.e. By the previous lemma, then g is a constant! $g(a) = f(a)$ then gives the result. □

L^p spaces

These are notes from Chapter 15 of [**1**].

Let $(X, \mathcal{A}, \mu)$ be a $\sigma-$finite meausre sapce. For $1 \leq p < \infty$ define the L^pnorm of f by:

$$\|f\|_p = \left(\int |f(x)|^p \, \mathrm{d}x \right)^{1/p}$$

For $p = \infty$ $\|f\|_\infty$ is the essential supremum of f. The space L^p is the set of functions whose L^pnorm is finite.

11.28. Norms

PROPOSITION. *(15.1.) (Holder's Inequality) If* $1 < p, q < \infty$ *and* $p^{-1} + q^{-1} = 1$ *then:*

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q$$

PROOF. Assume WOLOG that $\|f\|_p = \|g\|_q = 1$ and then integrat$F(x)G(x) \leq \frac{F(x)^p}{p} + \frac{G(x)^q}{q}$ to get the result. □

PROPOSITION. *(Minkowski's Inequality) If* $1 \leq p \leq \infty$ *then:*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

PROOF. Write:

$$\begin{aligned} |f+g|^p &= |f+g| \, |f+g|^{p-1} \\ &\leq |f| \, |f+g|^{p-1} + |g| \, |f+g|^{p-1} \end{aligned}$$

Now apply Holder's inequality with p and $q = \left(1 - \frac{1}{p}\right)^{-1}$ to get:

$$\begin{aligned} \int |f+g|^p &\leq \int |f| \, |f+g|^{p-1} + \int |g| \, |f+g|^{p-1} \\ &\leq \|f\|_p \left(\int |f+g|^{(p-1)q} \right)^{1/q} + \|g\|_p \left(\int |f+g|^{(p-1)q} \right)^{1/q} \end{aligned}$$

This gets us to:

$$\|f+g\|_p^p \leq \left(\|f\|_p + \|g\|_p \right) \|f+g\|_p^{p/q}$$

And dividing out gives the resutl. □

11.29. Completeness

THEOREM. *(15.4.) If $1 \le p \le \infty$ then L^p is complete.*

PROOF. Say f_n is a Cauchy sequence. Look at the subsequence n_k so that $\left\| f_{n_k} - f_{n_{k+1}} \right\| \le 2^{-k}$. We will show that $\sum_m \left(f_{n_m} - f_{n_{m-1}} \right)$ converges absolutely and this will be our limit point for the Cauchy sequence. Let $g_k = \sum_{i=1}^k \left| f_{n_m} - f_{n_{m-1}} \right|$By Minkowski we have:

$$\begin{aligned} \|g_j\|_p &\le \sum_{i=1}^{k} \left\| f_{n_m} - f_{n_{m-1}} \right\|_p \\ &\le \sum_{m=1}^{j} 2^{-m} < 1 \end{aligned}$$

Let g be the pointwise limit $g = \lim_{n\to\infty} g_n$. By Fatou, we have $\int |g(x)|^p \le \lim_{m\to\infty} \|g_j\|_p^p \le 1$. Hence g must be finite a.e.

Now have $f(x) = \lim_{K\to\infty} \sum_{m=1}^K \left(f_{n_m} - f_{n_{m-1}} \right)(x) = \lim_{K\to\infty} f_{n_K}(x)$ and by Fatou we have:

$$\|f - f_{n_k}\|_p^p = \int \left| f - f_{n_j} \right|^p \le \liminf_{K\to\infty} \int \left| f_{n_K} - f_{n_j} \right|^p \le \sum \left\| f_{n_{k+1}} - f_{n_k} \right\|_p^p \le 2^{(-j+1)p} \to 0$$

So $f_{n_{kl}} \to f$ in L^p as desired. □

PROOF. To prove a space is complete, it suffices to prove that every absolutly convergent sequence is convergent. Suppose h_n is a sequence with $\sum \|h_n\|_{L^p} < \infty$. We want to define $f(x) = \sum_{n=1}^\infty h_n(x)$ and show that this is finite a.e. and is the L_p limit of $\sum_{n=1}^\infty h_n(x)$.

Claim 1: The pointwise limit $f(x) = \lim_{N\to\infty} \sum_{n=1}^N h_n(x)$ exists and is finite a.e.

Pf: Let $g_N(x) = \sum_{n=1}^N |h_n(x)|$ and let $g(x) = \sup_N g_N(x) = \lim_{N\to\infty} g_N(x)$. By Fatou's lemm:

$$\begin{aligned} \int (g(x))^p \,\mathrm{d}x &= \int \lim_{N\to\infty} (g_N(x))^p \,\mathrm{d}x \\ &\le \liminf_{N\to\infty} \int (g_N(x))^p \,\mathrm{d}x \\ &\le \liminf_{N\to\infty} \sum_{n=1}^{N} \|h_n\|_p^p \text{ by Minkowski inequality} \\ &< \infty \text{ since } h_n \text{ is absolutly summable} \end{aligned}$$

Hence $g(x)$ is finite a.e.. Hence for a.e. x the sum $f(x) = \sum_{n=1}^\infty h_n(x)$ converges absolutly and is finite.

Claim 2: $f(x) = \lim_{N\to\infty} \sum_{n=1}^N h_n(x)$ in L^p

Pf: Define $f_N(x) = \sum_{n=1}^{N} h_n(x)$ so $f_N \to f$ pointwise for a.e. x. Now again by Fatou:

$$\begin{aligned}
\int |f - f_N|^p \, \mathrm{d}x &= \int \lim |f_M - f_N|^p \, \mathrm{d}x \\
&\leq \liminf_{M\to\infty} \int |f_M - f_N|^p \, \mathrm{d}x \\
&= \liminf_{M\to\infty} \int \left| \sum_{k=N+1}^{M} h_k(x) \right|^p \mathrm{d}x \\
&\leq \liminf_{M\to\infty} \sum_{n=N+1}^{M} \|h_n\|_p^p \\
&\leq \sum_{n=N+1}^{M} \|h_n\|_p^p \\
&\to 0 \text{ as } N \to \infty
\end{aligned}$$

□

PROPOSITION. *The set of continuous functions with compact support is dense in $L^p(\mathbb{R})$*

PROOF. Compact sets is clear since $\int \left| f - f 1_{[-n,n]} \right| \to 0$ by LDCT. To do continuous function, approximate f by simple functions. Approximate each measurable set by a closed set. Then approximate the characteristic function of a closed set by a continuous function. □

11.30. Convolutions

The **convolution** of two measurable functions is defined by:

$$f * g = \int f(x-y)g(y)\mathrm{d}y$$

PROPOSITION. *(15.7.) If $f, g \in L^1$ then $f * g$ is in L^1 and:*

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1$$

PROOF. Write:

$$\begin{aligned}
\|f * g\|_1 &\leq \int \int |f(x-y)| \, \mathrm{d}x \, |g(y)| \, \mathrm{d}y \\
&= \int \int |f(x)| \, \mathrm{d}x \, |g(y)| \, \mathrm{d}y \\
&= \|f\|_1 \|g\|_1
\end{aligned}$$

□

11.31. Bounded Linear Functionals

THEOREM. *(15.8) For $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$ we have:*

$$\|f\|_p = \sup \left\{ \int fg \, d\mu : \|g\|_q \leq 1 \right\}$$

PROOF. $\geq$ follows by the Holder inequality. To see the other way choose something like $g = (\text{sgn}(f(x)) \frac{|f(x)|^{p-1}}{\|f\|_p^{p/q}}$. □

PROPOSITION. *Actually it suffices to consider only simple functions:*

$$\|f\|_p = \sup\left\{\int fg d\mu : \|g\|_q \leq 1, \ g \textit{ is a simple function}\right\}$$

PROOF. Approximate f by simple functions and put again $g = (\text{sgn}(f(x)) \frac{|s_n(x)|^{p-1}}{\|s_n\|_p^{p/q}}$ □

PROPOSITION. *(15.10) If $1 < p < \infty$ with $p^{-1} + q^{-1} = 1$ and $g \in L^q$ then integration against g defines a bounded linear functional on L^p with $\|H\| = \|g\|_q$*

PROOF. By Holder H is bounded and by the above prop $\|H\| = \|g\|_q$. □

THEOREM. *(15.11) If $1 < p < \infty$ and H is a real valued bounded linear functional on L^p then there exists $g \in L^q$ such that $H(f) = \int fg$ and $\|g\|_q = \|H\|$*

PROOF. The idea is to define the measure $\nu(A) = H(\mathbb{1}_A)$ then show that ν is a measure with $\nu \ll \mu$. The function we want is $g = \frac{d\nu}{d\mu}$. Once we have established that $H(\mathbb{1}_A) = \int g d\mu$ it is easy to see that for simple functions we have $H(s) = \int sg d\mu$. Bu then $\|g\|_q = \sup\left(\int gs : \|s\| \leq 1, s \text{ is simple}\right) = \sup\left(H(s) : \|s\| \leq 1, \text{s is simple}\right) \leq \|H\|$. Hence $g \in L^q$. Finally then, by approximating any function f by simple functions one can check $H(s_n) \to H(f)$ which shows that $H(f) = \int fg$. □

Fourier Transforms

These are notes from Chapter 16 of [**1**].

12.32. Basic Properties

For $f \in L^1(\mathbb{R}^n)$ define:

$$\hat{f}(u) = \int_{\mathbb{R}^n} e^{iu\cdot x} f(x)\mathrm{d}x$$

PROPOSITION. *If f, g in L^1 then:*
i) $\hat{f}$ is bounded and continuous
ii) $\hat{\cdot}$ is linear

PROOF. Write:

$$\begin{aligned} \hat{f}(u+h) - \hat{f}(u) &= \int \left(e^{i(u+h)\cdot x} - e^{iu\cdot x}\right) f(x)\mathrm{d}x \\ \left|\hat{f}(u+h) - \hat{f}(u)\right| &\leq \int \left|e^{iu\cdot x}\right| \left|e^{ih\cdot x} - 1\right| \left|f(x)\right| \mathrm{d}x \end{aligned}$$

which is bounded by $2\,|f(x)|$. Since $e^{ix\cdot h} - 1 \to 0$ as $h \to 0$ the integral $\to 0$ by the LDCT. □

12.32.1. Derivatives.

PROPOSITION. *(16.2.) If $f \in L^1$ and $xf \in L^1$ then $\hat{f}$ is differentiable with:*

$$\hat{f}' = i \int e^{iu\cdot x} x f(x)\, dx$$

REMARK. In $\mathbb{R}^n$ the condition is that $x_j f \in L^1 \implies \hat{f}$ has a partial u_j deriavive.

PROOF. Write:

$$\frac{\hat{f}(u+h) - \hat{f}(u)}{h} = \int e^{iu\cdot x} \left(\frac{e^{ihx} - 1}{h}\right) f(x)\mathrm{d}x$$

And since $xf(x)$ is integrable, we use $\left|\frac{e^{ihx}-1}{h}\right| \leq |x|$ then apply LDCT and get the result. □

PROPOSITION. *If $f \in L^1$, f is absolutly continuous and $f' \in L^1$ then the Fourier transform of f' is $-iu\hat{f}(u)$.* |

PROOF. The trick is to do integration by parts:

$$\hat{f'} = \int_{-\infty}^{\infty} e^{iux} f'(x)\mathrm{d}x = \int iue^{iux} f(x)\mathrm{d}x + boundaryterms$$

The fact that $f'$is in L^1 can be used to help show that $f \to 0$ as $x \to \pm\infty$ that the boundary terms die. □

PROPOSITION. *If $f, g \in L^1$ then $\widehat{f * g} = \hat{f}\hat{g}$*

PROOF. Write it out and use $\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}$ □

12.33. The Inversion theorem

PROPOSITION. *Let $H_a : \mathbb{R} \to \mathbb{R}$ be the pdf of a Gaussian with mean 0 and variance a^2 then:*

$$H_a(x) = \frac{1}{a\sqrt{2\pi}} e^{-x^2/2a^2}$$

Then, $\hat{H}_a$ is the pdf of a Gaussian with mean 0 and variance $1/a^2$ and NO PRECONSTANTS:

$$\hat{H}_a(u) = e^{-a^2u^2/2}$$

Moreover, H_a is an approximate identity as $a \to 0$, i.e. we have:

$$\begin{aligned} \sup \|H_a\|_{L^1} &< \infty \\ \int H_a &= 1 \\ \int_{|x|>\delta} H_a &\to 0 \text{ as } a \to 0 \text{ for any fixed } \delta > 0 \end{aligned}$$

Consequenctly we have:

$$\begin{aligned} H_a * f &\xrightarrow{\|\cdot\|_\infty} f \text{ for all } f \in C^\infty(\mathbb{R}) \\ H_a * f &\xrightarrow{\|\cdot\|_p} f \text{ for all } f \in L^p(\mathbb{R}) \end{aligned}$$

Also notice that $\hat{H}_a(u) \to 1$ as $a \to 0$.

THEOREM. *If f and $\hat{f}$ are in L^1 then:*

$$f(y) = \frac{1}{2\pi} \int e^{-iu\cdot y} \hat{f}(u) du \text{ converges a.e.}$$

PROOF. The trick is to hit things with H_a, then unload the integration onto H, which we know is its own Fourier transform essentially.

Write:

$$\begin{aligned} \frac{1}{2\pi} \int e^{-iu\cdot y} \hat{f}(u)\mathrm{d}u = \lim_{a\to 0} \int e^{-iu\cdot y} \hat{f}(u)\hat{H}_a(u)\mathrm{d}u &= \int\int e^{iux} e^{-iu\cdot y} \hat{f}(u)\hat{H}_a(u)\mathrm{d}x\mathrm{d}u \\ &= C \int H_a(x-y) f(x)\mathrm{d}x \\ &= CH_a * f \\ &\to_{L_1} f \end{aligned}$$

□

Fourier Series

These are notes from Chapter 1 of [**2**]. These are rather informal and do not follow very closely with the textbook.

DEFINITION. For a function $f \in L^1(\mathbb{T})$ we define its n-th Fourier coefficient by:

$$\hat{f}(n) = \frac{1}{2\pi} \int f(t) e^{-int} \mathrm{d}t$$

And tis Fourier series is the <u>formal</u> trigonmentric series:

$$S[f] = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{int}$$

PROPOSITION. *Have:*

$$\left| \hat{f}(n) \right| \leq \|f\|_{L^1}$$

PROOF. Holds since $\left| e^{int} \right| = 1$ □

DEFINITION. A summability kernal is a seqeunce $\{k_n\}$ of 2π periodic functions such that:

$$\begin{aligned} \frac{1}{2\pi} \int k_n(t) \mathrm{d}t &= 1 \\ \frac{1}{2\pi} \int |k_n(t)| \, \mathrm{d}t = \|k_n\|_{L^1} &\leq C \text{ for some fixed } C \text{ and for all } n \\ \lim_{n\to\infty} \int_{\delta}^{2\pi-\delta} |k_n(t)| \, \mathrm{d}t &= 0 \text{ for every fixed } \delta \end{aligned}$$

LEMMA. *If k is a summability kernal then for <u>continuous</u> $\varphi \in C(\mathbb{T})$ we have: :*

$$\lim_{n\to\infty} \frac{1}{2\pi} \int k_n(s) \varphi(s) ds = \varphi(0)$$

PROOF. Given any $\epsilon > 0$, find δ so small by the continuity of φ so that $|\varphi(x) - \varphi(0)| < \epsilon$ whenever $|x| < \delta$. Then:

$$\begin{aligned}\left|\frac{1}{2\pi}\int k_n(s)\varphi(s)\mathrm{d}s-\varphi(0)\right| &= \left|\frac{1}{2\pi}\int k_n(s)\left(\varphi(s)-\varphi(0)\right)\mathrm{d}s\right| \\ &\leq \frac{1}{2\pi}\int_{\delta}^{2\pi-\delta}|k_n(s)|\,|\varphi(s)-\varphi(0)|\,\mathrm{d}s+\frac{1}{2\pi}\int_{-\delta}^{\delta}|k_n(s)|\,|\varphi(s)-\varphi(0)|\,\mathrm{d}s \\ &\leq 2\|\varphi\|_\infty\left(\frac{1}{2\pi}\int_{\delta}^{2\pi-\delta}|k_n(s)|\right)+\epsilon\frac{1}{2\pi}\int_{-\delta}^{\delta}|k_n(s)|\,\mathrm{d}s \\ &\to 0+\epsilon C\end{aligned}$$

□

PROPOSITION. *If $f\in L^1(\mathbb{T})$, and k_n a summability kernal, then we have:*

$$k_n * f \xrightarrow{L^1} f$$

PROOF. We first prove it for f continuous. If f is continuous, then for any ϵ find δ so small so that $|f(x)-f(y)|<\epsilon$ for $|x-y|<\delta$. Consider:

$$\begin{aligned}\left\|\frac{1}{2\pi}\int k_n(s)f(\cdot-s)\mathrm{d}s-f(\cdot)\right\|_{L^1} &= \frac{1}{(2\pi)^2}\int\left|\int k_n(s)\left(f(t-s)-f(t)\right)\mathrm{d}s\right|\mathrm{d}t \\ &= \frac{1}{(2\pi)^2}\int\left|\left(\int_{-\delta}^{\delta}k_n(s)\left(f(t-s)-f(t)\right)\mathrm{d}s+\int_{\delta}^{2\pi-\delta}k_n(s)\left(f(t-s)-f(t)\right)\mathrm{d}s\right)\right. \\ &\leq \frac{1}{(2\pi)^2}\int\left|\|k_n\|_{L^1}\epsilon+2\|f\|_\infty\int_{\delta}^{2\pi-\delta}k_n(s)\mathrm{d}s\right|\mathrm{d}t \\ &\to 0+\epsilon\sup_n\|k_n\|_{L^1}\end{aligned}$$

This proves it for continuous f. For $f\in L^1$, find g continuous with $\|f-g\|_{L^1}<\epsilon$ and consider:

$$\begin{aligned}\|k_n*f-f\|_{L^1} &\leq \|k_n*f-k_n*g\|_{L^1}+\|k_n*g-g\|_{L^1}+\|g-f\|_{L^1} \\ &\leq \left(\sup_n\|k_n\|_{L^1}\right)\|f-g\|_{L^1}+\|g-f\|_{L^1}+\|k_n*g-g\|_{L^1}\end{aligned}$$

□

DEFINITION. The Fejer Kernal is the summability kernal:

$$\begin{aligned}K_n(t) &= \frac{1}{n+1}\sum_{j=0}^{n}\left(\sum_{k=-j}^{j}e^{ikt}\right) \\ &= \sum_{j=-n}^{n}\left(1-\frac{|j|}{n+1}\right)e^{ijt}\end{aligned}$$

PROPOSITION. *Have the identity:*

$$K_n(t)=\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}}{\sin\frac{1}{2}t}\right)^2$$

PROOF. Use the identity:

$$\begin{aligned}\cos(t) &= \cos\left(2\left(\frac{t}{2}\right)\right)\\ &= \cos\left(\frac{1}{2}t\right)^2 - \sin\left(\frac{1}{2}t\right)^2\\ &= 1-2\sin\left(\frac{1}{2}t\right)^2\end{aligned}$$

So:

$$\sin\left(\frac{1}{2}t\right)^2 = \frac{1}{2}\left(1-\cos(t)\right) = -\frac{1}{4}e^{-it}+\frac{1}{2}-\frac{1}{4}e^{it}$$

Then check that:

$$\left(-\frac{1}{4}e^{-it}+\frac{1}{2}-\frac{1}{4}e^{it}\right)\sum_{j=-n}^{n}\left(1-\frac{|j|}{n+1}\right)e^{ijt}$$

turns it into a telescoping sum. □

DEFINITION. We define:

$$\sigma_n(f)(t) = \left(K_n * f\right)(t) = \sum_{k=-n}^{n}\left(1-\frac{|j|}{n+1}\right)\hat{f}(j)e^{ijt}$$

REMARK. Since K_n is a summability kernal, we know that:

$$\begin{aligned}\sigma_n(f) &\xrightarrow{ptwise} f \text{ for } f\in C(\mathbb{T})\\ \sigma_n(f) &\xrightarrow{L^1} f \text{ for } f\in L^1(\mathbb{T})\end{aligned}$$

The pointwise convergence is actually uniform (go back to our argument we used) and the convergence in L^1 can be moved up to L^p convergence with the same method of proof as before (just approximate the L^p function by a contunous function)

$$\begin{aligned}\sigma_n(f) &\xrightarrow{\|\cdot\|_\infty} f \text{ for } f\in C(\mathbb{T})\\ \sigma_n(f) &\xrightarrow{L^p} f \text{ for } f\in L^p(\mathbb{T})\end{aligned}$$

REMARK. (This is kind of off topic...) Using the same kind of naive-ish estimates as above, one can show that the Fourier series converge in C^α, the space of Holder continuous functions:

$$S_n(f)\xrightarrow{\|\cdot\|_\infty} f \text{ for } f\in C^\alpha(\mathbb{T})$$

13.33.1. Corollaries to Fejer Convergence.

THEOREM. *(Fejer Convergence Thm)*

$$\begin{aligned}\sigma_n(f) &\xrightarrow{\|\cdot\|_\infty} f \textit{ for } f\in C(\mathbb{T})\\ \sigma_n(f) &\xrightarrow{L^p} f \textit{ for } f\in L^p(\mathbb{T})\end{aligned}$$

COROLLARY. *Trig. polynomials are dense in $C(\mathbb{T})$, $L^p(\mathbb{T})$ for $1\le p<\infty$*

PROOF. Each $\sigma_n(f)$ is a trig. polynomial. □

COROLLARY. *The functions* $\left\{e^{int}\right\}_{n=-\infty}^{\infty}$ *form a maximal orthonormal set (i.e. an orthonormal set) for* $L^2(\mathbb{T})$.

PROOF. Its easy to check they are orthonormal. To see they are maximal, suppose by contradiction that $f \perp e^{int}$ for all n. Find g a trig poly. so that $\|f-g\|_{L^2} < \epsilon$. Then we have that:

$$\|f\|^2 = \langle f, f\rangle = \langle f, f-g\rangle + \langle f, g\rangle \leq \|f\|\,\|f-g\| + 0 \leq \epsilon\,\|f\|$$

So we get $\|f\| = 0$. □

COROLLARY. *We have:*

$$\begin{aligned} \|f\|_2^2 &= \sum_{n\in\mathbb{Z}} \left|\hat{f}(n)\right|^2 \\ f &= \lim_{n\to\infty} \sum_{i=-n}^{n} \hat{f}(n)e^{int} = \lim_{n\to\infty} \sum_{i=-n}^{n} \left\langle f, e^{int}\right\rangle e^{int} \\ \int f(x)\overline{g(x)}dx = \langle f, g\rangle_{L^2} &= \sum_{n\in\mathbb{Z}} \left\langle f, e^{int}\right\rangle \left\langle e^{int}, g\right\rangle = \sum_{n\in\mathbb{Z}} \hat{f}(n)\overline{\hat{g}(n)} \end{aligned}$$

PROOF. These follow by basic Hilbert space theory and are equivalent to the fact that e^{int} are a basis. The first one is Plancharel's identity, the second one is Riesz-Fisher, and the last one is Parseval's identity. □

COROLLARY. *(Uniquenss Theorem) If* $\hat{f}(n) = 0$ *for all* n *and* $f \in L^1(\mathbb{T})$ *then* $f \equiv 0$

PROOF. $\sigma_n(f) = 0$ for every f and $\sigma_n(f) \to f$ □

COROLLARY. *(Riemann-Lebesgue) If* $f \in L^1(\mathbb{T})$ *then* $\hat{f}(n) \to 0$ *as* $n \to \pm\infty$

PROOF. Given any $\epsilon > 0$, choose N so large so that $\|\sigma_N f - f\|_{L^1} < \epsilon$. Then for any $n > N$, we have that $\widehat{\sigma_n(f)} = 0$ and so:b vb

$$\left|\hat{f}(n)\right| = \left|\widehat{\sigma_N f(n)} - \hat{f}(n)\right| \leq \|\sigma_N f - f\|_{L^1} < \epsilon$$

□

REMARK. As we have seen in Hang's harmonic class, given any sequence $a_n \to 0$ there is an L^1 function whose fourier seires go to 0 slower than a_n. (You can construct it as long as its convex)

13.34. More Fourier Facts

PROPOSITION. *Fejer's Theorem: For continuous function the Fejer kernal* $F_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} D_k(x)$ *where* D_k *is the Dirichlet kernal* $D_k(x) = \sum_{s=-k}^{k} e^{isx}$*.(This can be written as* $F_n(x) = \frac{1}{n}\frac{1-\cos(nx)}{1-\cos(x)}$ *) has* $F_n * f \to f$ *uniformly in* $[-\pi, \pi]$ *if* f *is continuous.*

PROOF. Write

$$\begin{aligned}(F_n * f - f)(y) &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} F_n(x) f(y-x)\mathrm{d}x - f(y) \\ &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} F_n(x)\left[f(y-x) - f(y)\right]\mathrm{d}x \text{ since Fejer kernal integrates to 1} \\ &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} F_n(x)\left[f_{y-\cdot}(x) - f_{y-\cdot}(0)\right]\mathrm{d}x\end{aligned}$$

Where $f_{y-\cdot}(z) := f(y-z)$. Now divide the region of integration into two zones, one where x is small and one where x is large. When x is small, the continuity of f can be used to control it. When x large, the fact that the the Fejer kernal has $\int_{|x|>\delta} |F_n| \to 0$. □

PROPOSITION. *If $f \in L^2$ then the Fourier series approx for f converges in the L^2 sense $D_n * f \to f$ in L^2*

PROOF. This amounts to showing that the trigonometric polynomials e^{inx} form a BASIS for the Hilbert space L^2. Indeed, we first remark that these are dense in L^2 (by Stone-Weirestrass or by the Fejer kernal we know that these are dense in the continuous functions....and the continuous functions are dense in L^2). Next, suppose by contradiction that e^{inx} is not a basis for L^2. Then there exists an $f \in L^2$ perp to all e^{inx}. But then find a sequence of trigonometric polynomials $f_n \to f$ in L^2. We have then:

$$\begin{aligned}\|f\|^2 &= \langle f, f\rangle \\ &= \langle f - f_n, f\rangle \text{ since } \langle f_n, f\rangle = 0 \\ &\leq \|f - f_n\|\,\|f\| \\ &\to 0 \text{ since } f_n \xrightarrow{L^2} f\end{aligned}$$

□

PROPOSITION. *The fourier series of a continuous function does not nessisarily converge pointwise.*

PROOF. (Using the Banach Steinhouse principle). Notice that the Dirichlet kernal D_n has $\|D_n\|_1 \sim \log n \to \infty$. If $D_n * f \to f$ pointwise, then we would have that $\{x : \sup_n |D_n * f(x)| < \infty\} = [-\pi, \pi]$. By the uniform boundeness principle, we would have that $\sup_n \|D_n * f\|_\infty$ is uniformly bounded! However, $\|D_n * f\|_\infty = \|f\|_\infty \|D_n\|_1 \to \infty$ so this is a contradiction. □

Bibliography

[1] R.F. Bass. *Real Analysis for Graduate Students*. CreateSpace Independent Publishing Platform, 2013.

[2] Y. Katznelson. *An Introduction to Harmonic Analysis*. Cambridge Mathematical Library. Cambridge University Press, 2004.

[3] H.L. Royden. *Real Analysis*. Mathematics and statistics. Macmillan, 1988.