ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
“ΜΑΘΗΜΑΤΙΚΑ (ΑΝΑΛΥΣΗ - ΑΛΓΕΒΡΑ - ΓΕΩΜΕΤΡΙΑ)”

MAXIMAL OPERATORS ON EUCLIDEAN SPACES

Αθανάσιος Γκρέπης

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΙΩΑΝΝΙΝΑ, 2020

Το αφιερώνω στην οικογένεια μου.

Η παρούσα Μεταπτυχιακή Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στα

Μαθηματικά (Ανάλυση-Άλγεβρα-Γεωμετρία)

που απονέμει το Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων.

Εγκρίθηκε την 30/07/2020 από την εξεταστική επιτροπή:

Ονοματεπώνυμο	**Βαθμίδα**
Ελευθέριος Νικολιδάκης	Επίκουρος Καθηγητής (Επιβλέπων)
Ιωάννης Γιαννούλης	Αναπληρωτής Καθηγητής
Ανδρέας Τόλιας	Επίκουρος Καθηγητής

ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

Δηλώνω υπεύθυνα ότι η παρούσα διατριβή εκπονήθηκε κάτω από τους διεθνείς ηθικούς και ακαδημαϊκούς κανόνες δεοντολογίας και προστασίας της πνευματικής ιδιοκτησίας. Σύμφωνα με τους κανόνες αυτούς, δεν έχω προβεί σε ιδιοποίηση ξένου επιστημονικού έργου και έχω πλήρως αναφέρει τις πηγές που χρησιμοποίησα στην εργασία αυτή.

Αθανάσιος Γκρέπης

Ευχαριστιες

Σε αυτό το σημείο οφείλω να ευχαριστήσω τον επιβλέποντα καθηγητή μου κύριο Ελευθέριο Νικολιδάκη. Η συμβολή του ήταν πολύτιμη ώστε να πραγματοποιηθεί αυτή η εργασία. Επίσης να ευχαριστήσω τα άλλα δυο μέλη της εξεταστικής επιτροπής, τους κυρίους Ιωάννη Γιαννούλη και Ανδρέα Τόλια. Είχα την τύχη να έχω και τον κύριο Νικολιδάκη αλλά και τους κυρίους Γιαννούλη και Τόλια ως διδάσκοντες σε μαθήματα των προπτυχιακών και μεταπτυχιακών μου χρόνων. Επίσης θα ήθελα να ευχαριστήσω τους υπόλοιπους καθηγητές που είχα σε μεταπτυχιακό επίπεδο σε ανάλυση, άλγεβρα και γεωμετρία καθώς και γενικότερα το Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων για τα εφόδια που μου έχει προσφέρει. Επιπλέον ένα μεγάλο ευχαριστώ στους φίλους που απέκτησα στο τμήμα όλα τα χρόνια που ήμουν εκεί είτε ως προπτυχιακός είτε ως μεταπτυχιακός φοιτητής. Τέλος να ευχαριστήσω την οικογένεια μου για την στήριξη και φροντίδα που μου παρέχουν όλα τα χρόνια της ζωής μου.

Περιληψη

Το αντικείμενο της παρούσας μεταπτυχιακής διατριβής είναι η μελέτη μεγιστικών τελεστών και έχει την ακόλουθη δομή. Στο πρώτο κεφάλαιο εισάγουμε δύο έννοιες, αυτή της συνάρτησης κατανομής καθώς και της φθίνουσας αναδιάταξης μιας συνάρτησης. Κλείνοντας εισάγουμε τον πρώτο μας μεγιστικό τελεστή. Στο επόμενο κεφάλαιο εισάγουμε τους βασικούς μεγιστικούς τελεστές με τους οποίους θα ασχοληθούμε, τον Hardy-Littlewood και τον Δυαδικό. Όπως θα δούμε η μελέτη του πρώτου και πιο συγκεκριμένα το γεγονός ότι είναι ασθενώς-L^1 θα μας επιτρέψει να αποδείξουμε το θεώρημα διαφόρισης του Lebesgue, καθώς και να αποδείξουμε ότι είναι ασθενώς-L^p όμως για να φτάσουμε σε αυτό το σημείο θα χρειαστεί να μελετήσουμε πρώτα τον Δυαδικό μεγιστικό τελεστή. Μελετώντας τον δεύτερο θα συναντήσουμε σημαντικές μαθηματικές κατασκευές όπως την αποσύνθεση Calderon Zygmund και το θεώρημα παρεμβολής του Marcinkiewicz το οποίο μάλιστα θα το γενικεύσουμε ως ένα βαθμό. Μετά από αυτή τη διαδρομή θα προχωρήσουμε στην απόδειξη κάποιων θεωρημάτων κάλυψης με πιο «εντυπωσιακό» αυτό του Besicovitch. Στη συνέχεια θα μελετήσουμε συνθήκες για βάρη κάτω από τις οποίες ο μεγιστικός τελεστής Hardy-Littlewood είναι ασθενώς-L^p στον αντίστοιχο χώρο με βάρος. Η εργασία ουσιαστικά τελειώνει με τις λεγόμενες συναρτήσεις Bellman. Τα κύριο θέματα αυτού του κεφαλαίου είναι ένα θεώρημα σχετικά με τον δυαδικό μεγιστικό τελεστή καθώς και η εύρεση της ακριβής μορφής της συνάρτησης Bellman τριών μεταβλητών του δυαδικού μεγιστικού τελεστή. Στο τέλος υπάρχει ένα παράρτημα που περιλαμβάνει τη βασική θεωρία μέτρου που είναι προαπαιτούμενο για να κατανοήσει ένας αναγνώστης το τι πραγματεύεται η εργασία.

Abstract

The subject of this master thesis is the study of maximal operators. The structure of this thesis is the following. In the first chapter we introduce two notions that of the distribution function as well as that of the decreasing rearrangement of a given function. We will end this chapter by introducing our first maximal operator. In the next chapter we will introduce the primary maximal operators we will concern ourselves with, the Hardy-Littlewood and the Dyadic one. As we are going to see by studying the first and more specifically its weak-L^1 condition, will allow us to prove Lebesgue's Differentiation theorem and its strong L^p-boundedness but in order to reach to that point we first need to study the Dyadic Maximal Operator. By studying the Dyadic Maximal Operator we will deduce important mathematical constructions like the Calderon-Zygmund Decomposition and Marcinkiewicz's Interpolation Theorem which we will also generalize to some extend. After this journey we will move on showing some covering lemmas with the most "significant" that of Besicovitch. Next we will be studying conditions for weights under which the Hardy-Littlewood Maximal Operator is weak-L^p in the respective weighted space. This thesis essentially ends with the Bellman functions. The main themes of this chapter are a theorem regarding the Dyadic Maximal Operator as well as finding the exact form of the Bellman function of three variables of the Dyadic Maximal Operator. At the end there is an appendix presenting results of the elementary measure theory one needs to know in order to understand what this thesis is concerned with.

Introduction

This is a simple introduction mentioning some of the results concerning this thesis. For more details see inside.

Suppose we want prove Lebesgue's Differentiation theorem which states that for every $f \in L^1_{loc}(\mathbb{R}^n)$,

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f - f(x)| \to 0, \quad \text{for a.e. } \; x \in \mathbb{R}^n.$$

By setting T as follows,

$$Tf(x) := \limsup_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f - f(x)|$$

we will prove that $Tf \leq T^*f + |f|$. Here T^* stands for the Hardy-Littlewood Maximal Operator which we will define down below.

Now using the fact that the space $C_c((R)^n)$ is dense in L^1 and a useful inequality which we are going to give a great deal of time studying, the weak-L^1 inequality we will have a complete proof of the theorem.

A semi-linear operator T defined on the space $L^p(X)$ where (X, μ) a measure space and taking values on the space of measurable functions of a space (Y, ν) is called weak-L^p $(p \geq 1)$ if and only if there is a constant C_p such that

$$\nu(\{Tf > \lambda\}) \leq \left(C_p \frac{\|f\|_{L^p}}{\lambda} \right)^p, \quad \text{for every } f \in L^p(X).$$

In case $p = \infty$ we identify the weak-L^∞ condition with the strong-L^∞ one.

We now give a brief description of each chapter of this thesis.

The first chapter concerns with the basic properties of the distribution function and the decreasing rearrangement of an initial given function. The *distribution function* of a function f defined on a measure space (X, μ) is defined as follows,

$$\mu_f(\lambda) = \mu\left(\{x : |f(x)| > \lambda\}\right), \quad \text{for every } \lambda > 0.$$

With the above definition in mind the *decreasing rearrangement* of f (denoted as f^*) is defined as follows,

$$f^*(t) = \inf\{\lambda > 0 : \mu_f(\lambda) \leq t\}.$$

A basic property of the distribution function and the decreasing rearrangement of a function f which we will use is that we can calculate the L^p norm of f using them. Indeed the following identity holds true

$$\|f\|^p_{L^p(X)} = \int_0^\infty p\lambda^{p-1}\mu_f(\lambda)\,d\lambda = \int_0^\infty f^*(t)^p\,dt.$$

In the second chapter we introduce the two maximal operators we are concerned with throughout this thesis, the Hardy-Littlewood Maximal operator and the Dyadic Maximal Operator. For a $L^1_{loc}(\mathbb{R}^n)$ function f the *Hardy-Littlewood Maximal Operator* is defined as follows[1]

$$\mathcal{M}f(x) := \sup_{Q \ni x} \frac{1}{|Q|}\int_Q |f|, \quad x \in \mathbb{R}^n$$

where the supremum is taken over all cubes Q with faces parallel to the axes that contain x and the *Dyadic Maximal Operator* of f as

$$\mathcal{M}_d f(x) := \sup_{D \ni x} \frac{1}{|D|}\int_D |f|, \quad x \in \mathbb{R}^n$$

where D denotes a dyadic cube that contains x. The dyadic cubes are the cubes that are formed from the grids $2^{-m}\mathbb{Z}^n$ where $m \in \mathbb{Z}$. Regarding the Hardy-Littlewood Maximal Operator the one dimensional case was introduced by Hardy and Littlewood [18] and the multi-dimensional case by Wiener [40].

It also worth mentioning that in this chapter we prove the Marcinkiewicz Interpolation theorem as well as the Lebesgue's Differentiation theorem which we mentioned at the beginning.

In the third chapter we introduce some covering lemmas that are of particular interest in Harmonic Analysis such as the Besicovitch Covering Lemma and the Vitalli Covering Lemma. The Besicovitch Covering Lemma states that for every subset A of $\mathbb{R}^n$ and every covering of A consisting of balls such that each point of A is the center of at least one member of the covering, there exists a number C_n depending only on the dimension of the space and a countable subcovering such that its elements can be distributed over C_n subfamilies so that the elements of each of these families are pairwise disjoint.

In the fourth chapter we will move to studying the A_p weights. A weight w is called an A_p *weight* $(1 \le p < \infty)$ if the Hardy-Littlewood Maximal Operator is weak-L^p_w[2], i.e. if and only if

$$w(\{x : \mathcal{M}_q f(x) > \lambda\}) \le \frac{C}{\lambda^p}\int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx, \quad \text{for every } f \in L^p_w(\mathbb{R}^n)$$

where for a Lebesgue measurable subset A of $\mathbb{R}^n$ and $w(A) := \int_A w\,dx$. Equivalently a weight w is an A_p weight if and only if w is locally integrable and

- If $p = 1$ then $\frac{w(Q)}{|Q|} \le Cw(x)$, for every cube Q and a.e. $x \in Q$;

[1] Some authors define it with balls in place of cubes and they may even restrict themselves to cubes or balls centered at each point $x \in \mathbb{R}^n$

[2] Here we mean the weighted Lebesgue space $L^p_w = \{f \text{ measurable}: \int_{\mathbb{R}^n} |f|^p w\,dx < \infty\}$

- if $p > 1$ then $\frac{w(Q)}{|Q|}\left(\int_Q w^{1-p'}\right)^{p-1} \leq C, \quad$ for every cube Q

for some constant $C > 0$.

For each $p \in [1, \infty)$ the condition described above is called the "A_p condition".

In the fifth and final chapter we concern ourselves with the notion of the Bellman function. More specifically we will prove that the exact form of the Bellman function (of three variables) of the Dyadic Maximal Operator which is defined as

$$B_p^{\mathcal{T}}(f, F, L) = \sup\left\{\int_X \{\mathcal{M}_{\mathcal{T}}\phi, L\}^p : \phi \geq 0, \int_X \phi = f, \int_X \phi^p = F\right\}$$

for all f, F, L such that $0 \leq f^p \leq F$ and $L \geq f$, is the following

$$B_p^{\mathcal{T}}(F, f, L) = \begin{cases} F\omega_p\left(\frac{pL^{p-1}f-(p-1)L^p}{F}\right)^p, & \text{if } f \leq L < \frac{p}{p-1}f \\ L^p + \left(\frac{p}{p-1}\right)^p (F - f^p), & \text{if } L \geq \frac{p}{p-1}f. \end{cases}$$

Contents

CHAPTER 1

Distribution functions and Rearrangements

Let us begin by introducing some of the first new notions that we are going to need, that of the distribution function and the decreasing rearrangement. Let us note that from here on our spaces are going to be σ-finite. This chapter is primary influenced by the book "Interpolation of Operators" of Bennet and Sharpley [2].

1.1 Distribution function

Definition 1.1. Let (X, μ) be a σ-finite measure space and f be a measurable a.e. finite function.

The *distribution function* of f is the function $\mu_f : [0, \infty) \to [0, \infty]$ defined as follows

$$\mu_f(\lambda) := \mu(\{|f| > \lambda\})^1$$

Before we continue with new definitions let us calculate the distribution function of a simple function, which will help us understand the notion better and will be used in the sequel.

Example 1.2 (Distribution function of a simple function)**.** Let (X, μ) be a σ-finite measure space and $f : X \to [0, \infty)$ be a simple function of the form

$$f = \sum_{i=1}^{n} a_i \chi_{A_i},$$

where $\infty > a_1 > a_2 > \cdots > a_n > 0$ and the sets A_i are pairwise disjoint and of finite measure.

Now to find its distribution function.

1. First consider a λ which is greater or equal to a_1. Then we have that $\mu_f(\lambda) = \mu(\{f > \lambda\}) = \mu(\emptyset) = 0$

[1] By that we mean the set $\{x \in X : |f(x)| > \lambda\}$.

2. Then consider all the λ that lie in the interval $[a_2, a_1)$.Then we have that $\mu_f(\lambda) = \mu(A_1)$

3. Similarly we find that for $\lambda \in [a_3, a_2)$ $\mu_f(\lambda) = \mu(A_1 \cup A_2)$, for $\lambda \in [a_4, a_3)$, $\mu_f(\lambda) = \mu(A_1 \cup A_2 \cup A_3)$ and so on.

So by setting $m_i := \mu(A_1 \cup \cdots \cup A_i)$, for $i = 1, \ldots, n$ the distribution function of f can be written as follows,

$$\mu_f(\lambda) = \sum_{i=0}^{n} m_i \chi_{[a_{i+1}, a_i)},$$

where we define $a_0 := \infty$, $a_{n+1} := 0$ and $m_0 := 0$.

Now with the above definition in mind we can define what it means for two a.e. measurable finite functions to be equimeasurable to each other.

Definition 1.3. Let (X, μ), (Y, ν) be two σ-finite measure spaces and f and g be two a.e. measurable finite functions defined on X and Y respectively. The functions f, g are called rearrangements of each other or equivalently *equimeasurable* to each other if they have the same distribution function.

Remark 1.4. Note here that two equimeasurable functions are *NOT* defined in the same measure space. Shortly we are going to define for every function f as above a new function f^* equimeasurable with f defined on $[0, \infty)$. Also it straightforward to see that equimeasurability constitutes an equivalence relation.

Theorem 1.5. *Let (X, μ) be a σ-finite measure space f a measurable a.e. finite function on X and $\mu_f : [0, \infty) \to [0, \infty]$ be its respective distribution function, then μ_f is non-negative, decreasing and right continuous.*

Proof. The first two properties stem directly from the definition. As for the right continuity, due to the fact that μ_f is decreasing it is enough to show that

$$\mu_f\left(\lambda + \frac{1}{n}\right) \to \mu_f(\lambda), \quad \text{for every } \lambda \geq 0.$$

If we set $E(\lambda) := \{|f| > \lambda\}$ for any $\lambda \geq 0$ then

$$E(\lambda) = \bigcup_{n \in \mathbb{N}} E\left(\lambda + \frac{1}{n}\right).$$

The sequence to the right is increasing and so by the standard properties of a measure

$$\mu_f\left(\lambda + \frac{1}{n}\right) = \mu\left(E\left(\lambda + \frac{1}{n}\right)\right) \to \mu(E_\lambda)^+ = \mu_f(\lambda). \qquad \square$$

Now we find some properties regarding the distribution function.

Theorem 1.6 (Basic properties of the Distribution function)**.** *Let (X, μ) be a σ-finite measure space and f, g finite a.e. functions. Then the followings hold true.*

1. *If $|f| \leq |g|$ a.e. then $\mu_f \leq \mu_g$.*
2. *If $a \in \mathbb{R}\backslash\{0\}$ then*
$$\mu_{af}(\lambda) = \mu_f\left(\frac{\lambda}{|a|}\right), \quad \text{for every } \lambda \geq 0.$$
3. *For every $\lambda_1, \lambda_2 \geq 0$, $\mu_{f+g}(\lambda_1 + \lambda_2) \leq \mu_f(\lambda_1) + \mu_g(\lambda_2)$.*
4. *If $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable a.e. finite functions such that*
$$|f| \leq \liminf_{n\to\infty} |f_n| \ \mu\text{-a.e.}$$
then
$$\mu_f \leq \liminf_{n\to\infty} \mu_{f_n}.$$
In particular if $f_n(x) \to f^+(x)$ for a.e. $x \in X$ then
$$\lim_{n\to\infty} \mu_{f_n}(\lambda) = \mu_f(\lambda), \quad \text{for every } \lambda \geq 0.$$

Proof. 1. Let $F_\lambda := \{|f| > \lambda\}$ and $G_\lambda := \{|g| > \lambda\}$ then we have that $F_\lambda \preccurlyeq$ [2]G_λ and as a result
$$\mu_f(\lambda) = \mu(F_\lambda) \leq \mu(G_\lambda) = \mu_g(\lambda).$$

2. Let $a \neq 0$ then
$$\mu_{af}(\lambda) = \mu(\{|af| > \lambda\}) = \mu(\{|a||f| > \lambda\}) = \mu\left(\left\{|f| > \frac{\lambda}{|a|}\right\}\right) = \mu_f\left(\frac{\lambda}{|a|}\right)$$

3. We have that $\{x \in X : |f+g| > \lambda_1 + \lambda_2\} \subset \{x \in X : |f(x)| > \lambda_1\} \cup \{x \in X : |g(x)| > \lambda_2\}$ and so passing to the measures we get our result.

4. For $\lambda \geq 0$ let
$$E_\lambda := \{|f| > \lambda\} \quad \text{and} \quad E_n := \{|f_n| > \lambda\}, \quad \text{for } n = 1, 2, \ldots.$$
Then
$$E \preccurlyeq \bigcup_{m=1}^{\infty} \bigcap_{n>m} E_n.$$
Now since
$$\mu\left(\bigcap_{n>m} E_n\right) \leq \inf_{n>m} \mu(E_n) \leq \sup_m \inf_{n>m} \mu(E_n) = \liminf_{n\to\infty} \mu(E_n)$$

[2] here we mean that the set of points of F_λ that lie outside of G_λ has measure equal to zero.

and the sequence $\cap_{n>m} E_n$ is decreasing by considering the basic properties of any measure we obtain that

$$\mu(E) \le \mu\left(\bigcup_{m=1}^{\infty} \bigcap_{n>m} E_n\right) = \lim_{m\to\infty} \mu\left(\bigcap_{n>m} E_n\right) \le \liminf_{n\to\infty} \mu(E_n).$$

Now if also $f_n(x) \to f^+(x)$ for a.e. $x \in X$, from the second property we have

$$\mu_{f_n}(\lambda) \le \mu_f(\lambda), \quad \text{for every } \lambda \ge 0$$

and so

$$\limsup \mu_{f_n}(\lambda) \le \mu_f(\lambda), \quad \text{for every } \lambda \ge 0.$$

Combining this with that we proved just before gives us the result we sought. □

1.2 Decreasing Rearrangement

With the notion of the distribution function of a function f in our hands we now introduce a certain new function f^*(called the decreasing rearrangement of f) which turns out to be the unique non-negative, decreasing and right continuous function which is equimeasurable with f. The most simple example of a decreasing rearrangement is to consider a random finite sequence $\{a_n\}_{n=1}^m$ of real numbers, then its decreasing rearrangement is simply a new sequence $\{a_n^*\}_{n=1}^m$ which has the exact same values as the first but ordered in a decreasing manner.

Definition 1.7. Let (X, μ) be a σ-finite measure space and f be an a.e. finite measurable function on X. The *decreasing rearrangement* of f is the function $f^* : [0, \infty) \to [0, \infty]$ defined by

$$f^*(t) := \inf\{\lambda > 0 : \mu_f(\lambda) \le t\}, \quad t \ge 0.$$

In case the set is empty we use the convention $\inf(\emptyset) := \infty$.

Theorem 1.8. *Let (X, μ) be a σ-finite measure space and $f : X \to [-\infty, \infty]$ be an measurable a.e. finite function. Then its decreasing rearrangement is the only non-negative, decreasing and right continuous function that is equimeasurable to f.*

Proof. Since it is straightforward that f^* is nonnegative and decreasing, *we will concern ourselves only with its right continuity and uniqueness. For the right continuity* we just need to make the following observation. Since μ_f is decreasing, f^* can also be expressed as

$$f^*(t) = \inf\{\lambda > 0 : \mu_f(\lambda) \le t\} = \mu(\{\mu_f > t\}).$$

So it is in fact the distribution function of μ_f and we have already proved that distributions functions are right continuous.

Now to prove the uniqueness of f^*. Let f' be another non negative,decreasing and right continuous that is equimeasurable with f. Since both of them are equimeasurable to f they are equimeasurable to each other so

$$|\{f^* > \lambda\}| = |\{f' > \lambda\}|, \quad \text{for every } \lambda \geq 0.$$

Since both of them are decreasing the respective sets above provided they are not empty, they constitute intervals and more specifically

$$\{f^* > \lambda\} = [0, a) \quad \text{and} \quad \{f' > \lambda\} = [0, b),$$

where $a, b > 0$. The reason both of these sets are open to the right is due to their right continuity. Combining these two results gives us that $a = b$.

Now to prove that f^* and f' coincide. Suppose to the contrary there exists a point t_0 such that $f^*(t_0) \neq f'(t_0)$ and without loss of generality let us also assume that $f^*(t_0) > f'(t_0)$. Then we would have

$$\{f^* > f'(t_0)\} = [0, c) \quad \text{and} \quad \{f' > f'(t_0)\} = [0, d)$$

for some $c > t_0$, $d < t_0$.

The first equality is a consequence of the decreasing monotonicity and the right continuity of f^* and the second stems directly from the fact that f' is decreasing, But we reached a contradiction since their level sets are equal. □

Let us now calculate the decreasing rearrangement of a simple function

Example 1.9 (Decreasing Rearrangement of simple function)**.** Let (X, μ) be a σ-finite measure space and $f : X \to [0, \infty)$ be a simple function of the form

$$f := \sum_{i=1}^{n} a_i \chi_{A_i},$$

where $\infty > a_1 > a_2 > \cdots > a_n > 0$ and A_i are pairwise disjoint measurable sets of finite measure.

As we have shown before its distribution function is the following

$$\mu_f(\lambda) := \sum_{i=0}^{n} m_i \chi_{[a_{i+1}, a_i)}.$$

As for its decreasing rearrangement

1. First let us consider all the non-negative t such that $t < m_1$. In order for $\mu_f(\lambda) \leq t$, λ *must be greater or equal to* a_1.

2. For $t < m_2$ and $t \geq m_1$ in order for $\mu_f(\lambda) \leq t$, λ *must be greater or equal to* a_2.

3. For $t < m_3$ and $t \geq m_2$ in order for $\mu_f(\lambda) \leq t$, λ *must be greater or equal to* a_3 and so on.

So its decreasing rearrangement has the following form

$$f^* = \sum_{i=1}^{n} a_i \chi_{[m_{i-1}, m_i)}.$$

Let us now present another form of the distribution function of a simple function which we will need later on.

Example 1.10 (Alternative form of the distribution function of a simple function)**.** Let f be simple function defined as in Example 1.9. Then f can also be written as follows

$$f = \sum_{i=1}^{n} b_i \chi_{F_i},$$

where $b_k := a_k - a_{k+1}$ and $F_k := \cup_{i=1}^{k} A_i$.

Then the decreasing rearrangement can be written as

$$f^* = \sum_{i=1}^{n} b_i \chi_{[0, \mu(F_i))}.$$

We are now going to study some basic properties of f^*. One may notice that they are similar to the basic properties that we derived above for the distribution function (Theorem 1.6). It is not a coincidence since as we showed the decreasing rearrangement of a function is the distribution function of its distribution function.

Theorem 1.11 (Basic Properties of f^*)**.** *Let X be a σ-finite measure space and f, g be a.e measurable finite functions on X. Then the following hold true.*

1. *If $\lambda, t \geq 0$ then*
$$f^*(\mu_f(\lambda)) \leq \lambda \quad \text{and} \quad \mu_f(f^*(t)) \leq t.$$
2. *If $|f| \leq |g|$ a.e. then $f^* \leq g^*$.*
3. $(af)^* = |a| f^*$.
4. $(f+g)^*(t_1 + t_2) \leq f^*(t_1) + g^*(t_2), \quad t_1, t_2 \geq 0.$
5. *If $\{f_n\}_{n \in \mathbb{N}}$ is a sequence of measurable a.e. finite function on X such that*
$$f(x) \leq \liminf_{n \to \infty} f_n(x), \quad \text{for a.e. } x$$
then,
$$f^*(\lambda) \leq \liminf_{n \to \infty} f_n^*(\lambda), \quad \text{for every } \lambda \geq 0.$$
In particular if $f_n(x) \to f^-(x)$ for a.e. $x \in X$ then,
$$f_n^*(\lambda) = f^{*-}(\lambda), \quad \text{for every } \lambda \geq 0.$$

Proof. As we mentioned above the proofs of these properties are based on the basic properties of the distribution function we proved before (Theorem 1.6).

1. The first inequality stems directly from the definition of $f^*(\mu_f(\lambda))$.

   As for the other inequality, we may find a decreasing sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ such that
   $$\mu_f(\lambda_n) \leq t \quad \text{and} \quad \lambda_n \to f^*(\mu_f(\lambda))$$
   and so our result stems from the right continuity of f^*.

2. Since $|f| \leq |g|$ we have $\mu_f \leq \mu_g$ and so
   $$\{\lambda > 0 : \mu_g(\lambda) \leq t\} \subset \{\lambda > 0 : \mu_f(\lambda) \leq t\}.$$
   Considering the respective infima we have our result.

3. For $a \neq 0$
   $$(af)^*(t) = \inf\{\lambda > 0 : \mu_{af}(\lambda) \leq t\}$$
   and from what we proved before for the distribution function (Theorem 1.6) gives us
   $$(af)^*(t) = \inf\left\{\lambda > 0 : \mu_f\left(\frac{\lambda}{|a|}\right) \leq t\right\}.$$
   Equivalently
   $$(af)^*(t) = |a| \inf\left\{\frac{\lambda}{|a|} > 0 : \mu_f(\frac{\lambda}{|a|}) \leq t\right\} = |a|f^*(t).$$

4. Let us set $\lambda := f^*(t_1) + g^*(t_2)$.

   Then
   $$\begin{aligned} t := \mu_{f+g}(f^*(t_1) + g^*(t_2)) &\leq \mu_f(f^*(t_1)) + \mu_g(g^*(t_2)) \\ &\leq t_1 + t_2. \end{aligned}$$
   So taking into consideration the decreasing monotonicity of the decreasing rearrangement we have
   $$\begin{aligned} (f+g)^*(t_1+t_2) \leq (f+g)^*(t) &= (f+g)^*(\mu_{f+g}(\lambda)) \\ &\leq \lambda = f^*(t_1) + g^*(t_2). \end{aligned}$$

5. As we have shown if $f_n(x) \to f^-(x)$ for a.e. $x \in X$ then $\mu_{f_n}(\lambda) \to \mu_f^-(\lambda)$ for every $\lambda \geq 0$. For the same reason
   $$\mu_{\mu_{f_n}}(\lambda) \to \mu_{\mu_f}^-(\lambda), \quad \text{for evert } \lambda \geq 0$$
   which is equivalent to saying that
   $$f_n^*(\lambda) \to f^{*-}(\lambda), \quad \text{for every } \lambda \geq 0.$$ □

Now let us state a basic theorem that express the L^p norm of a function in terms of its distribution function and its decreasing rearrangement.

Theorem 1.12. *Let (X, μ) be a σ-finite measure space and $f : X \to \mathbb{R}$ be a measurable a.e. finite function. Then*

$$\|f\|^p_{L^p(X)} = \int_0^\infty p\lambda^{p-1}\mu_f(\lambda)\,d\lambda = \|f^*\|^p_{L^p(\mathbb{R})}, \quad \textit{for } p < \infty$$

and

$$\|f\|_{L^\infty(X)} = f^*(0).$$

Proof. To prove the left part of the first identity we are going to need Fubini's theorem, so

$$\begin{aligned}
\|f\|^p_{L^p(X)} = \int_X |f|^p\,d\mu &= \int_X \int_0^{|f(x)|} p\lambda^{p-1}\,d\lambda\,d\mu \\
&= \int_X \int_0^\infty p\lambda^{p-1}\chi_{\{|f|>\lambda\}}\,d\lambda\,d\mu \\
&= \int_0^\infty \int_X p\lambda^{p-1}\chi_{\{|f|>\lambda\}}\,d\mu\,d\lambda \\
&= \int_0^\infty p\lambda^{p-1} \int_X \chi_{\{|f|>\lambda\}}\,d\mu\,d\lambda \\
&= \int_0^\infty p\lambda^{p-1}\mu_f(\lambda)\,d\lambda.
\end{aligned}$$

As for the right side since f^* and f are equimeasurable to each other, they have the same distribution function.

As for the second identity we have

$$f^*(0) = \inf\{\lambda > 0 : \mu(\{|f| > \lambda\}) = 0\} = \inf\{\lambda > 0 : f \leq \lambda \text{ a.e.}\} = \|f\|_{L^\infty}(X). \quad \square$$

We now present a case which we will need later on.

Example 1.13. If f be an non-negative a.e. finite function such that $f^* = 1$ on a set of measure equal to $t < \infty$ and zero elsewhere then f is a.e. equal to the characteristic function of a set of measure t.

To prove that let us first remember that since f and f^* are equimeasurable

$$\mu(\{f > 0\}) = \mu(\{f^* > 0\}) = t.$$

Also from Theorem 1.12 above

$$\|f\|_{L^\infty(X)} = f^*(0) = 1$$

As a result $f \leq 1$ a.e..

Now to prove that $f = 1$ *on a set of measure equal to* t *and* 0 *almost elsewhere.*

If we assume the contrary that there exist an $n \in \mathbb{N}$ such that

$$\mu\left(\left\{f \in \left(0, 1 - \frac{1}{n}\right)\right\}\right) > 0$$

then using the fact that

$$t = \mu\left(\{f > 0\}\right) = \mu\left(\left\{f > 1 - \frac{1}{n}\right\}\right)$$

we reach a contradiction. As a result $f = 1$ on a set of measure equal to t and zero almost elsewhere.

1.3 Hardy's Inequality

In this section we are going to investigate an inequality due to G.H Hardy. Basically our next result is a generalization of Hardy's Inequality for finite sequences i.e.

$$\sum_{i=1}^{n} a_i b_i \leq \sum_{i=1}^{n} a_i^* b_i^*,$$

where the sequences in the right side are just the terms of the respective sequences in the left arranged in a decreasing order.

Theorem 1.14 (Hardy's Inequality)**.** *Let* (X, μ) *be a* σ*-finite measure space and* $f, g : X \to \mathbb{R}$ *be measurable a.e. finite functions. Then the following inequality is true,*

$$\int_X |f||g| \, d\mu \leq \int_{\mathbb{R}} f^* g^* \, dx. \tag{1.1}$$

Proof. Since the inequality above depends only on the absolute values of the function we may assume that f, g are non-negative. We can also assume that f, g are simple functions since then the general case then stems directly from the Monotone Convergence Theorem.

We are going to prove this inequality in two steps.

First we are going to prove a special case of the inequality above and then using that we are going to prove the more general case.

So let $g : X \to \mathbb{R}$ be a non-negative simple function. *We prove that*

$$\int_E g \, d\mu \leq \int_0^{\mu(E)} g^* \, dx$$

for every $E \subset X$ measurable. To prove the above inequality we will use the following representation of g as a simple function

$$g = \sum_{i=1}^{n} b_i \chi_{F_i}, \quad \text{where } F_1 \subset \cdots \subset F_n \text{ and } b_i > 0 \text{ for } i = 1, \ldots, n.$$

With that in mind

$$\int_E g\,d\mu = \sum_{i=1}^n b_i\mu(F_i\cap E) \le \sum_{i=1}^n b_i \min\{\mu(E),\mu(F_i)\}$$
$$= \sum_{i=1}^n b_i \int_0^{\mu(E)} \chi_{[0,\mu(F_i))}(s)\,ds$$
$$= \int_0^{\mu(E)} f^*(s)\,ds.$$

Now we prove inequality (1.1) itself. Let $f = \sum_{i=1}^n a_i\chi_{B_i}$, where $B_1 \subset \cdots \subset B_n$. As we have seen (Example 1.10)

$$f^* = \sum_{i=1}^n a_i\chi_{[0,\mu(B_i))}$$

and so

$$\int_X |f||g|\,d\mu = \sum_{i=1}^n a_i \int_{B_i} g\,d\mu \le \sum_{i=1}^n a_i \int_0^{\mu(B_i)} g^*\,dx$$
$$= \int_0^\infty \sum_{i=1}^n a_i\chi_{[0,\mu(B_i))}(x) g^*(x)\,dx$$
$$= \int_0^\infty f^*g^*\,dx. \qquad \square$$

Note 1.15. With the above inequality in mind one can deduce the following. For every g' equimeasurable with g the same inequality holds with g' in place of g. So one might wonder, under which circumstances if any, there exists a g' equimeasurable to g so that equality occurs. That's the topic we are going to address next.

Definition 1.16. Let (X,μ) be a measure space. A measurable subset B of X is called an *atom* if it is of positive measure and all its measurable subsets have measure equal to it or zero, i.e.

$$\text{for every } A\subset B \text{ so that } A \text{ measurable,} \quad \mu(A)=\mu(B) \quad \text{or} \quad \mu(A)=0.$$

Definition 1.17. A measure space that has no atoms will be called *non-atomic measure space*. If all its measurable subsets of positive measure are atoms then it will be called *completely atomic*.

Theorem 1.18. *Let (X,μ) be a measure space and A an atom of X of finite measure. Then every measurable a.e. finite function $f: X\to\mathbb{R}$ is a.e. constant on A.*

Proof. Let us assume that A is an atom and let us also assume for the beginning that f is a non-negative function defined on X and for $\lambda>0$ set

$$A_\lambda := \{f\ge\lambda\}\cap A.$$

Since A is an atom we have that its measure is either $\mu(A)$ or zero. If for every $\lambda > 0$ the correspondent set has measure $\mu(A)$ then since $\mu(A) < \infty$ and A_λ are decreasing with respect to λ

$$\mu(\{x \in A : f(x) = \infty\}) = \lim_{\lambda\to\infty} \mu(A_\lambda) = \mu(A) > 0.$$

This is not possible since f in finite a.e.. Thus f is essentially bounded from above. If we denote with $\|f|_A\|_{L^\infty}$ the ess sup of f on A, then we have $f(x) \in [0, \|f|_A\|_{L^\infty}]$ for a.e $x \in A$ with $\|f|_A\|_{L^\infty} < \infty$. Dividing the interval to half then gives us two subsets I_1 and I_2 of A defined as

$$I_1 = \left\{ f|_A \in \left[0, \frac{\|f|_A\|_{L^\infty}}{2}\right]\right\} \quad \text{and} \quad I_2 = \left\{ f|_A \in \left[\frac{\|f|_A\|_{L^\infty}}{2}, \|f|_A\|_{L^\infty}\right]\right\}.$$

Now since A is an atom one of them has measure equal to $\mu(A)$ and the other zero measure. Let us assume for simplicity that I_1 has measure equal to $\mu(A)$. Then again we split it in half and obtain two subsets of I_1

$$I_{11} = \left\{ f|_A \in \left[0, \frac{\|f|_A\|_{L^\infty}}{4}\right]\right\} \quad \text{and} \quad I_{12} = \left[f|_A \in \left\{\frac{\|f\|_A|_{L^\infty}}{4}, \frac{\|f\|_A||_{L^\infty}}{2}\right]\right\}.$$

Again now one has measure equal to $\mu(A)$ and the other zero. We select the one with the positive measure and repeat the same process.

Considering the intersection of the sets we elected in each step gives us a set of the form $\{f = a\}$ where $0 \leq a < \infty$ and since A is an atom by using elementary properties of a measure we obtain that its measure is equal to $\mu(A)$.

Generally if f takes also negative values, we just proved that $|f|$ is a.e. constant on A and so if a is its a.e. constant then we must have either $|\{f|_A = a\}| = \mu(A)$ or $|\{f|_A = -a\}| = \mu(A)$ and by that the other has zero. $\square$

Remark 1.19. The same result doesn't hold true when the atom has infinite measure. To prove this we consider the σ-algebra $S = \{A, B, C, \emptyset\}$ where $A = B \cup C$ and $B \cap C = \emptyset$ with all of them except of course the null set having infinite measure. That gives us that A is an atom, but we may consider measurable functions that take different values on B and C.

Theorem 1.20. *If (X, μ) is a non-atomic measure space then for every nonnegative number a such that $a < \mu(X)$ there exists a measurable subset B of X such that $\mu(B) = t$.*

For the proof we are going to need the following lemma

Lemma 1.21. *Let (X, μ) be a finite measure space. Then for every $\epsilon > 0$ there exists a finite partition of X of measurable sets $X_1, \ldots, X_n$ such that either $\mu(X_i) \leq \epsilon$ or X_i is an atom such that $\mu(X_i) > \epsilon$.*

Proof. Since the measure space is finite there are finitely many nonequivalent atoms $A_1, \dots, A_k$ of measure greater than ϵ. And so the space $Y \setminus \cup_{i=1}^k A_i$ has no atoms of measure greater than ϵ.

Now let us show that every B measurable, $B \subset Y$ having positive measure contains a set C such that $0 < \mu(C) \leq \epsilon$.

If this is not true then we must have that $\mu(B) > \epsilon$, for else the result holds for $C = B$ and so B can't be an atom. So there exists a measurable set B_1, with $B_1 \subset B$ such that

$$\epsilon < \mu(B_1) < \mu(B).$$

Now for the set $B \setminus B_1$ we may find set $B_2 \subset B \setminus B_1$ such that

$$\epsilon < \mu(B_2) < \mu(B \setminus B_1)$$

and again we can find a set $B_3 \subset B \setminus (B_1 \cup B_2)$ such that

$$\epsilon < \mu(B_3) < \mu(B \setminus \bigcup_{i=1}^{3} B_i).$$

Repeating the same process gives us a sequence of pairwise disjoint sets of measure greater than $\epsilon > 0$, but that contradicts the fact that X is finite.

Now if Y has measure less or equal to ϵ then

$$\{A_1, \dots, A_k, Y\}$$

is the desired partition.

If $\mu(Y) > \epsilon$ let us set for a measurable A such that $A \subset Y$

$$\eta(A) := \sup\{\mu(B) : B \subset A, \mu(B) \leq \epsilon\}.$$

We just proved that $0 < \eta(A) \leq \epsilon$ provided that $\mu(A) > 0$. With that in mind we can find a sequence (which can be finite) of pairwise disjoint measurable sets B_i such that $B_i \subset Y$ and

$$\frac{1}{2}\eta\left(Y \setminus \bigcup_{i=1}^{n} B_i\right) \leq \mu(B_{n+1}) \leq \epsilon.$$

If at some step m we can't continue the procedure i.e. the remaining set has measure equal to zero we adjoin that remaining zero measure set to B_{m-1} so that their union is the space itself. So the desired partition is the following

$$\{A_1, \dots, A_k, B_1, \dots, B_{m-1}\}.$$

If not we set $B_0 := Y \setminus \cup_{i=1}^{\infty} B_i$. Then since

$$\eta(B_0) \leq \eta\left(Y \setminus \bigcup_{i=1}^{n} B_i\right) \leq 2\mu(B_{n+1})$$

and $\sum_{i=1}^{\infty} \mu(B_i) \leq \mu(Y) < \infty$ we have that $\lim_{n\to\infty} \mu(B_n) = 0$ so $\eta(B_0) = 0$ and by the property that is proved above for Y

$$\mu(B_0) = 0.$$

Considering a number l large enough so that $\sum_{i=l}^{\infty} \mu(B_i) < \epsilon$ gives us that

$$\{A_1, \dots, A_k, B_1, \dots, B_{l-1}, B_l, \bigcup_{i=l+1}^{\infty} B_i, B_0\}$$

is a partition with the desired properties. □

Proof of Theorem. Using Lemma 1.21 we proved just above we are going to construct a increasing sequence of measurable subsets of A, $\{A_n\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.

First we partition X into finitely many measurable sets $\{X_1, \dots, X_m\}$ such that $\mu(X_i) \leq \frac{1}{2}$ and consider the greatest number l such that $\mu(\bigcup_{i=1}^{l} X_i) \leq a$ and set

$$A_1 := \bigcup_{i=1}^{l} X_i.$$

If $\mu(A_1) = a$ then we are done. If not let us first notice that

$$\mu(A_1) \geq a - \frac{1}{2}. \tag{1.2}$$

Inequality (1.2) holds for else we could adjoint another X_i to A_1. Now using Lemma 1.21 we find a measurable finite partition of $X \setminus A_1$, $\{\tilde{X}_1, \dots, \tilde{X}_k\}$ such that $\mu(\tilde{X}_i) \leq \frac{1}{3}$ for $i = 1, 2, \dots, k$ and pick the greatest j such that

$$\mu\left(\bigcup_{i=1}^{j} \tilde{X}_j\right) \leq a - \mu(A_1). \tag{1.3}$$

We now set $B_1 = \cup_{i=1}^{j} \tilde{X}_i$ and $A_2 = A_1 \cup B_1$.s If $\mu(A_2) = a$ then we are finished. If not we similarly have

$$\mu(B_1) \geq a - \mu(A_1) - \frac{1}{3}$$

and proceed accordingly.

If at some step m the procedure stops that means that $\mu(A_m) = a$ and we are finished. If not then during this procedure we elected an increasing sequence of measurable sets A_i such that

$$a \geq \mu(A_i) \geq a - \frac{1}{i+1}$$

and so the set $A := \cup_{i=1}^{\infty} A_i$ is the desired one. □

Corollary 1.22. *Let (X,μ) be a finite non-atomic measure space. Then for every t such that $(0 < t < \mu(X))$ there exists a set E_t of measure t such that for every $t, s \in (0,1)$ with $t \leq s$, $E_t \subset E_s$.*

Proof. We may assume without loss of generality that $\mu(X) = 1$ and let t such that $0 < t < 1$. We will first construct a set E_t such that $\mu(E_t) = t$.

1. First using Theorem 1.20 we partition X into two subsets A_1^1 and A_2^1 such that
$$\mu(A_1^1) = \mu(A_2^1) = \frac{1}{2}.$$
If now $t \geq 1/2$ then we set $E_1 := A_2^1$ and proceed to the next step. *Otherwise* we set $E_1 := \emptyset$ and again proceed to the next step.

2. For the next step we first partition the set A_1^1 and A_2^1 as follows
$$A_1^1 := A_1^2 \cup A_2^2 \quad \text{and} \quad A_2^1 := A_3^2 \cup A_4^2$$
with $\mu(A_i^2) = 1/4$ for $i = 1,2,3,4$.
For the case where $t \geq 1/2$.
If $t \geq 3/4$ we set $E_2 := E_1 \cup A_3^2$ and proceed to the next step. otherwise i.e. if $t \in [1/2, 3/4)$ we set $E_2 := E_1$ and again proceed to the next step.
As for the case where $t < 1/2$.
If $t \in [1/4, 1/2)$ then we set $E_2 := A_1^2$ and proceed to the next step otherwise we set $E_2 := \emptyset$ and also proceed to the next step.

3. Suppose now we have repeated the process until the m step where $m \geq 2$. From that point we partition X as follows
$$A_1^{m+1}, A_2^{m+1}, \ldots, A_{2^{m+1}}^{m+1}$$
so that
$$A_1^m = A_1^{m+1} \cup A_2^{m+1}, \quad A_2^m = A_3^{m+1} \cup A_4^{m+1}, \ldots, A_{2^m}^m = A_{2^{m+1}-1}^{m+1} \cup A_{2^{m+1}}^{m+1}$$
and $\mu\left(A_i^{m+1}\right) = 1/2^{m+1}$ for $i = 1, \ldots, 2^{m+1}$. Now let $k = 1, 2, \ldots, 2^{m+1}$ be the unique natural number such that $t \in [(k-1)/2^m, k/2^m)$. Consequently $E_m = \cup_{i=1}^{k-1} A_i^m$ and we turn our attention to the set A_k^m for which there exists an i such that
$$A_k^m = A_i^{m+1} \cup A_{i+1}^{m+1}.$$
If t is greater or equal to the middle point of $[(k-1)/2^m, k/2^m)$ then we set $E_{m+1} := E_m \cup A_i^{m+1}$ otherwise we set $E_{m+1} := E_m$ and proceed accordingly.

Now by setting $E_t := \cup_{i=1}^{\infty} E_i$ we have $\mu(E_t) = t$ since the sequence $\{E_i\}_{i=1}^{\infty}$ is increasing and

$$|\mu(E_i) - t| \leq \frac{1}{2^i}, \quad \text{for every } i \in \mathbb{N}.$$

As for the second property to hold we just have to use the same sets A_j^i in each step *for every* $t \in (0,1)$. □

Definition 1.23. A σ-finite measure space (X, μ) will be called *resonant* if for every pair of measurable functions $f, g : X \to \mathbb{R}$

$$\sup\left\{\int_X fg'\,d\mu : g' \text{ equimeasurable with } g\right\} = \int_0^{\mu(X)} f^* g^*\,dx.$$

In particular if for every pair f, g there is in fact g' equimeasurable to g so that equality occurs i.e.

$$\int_X fg'\,d\mu = \int_0^{\mu(X)} f^* g^*\,dx$$

then the space will be called *strongly resonant*.

Theorem 1.24. *A σ-finite measure space (X, μ) is strongly resonant if and only if it is of finite measure and satisfies one of the conditions below*

- *It is completely atomic with its atoms having the same measure.*
- *It is non-atomic.*

For the proof of this theorem we will use the following lemma.

Lemma 1.25. *Let (X, μ) be a finite nonatomic measure space and f an a.e. finite function on X. Then for every real number t such that $0 \leq t \leq \mu(R)$ there exist measurable subsets of X, E_t such that*

$$\int_{E_t} |f|\,d\mu = \int_0^t f^*(s)\,ds.$$

Moreover E_t, $0 \leq t \leq \mu(X)$ can be chosen so that

$$E_s \subset E_t, \quad \text{if } s \leq t.$$

Proof. In order to prove that we will study separately the cases where the number t is a value of the distribution function of f and isn't.

Let us first consider the case where t is a value of the distribution function. Then we do have the following for its decreasing rearrangement,

$$f^*(t) = \inf\{\lambda > 0 : \mu_f(\lambda) = t\}$$

and so we have that $\mu_f(f^*(t)) = t$. So if we set $E_t := \{x : |f(x)| > f^*(t)\}$, we have equivalently $\mu(E_t) = t$.

So now if we consider the restrictions f_t, of f to E_t we have for its distribution function which we denote μ_{f_t} that

$$\begin{aligned}\mu_{f_t}(\lambda) = \mu(\{x \in E_t : |f(x)| > \lambda\}) &= \mu(\{|f| > f^*(t)\} \cap \{|f| > \lambda\}) \\ &= \mu(\{|f| > \max(f^*(t), \lambda)\}).\end{aligned}$$

Consequently

$$\mu_{f_t}(\lambda) = \begin{cases} t, & \text{if } \lambda \leq f^*(t) \\ \mu_f(\lambda), & \text{if } \lambda > f^*(t). \end{cases}$$

We also have for the distribution function of f^* restricted to $[0, t]$

$$|\{x \in [0,t] : f^*(x) > \lambda\}| = \begin{cases} t, & \text{if } \lambda \leq f^*(t) \\ \mu_{f^*}(\lambda), & \text{if } \lambda > f^*(t) \end{cases}$$

and since f and its decreasing rearrangement are equimeasurable we have that f_t and f^* are also equimeasurable. As a result

$$\int_{E_t} |f| \, d\mu = \int_0^t f^* \, dx$$

in case t is value of the distribution function.

Moreover if s is another value of μ_f with $s > t$ then we have that $f^*(s) \leq f^*(t)$ and by extension

$$E_s = \{x \in X : |f(x)| > f^*(s)\} \supset \{x \in X : |f(x)| > f^*(t)\} = E_t.$$

Now we treat the case where t is not a value of the distribution function of μ_f. Let us first notice that since X is of finite measure and f is an a.e. finite function

$$\lim_{n\to\infty} \mu_f(n) = \lim_{n\to\infty} \mu(\{|f| > n\}) = \mu(\{|f| = \infty\}) = 0$$

and as a result

$$\lim_{\lambda\to\infty} \mu_f(\lambda) = 0$$

since μ_f is decreasing. Consequently $f^*(t) = \inf\{\lambda > 0 : \mu_f(\lambda) \leq t\} < \infty$.

We now set $\lambda_0 := f^*(t)$ and examine separately the cases where λ_0 is equal to zero or not but *before we move on let us to fix for each set $G_{\lambda_0} := \{|f| = \lambda_0\}$ where λ_0 is a value of f^*, a family of sets with the same properties as Corollary 1.22 by applying it to each of the G_{λ_0}.*

Let us first consider the case where $\lambda_0 > 0$. Then since μ_f is decreasing

$$t_0 := \mu_f(\lambda_0) < t \leq \mu_f(\lambda_0-) := t_1$$

where the last one denotes the left-side limit of μ_f at λ_0. Then we obtain

$$f^*(s) = \lambda_0, \quad \text{if } s \in [t_0, t_1).$$

Now let us prove that

$$t_1 = \mu\{x : |f(x)| \geq \lambda_0\}. \tag{1.4}$$

If we set

$$F_n := \left\{ |f| > \lambda_0 - \frac{1}{n} \right\}$$

for every $n \in \mathbb{N}$ then

$$\{|f| \geq \lambda_0\} = \bigcup_{n=1}^{\infty} F_n$$

and since X is finite we obtain

$$\mu(\{|f| \geq \lambda_0\}) = \lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} \mu_f \left(\lambda_0 - \frac{1}{n} \right) = \mu_f(\lambda_0^-) = t_1.$$

Now expressing $G_{\lambda_0} = \{|f| = \lambda_0\}$ as follows

$$G := \{|f| \geq \lambda_0\} \setminus \{|f| > \lambda_0\}$$

and combining (1.4) with the definition of t_0 gives us

$$\mu(G) = t_1 - t_0.$$

We now select a set F_t of G_{λ_0} of measure equal to $t - t_0$ of the family of subsets we fixed at the beginning and set

$$E_t := \{|f| > \lambda_0\} \cup F_t.$$

We then have that $\mu(E_t) = \mu_f(\lambda_0) + (t - t_0) = t$.

Moreover

$$\begin{aligned} \int_{E_t} |f| \, d\mu &= \int_{\{|f|>\lambda_0\}} |f| \, d\mu + \int_{F_t} |f| \, d\mu \\ &= \int_0^{t_0} f^*(s) \, ds + \lambda_0(t - t_0) \end{aligned}$$

and since f^* is equal to λ_0 on (t_0, t) the right side is equal to $\int_0^t f^*(s) \, ds$.

As for the case when $f^*(t) = \lambda_0 = 0$ again from the fact that μ_f is decreasing we have

$$\mu(\{|f| > 0\}) := t_0 < t.$$

In this case we select a set F_t of the initial fixed family of $G_0 = \{f = 0\}$ having measure equal to $t - t_0$ and set likewise

$$E_t := \{|f| > 0\} \cup F_t.$$

which then gives us

$$\int_{E_t} f \, d\mu = \int_{\{|f|>0\}} = \int_0^{t_0} f^*(s) ds.$$

As for the decreasing property of all the sets we elected it is a straightforward consequence of Corollary 1.22 which we applied during their construction. □

Proof of theorem. First let us prove that X has to comply with one of the two conditions above in order to be strongly resonant.

Let us assume that X has two atoms a,b of different measure and without loss of generality let us assume that $\mu(b) < \mu(a)$. Then if we consider their respective characteristic functions χ_a and χ_b set them in place of f and g in Hardy's Inequality one has

$$0 = \int_X fg \, d\mu \leq \mu(b)$$

where $f = \chi_a$ and $g = \chi_b$.
If we consider a consider another function g' that is equimeasurable with g then g' would be a.e. equal to a characteristic function of a set B of measure $\mu(b)$ so if we set $N := \{f \neq 0\} \cap \{g' \neq 0\}$ then

$$N \preccurlyeq b,$$

by extension $\mu(N) = \mu(b)$ or $\mu(N) = 0$. But N can't have measure equal to $\mu(B)$ since that contradicts the atomic structure of a, so $\mu(N) = 0$. As a result the integral $\int_X fg' \, d\mu$ always equal to zero when g' is equimeasurable to g.

Now for the necessity of the finiteness. Let us assume to the contrary that X has infinite measure. Then X can be written as follows

$$X = \bigcup_{i \in I} X_i$$

where X_i are pairwise disjoint sets of equal and finite measure.

Indeed if X is purely atomic then X can be expressed directly in this form due the fact that all its atoms must have the same finite measure. If X is nonatomic we use Theorem 1.20 for this decomposition.

Now considering a function f as follows

$$f := \sum_{n \in \mathbb{N}} \left(1 - \frac{1}{n}\right) \chi_{X_n}.$$

gives us that $f^* = 1$.

Now let g to be the characteristic of set of measure c. If g' is another function equimeasurable to g then $g^{'}$ is equal a.e. to the characteristic function of a set of measure c.

Hardy's Inequality gives us

$$\int_X fg' \, d\mu \leq \int_0^\infty f^* g^* \, dt.$$

The right side of the inequality is equal to c, as opposed to the left side which is always less than c.

And now to prove the sufficiency of the conditions

1. *Let us concentrate on the sufficiency of the first condition first.* If it is completely atomic with its atoms having equal measure then it has only a finite number of distinct atoms $\{A_i\}_{i=1}^n$ since the space is of finite measure. Since every measurable function is a.e. constant on atoms of finite measure

$$f = \sum_{i=1}^{n} f_i \chi_{A_i} \text{ a.e.} \quad \text{and} \quad g = \sum_{i=1}^{n} g_i \chi_{A_i} \text{ a.e.}$$

and as a result

$$\int_X fg \, d\mu = \sum_{i=1}^{n} f_i g_i \mu(A_i).$$

Now a straightforward calculation of the right side of Hardy's Inequality shows us that

$$\sum_{i=1}^{n} f_i g_i \mu(A_i) \leq \sum_{i=1}^{n} f_i^* g_i^* \mu(A_i)$$

where $f_1^* \geq \cdots \geq f_n^*$ and $g_1^* \geq \cdots \geq g_n^*$. This is practically the case of Hardy's Inequality for finite sequences and so we just need to consider an appropriate arrangement of g.

2. *And now for the sufficiency of the second condition.* To prove that let us first remember that we can find an increasing sequence of simple functions $\{g_n\}_{n\in\mathbb{N}}$ that converges pointwise to g.

By having such a sequence in our hands we are going to find an increasing sequence $\{g_n'\}_{n\in\mathbb{N}}$ equimeasurable per term with $\{g_n\}_{n\in\mathbb{N}}$ and such that

$$\int_X |f||g_n'| \, d\mu \ = \int_0^\infty f^* {g_n}^* \, dx$$

for every $n \in \mathbb{N}$.

Then by considering the point set-limit function g' of the sequence $\{g_n'\}_{n\in\mathbb{N}}$ we have that it is equimeasurable with g and so a simple application of the monotone convergence theorem gives us the result we want.

Now to find such a sequence $\{g_n'\}_{n\in\mathbb{N}}$ let us first consider a term of the sequence $\{g_n\}_{n\in\mathbb{N}}$ and let us set $h := g_n$ in order to simplify our notation and express h as follows

$$h = \sum_{i=1}^{m} b_i \chi_{B_i}$$

where $b_i > 0$ for $i = 1, \ldots, m$ and $B_1 \subset \cdots \subset B_m$.

From Lemma 1.25 we have previously proven there exist sets $E_1, \ldots, E_n$ such that $\mu(E_i) = \mu(B_i)$ and such as

$$\int_{E_i} f \, d\mu = \int_0^{\mu(B_i)} f^*(x) \, dx.$$

So by setting

$$h' := \sum_{i=1}^{m} b_i \chi_{E_i}$$

we have that h and h' are equimeasurable since

$$h^* = \sum_{i=1}^{m} b_i \chi_{[0,\mu(B_i))} = \sum_{i=1}^{m} b_i \chi_{[0,\mu(E_i))} = (h')^*.$$

Consequently

$$\begin{aligned}\int_X fh' \, d\mu = \sum_{i=1}^{n} b_i \int_{E_i} f \, d\mu &= \sum_{i=1}^{m} b_i \int_0^{\mu(B_i)} f^* \, dx \\ &= \int_0^\infty \sum_{i=1}^{m} b_i \chi_{[0,\mu(B_i)} f^* \, dx \\ &= \int_0^\infty f^* h^* \, dx.\end{aligned}$$

□

Theorem 1.26. *A σ-finite measure space (X, μ) is resonant if and only if it satisfies one the conditions below.*

1. *It is non-atomic.*
2. *It is completely atomic with its atoms having equal measure.*

Proof. The first part of the previous proof proves the necessity of these conditions.

So let us concentrate on the sufficiency of these conditions. If X is of finite measure we have already proved it above so let us assume that the space is of infinite measure.

Let f, g be finite a.e. functions and let's also assume that they are non-negative. If the integral on the right hand side of Hardy's Inequality is equal to zero then equality surely holds so let us assume it is not.

It is enough to show that for every positive number a such that

$$a < \int_0^\infty f^* g^* \, dx$$

there exists a function g' equimeasurable with g such that

$$a < \int_X fg' \, d\mu.$$

Since X is σ-finite X it can be written as $X = \cup_{i \in I} X_i$ where X_i is an increasing sequence such that $\mu(X_i) < \infty$ and by extension we may find an increasing sequences $\{f_n\}_{n\in\mathbb{N}}, \{g_n\}_{n\in\mathbb{N}}$ of simple functions that converge pointwise to f and g respectively such that the support[3] of the i term of each one lies within X_i and so from the Dominated Convergence Theorem

$$a \leq \int_0^\infty f^*_{n_0} g^*_{n_0}\, dx$$

for some $n_0 \in \mathbb{N}$.

Since the space X_{n_0} is of finite measure, Theorem 1.24 ensures us that it is strongly resonant. As a result there exists a function h defined on X_{n_0} equimeasurable to g_{n_0} such that

$$\int_{X_{n_0}} fh\, d\mu = \int_0^{\mu(X_{n_0})} f^*(g_{n_0})^*\, dx$$

and so

$$a \leq \int_X fh\, d\mu.$$

Now if we set $g' := h\chi_{X_{n_0}} + g\chi_{X\setminus X_{n_0}}$, due to the fact that $g' \geq h$ and g' is equimeasurable with g gives us

$$a \leq \int_X fh\, d\mu \leq \int_X fg'\, d\mu. \qquad \square$$

1.4 The elementary maximal operator f^{**}

Consider Hardy's Inequality for an initial fixed function f. By setting g to be equal to the characteristic function of a set E of measure equal to t we have

$$\frac{1}{\mu(E)} \int_E |f|\, d\mu \leq \frac{1}{t} \int_0^t f^*(s)\, ds.$$

So the quantity to the right above constitutes an upper bound for all the averages of f over sets of measure t and that leads us to the definition we present below.

Definition 1.27. Let (X, μ) be a σ-finite measure space and f be a measurable a.e. finite function. We define *the maximal operator of* f^* denoted by f^{**} as follows

$$f^{**}(t) = \frac{1}{t} \int_0^t f^*(s)\, ds.$$

Remark 1.28. By using the inequality above or the fact that since f^* is decreasing so is $\frac{1}{t}\int_0^t f^*(s)ds$ we reach the fact that the average f^{**} is the biggest of the averages of f^* over sets of measure t and that leads us to use the term maximal in Definition 1.27.

Theorem 1.29 (Basic properties of f^{**})**.** *Let (X, μ) be a σ-finite measure space and and f, g be measurable a.e. finite functions defined on X, then*

1. $f^{**} = 0 \iff f = 0$ *a.e..*

[3] the set of points that differ from zero

2. $f^* \leq f^{**}$.

3. If $|f| \leq |g|$ *a.e. then* $f^{**} \leq g^{**}$.

4. For every $a \in \mathbb{R}$, $(af)^{**} = |a| f^{**}$.

Proof. All of these proofs stem from the basic properties of f^* (Theorem 1.11) we proved before.

1. If $f = 0$ then it is straightforward that $f^{**} = 0$. Now if $f^{**} = 0$ then that means that $f^* = 0$ a.e. and since f^* is decreasing if there was a $t > 0$ such that $f^*(t) > 0$ then we would have that $f^{**} \geq f^*(s)$ for $s \in [0, t]$ which is impossible, so for every $t > 0$ $f^*(t) = 0$. Now from the right continuity of f^* the same holds true for $t = 0$ and so $f = 0$ a.e..

2. From the definition of f^{**} and since f^* is decreasing
$$f^{**}(t) = \frac{1}{t} \int_0^t f^*(s)\, ds \geq \frac{1}{t} t f^*(t) = f^*(t).$$

3. We previously proved that if $f \leq g$ a.e. then $f^* \leq g^*$. So
$$f^{**}(t) = \frac{1}{t} \int_0^t f^*(s)\, ds \leq \frac{1}{t} \int_0^t g^*(s)\, ds = g^{**}(s).$$

4. For every $a \in \mathbb{R}$ we have $(af)^* = |a| f^*$ and consequently
$$(af)^{**}(t) = \frac{1}{t} \int_0^t (af)^*(s)\, ds = \frac{1}{t} \int_0^t |a| f^*(s)\, ds = |a| f^{**}(t).$$

□

Theorem 1.30. *Let* (X, μ) *be a* σ*-finite measure space and let* t *be a value of* μ *and let* f *be an a.e. finite measurable function. If* (X, μ) *is a resonant space then*
$$f^{**}(t) = \frac{1}{t} \sup \left\{ \int_A |f|\, d\mu : \mu(A) = t \right\}.$$

If it is strongly resonant then the exists a set A *with measure* t *such that*
$$f^{**}(t) = \frac{1}{t} \int_A |f|\, d\mu.$$

Proof. Since t is a value of μ there exists a set B such that $\mu(B) = t$. Setting $g := \chi_B$ and taking into consideration the fact that every g' equimeasurable to g is a.e. equal to the characteristic function of a set of measure t we have,
$$f^{**}(t) = \frac{1}{t} \sup \left\{ \int_X |f| \chi_B\, d\mu : \mu(B) = t \right\} = \frac{1}{t} \sup \left\{ \int_B f\, d\mu : \mu(A) = t \right\}.$$

provided that X is resonant.

Now if the space is strongly resonant that would mean that there is a g' equimeasurable to g and by extension be a.e. equal to a characteristic of a set A of measure t for which the respective measure is attained. Consequently

$$f^{**}(t) = \frac{1}{t}\int_X |f| g \, d\mu = \frac{1}{t}\int_A |f| \, d\mu. \qquad \square$$

By the theorem above we reach the following sub-additive property for f^{**} whenever (X, μ) is a resonant space and t is value of μ since

$$\begin{aligned}(f+g)^{**}(t) &= \frac{1}{t}\sup\left\{\int_A |f+g| \, d\mu : \mu(A) = t\right\} \\ &\leq \frac{1}{t}\sup\left\{\int_A |f| \, d\mu : \mu(A) = t\right\} + \frac{1}{t}\sup\left\{\int_A |g| \, d\mu : \mu(A) = t\right\} \\ &= f^{**}(t) + g^{**}(t).\end{aligned}$$

Now let us also suppose that X is nonatomic and at first let us suppose that X has infinite measure. Then μ takes all the values of $[0, \infty]$ and consequently the subadditivity holds for every $t \geq 0$.

If X is of finite measure with $\mu(X) := t_0 < \infty$, then for $t > t_0$ we have that $f^*(t) = g^*(t) = 0$ and so

$$f^{**}(t) = \frac{1}{t}\int_0^{t_0} f^*(s) \, ds \quad \text{and} \quad g^{**}(t) = \frac{1}{t}\int_0^{t_0} g^*(s) \, ds, \quad \text{for } t \geq t_0.$$

Equivalently

$$f^{**}(t) = \frac{t_0}{t} f^{**}(t_0) \quad \text{and} \quad g^{**}(t) = \frac{t_0}{t} g^{**}(t_0), \quad \text{for } t \geq t_0$$

and so since the subadditivity holds for t_0 it also holds for $t > t_0$.

As for the more general case we will describe here a procedure that will allow us to prove the subadditivity for every measure space, called *method of retracts*. What we are going to do is embed our arbitrary σ-finite measure space into a nonatomic measure space which as we have just shown has that property.

To do that let analyze our fixed random σ-finite measure space X as follows

$$X = X_0 \cup \bigcup_{j \in J} X_j$$

where the X_i are distinct atoms of X and X_0 is the nonatomic part of X. We can express X this way since X is σ-finite and so it has at most countable disjoint atoms.

Now let us define a new space $\overline{X}$ as follows

$$\overline{X} := X_0 \cup \bigcup_{i \in J} I_i$$

where I_i are disjoint intervals of $\mathbb{R}$ such that $|I_i| = \mu(X_i)$. If X_0 happens to be partially an interval of $\mathbb{R}$ then we just have to choose intervals disjoint from X_0.

So in order for $\overline{X}$ to be a measure space we just need to define a σ-algebra on $\overline{X}$ and a measure.

A subset of $\overline{X}$ will be called measurable if and only if its intersection with X_0 is a measurable subset of X and its intersection with each of the intervals I_i is Lebesgue measurable. The collection of these sets is indeed σ-algebra on $\overline{X}$.

Now let's define the measure on this σ-algebra by a natural way. If A is measurable subset of $\overline{X}$ we define its respective measure which we will denote by $\overline{\mu}(A)$ as follows

$$\overline{\mu}(A) := \mu(A \cap X_0) + \sum_{j \in J} |A \cap I_j|$$

and so the the measure space $(\overline{X}, \overline{\mu})$ is σ-finite non-atomic measure space since every measurable set A can be partitioned into two measurable subsets the part that lies within X_0 and the part that lies within the union of the intervals we defined above and both of them are nonatomic.

Now consider a measurable a.e. finite function f defined X. We have proved that f a.e. constant on atoms of finite measure and with that in mind we define a new function let us call it $\overline{f}$ defined on $\overline{X}$ by the following way.

We set $\overline{f}$ equal f on X_0 and in each of the disjoint intervals of X we set it to be equal to the a.e. constant value that f takes in the corresponding atoms of X.

These two functions are equimeasurable and so

$$(\overline{f})^* = f^*.$$

With the preparation above we are now ready to prove the subadditive property for a random σ-finite measure space X.

Theorem 1.31 (Subadditivity of f^{**})**.** *Let X be a σ-finite measure space and f, g be finite a.e. measurable functions defined on X then*

$$(f+g)^{**} \leq f^{**} + g^{**}.$$

Proof. Suppose that f, g are measurable a.e. finite functions. Then if we express X as

$$X = X_0 \cup \bigcup_{i \in I} X_i$$

as before, then in each of the atoms X_j, $f = f_j$ a.e. and $g = g_j$ a.e., where $f_j, g_j \in \mathbb{R}$. As a result $f + g = f_j + g_j$ a.e. and so $\overline{f+g}$ is equal to $f_j + g_j$ on the corresponding intervals I_j and equal to $f + g$ on X_0. Consequently $\overline{f+g} = \overline{f} + \overline{g}$.

Now using that gives us

$$(f+g)^{**} = (\overline{f+g})^{**} = (\overline{f} + \overline{g})^{**} \leq (\overline{f})^{**} + (\overline{g})^{**} = f^{**} + g^{**}. \qquad \square$$

CHAPTER 2

Maximal Operators

2.1 Hardy-Littlewood Maximal Operator

Consider a locally integrable function defined on $\mathbb{R}^n$ and for each point in $\mathbb{R}^n$ consider the integral average of the function over a cube (or a ball) that contains it. The Hardy-Littlewood Maximal Operator of that function at any given point is just the supremum of the averages over all the cubes (or balls) that contain the point. Before we use the theory of the decreasing rearrangements that we have already developed we will study some basic properties of this maximal operator in a stripped way and then proceed with it. The source of context for the first three sections is from the book by Duoandikoetxea "Fourier Analysis" [9].

Definition 2.1. The *Hardy-Littlewood Maximal Operator* of a function $f \in L^1_{loc}(\mathbb{R}^n)$ is defined as follows

$$\mathcal{M}f(x) := \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f|, \quad \text{for every } x \in \mathbb{R}^n$$

where the supremum is taken over all cubes with sides parallel to the coordinate axes that contain x.

Note 2.2. The 1-dimensional case of the above operator was introduced by Hardy and Littlewood [18] and its generalization in greater dimensions by Wiener [40].

Remark 2.3. If we replace the cubes with balls in the definition above we obtain another kind of maximal operator. But both of them are point-wise equivalent for instance if $\mathcal{M}_q$ is the Hardy-Littlewood maximal operator defined above and $\mathcal{M}_b$ is the one that stems if we replace in the definition the cubes with balls then

$$\mathcal{M}_b f \leq \frac{1}{v_n} \mathcal{M}_q f \leq \frac{(\sqrt{n})^n}{v_n} \mathcal{M}_b f$$

where v_n stands for the volume of the unit ball of $\mathbb{R}^n$, since every ball of radius $r > 0$ is contained inside a cube with the same center and length $2r$ and every such cube is contained in a ball again with the same center and radius $\sqrt{n}r$.

Similar results hold even if we restrict ourselves to the cubes (or the balls) centered about x. In the context of the theory we are studying it makes no difference which one of them we pick unless otherwise stated. *It is a matter of convenience.*

Our first observation is $\mathcal{M}f$ is not integrable unless f is a a.e. zero function.

Theorem 2.4. *Let $f \in L^1(\mathbb{R}^n)$ such that $f \neq 0$ at a set of positive measure. Then there are constants c and r s.th.*

$$\mathcal{M}f(x) \geq \frac{c}{|x|^n}, \quad \text{for } |x| \geq r.$$

Proof. For convenience we will be using the ball variant. From the hypothesis we may find a ball B_r of radius $r > 0$ centered at 0 such that f is not a.e. equal to zero on B_r. For $|x| \geq r$ we have then that $x \in B(0, 2|x|)$ and $B(0, 2|x|) \supset B_r$. Consequently

$$\mathcal{M}f(x) \geq \frac{\int_{B(0,2|x|)} f(y)\, dy}{|B(0, 2|x|)|} \geq \frac{\int_{B_r} f(y)\, dy}{2^n |x|^n} := \frac{C}{|x|^n}. \qquad \square$$

As for other spaces we will be prove that the Hardy-Littlewood Maximal Operator maps L^p to L^p when $p > 1$ using an interpolation theorem. Before that we will define a substitute of the L^p-boundedness, that for instance the Hardy-Littlewood Maximal Operator doesn't have for $p = 1$ as we have just proved, the weak-L^p condition.

More precisely if (X, μ) and(Y, ν) are measure spaces and $1 \leq p < \infty$, we say that an operator T from $L^p(X)$ to the space of the measurable functions defined on Y is *weak-L^p* if there is a constant C_p such that

$$\nu(\{Tf > \lambda\}) \leq \left(C_p \frac{\|f\|_{L^p}}{\lambda} \right)^p, \quad \text{for every } f \in L^p(X).$$

Note that the strong-L^p boundedness implies the weak-L^p boundedness for if we set $E_\lambda := \{Tf > \lambda\}$, then

$$\nu(E_\lambda) = \int_{E_\lambda} 1\, d\nu \leq \int_{E_\lambda} \frac{|Tf|^p}{\lambda^p}\, d\nu \leq \frac{\|Tf\|^p_{L_p}}{\lambda^p} \leq c \frac{\|f\|^p_{L_p}}{\lambda^p}.$$

In case $p = \infty$ we identify the weak-L^∞ with the strong L^∞, the reason for it will become apparent later on. We are now ready to state and prove the Marcinkiewicz Interpolation Theorem.

Theorem 2.5 (Marcinkiewicz Interpolation Theorem)**.** *Let X, Y be two σ-finite measure spaces and $p_1, p_2 \in [1, \infty]$ with $p_1 < p_2$ and let T be a sub-linear operator from $L^{p_1}(X) + L^{p_2}(X)$ to the space of measurable functions defined on Y, that is weak-L^{p_1} and weak-L^{p_2}. Then it is strong-L^p for every p such that $p_1 < p < p_2$.*

Proof. First let us consider a function $f \in L^p(X)$ and for a fixed $\lambda > 0$ consider the following "decomposition" of f with respect to λ,

$$f^1_\lambda := f\chi_{\{|f| > c\lambda\}};$$
$$f^2_\lambda := f\chi_{\{|f| \leq c\lambda\}},$$

where c is a constant which we will define later. Then we have that $f_\lambda^1 \in L^{p_1}$ and $f_\lambda^2 \in L^{p_2}$ since,

$$\|f_\lambda^1\|_{L^{p_1}}^{p_1} = \int_X |f_\lambda^1|^{p_1}\, d\mu = \int_{\{|f|\geq c\lambda\}} |f|^p |f|^{p-p_1}\, d\mu \leq (c\lambda)^{p-p_1} \int_X |f|^p\, d\mu < \infty$$
$$\|f_\lambda^2\|_{L^{p_2}}^{p_2} = \int_X |f_\lambda^2|^{p_2}\, d\mu = \int_X |f_\lambda^2|^p |f_\lambda^2|^{p_2-p}\, d\mu \leq (c\lambda)^{p_2-p} \int_X |f|^p\, d\mu < \infty$$

provided that $p_2 < \infty$. As for the case where $p_2 = \infty$ the result stems immediately from the fact if f is L^∞ so are its restrictions.

Now due to the fact that our operator is weak-L^{p_1} and $p_1 < p_2$ gives us $p_1 < \infty$. Consequently

$$\mu_{Tf_\lambda^1}\left(\frac{\lambda}{2}\right) \leq \left(\frac{2C_{p_1}}{\lambda}\right)^{p_1} \int_X |f_\lambda^1|^{p_1}\, d\mu.$$

The case where $p_2 = \infty$ and the case where $p_2 < \infty$ will be examined separately.

Let us first examine the case where $p_2 = \infty$. In this case we have that

$$\|Tf\|_\infty \leq C_\infty \|f\|_\infty, \quad \text{for every } f \in L^p.$$

So setting $c := 1/2C_\infty$ gives us $\mu_{Tf_\lambda^2}(\lambda/2) = \mu(Tf_\lambda^2 > \lambda/2) = 0$.

Let us now consider a $p \in (p_1, \infty)$. Using the formula that expresses the L_p norm a function in terms of its distribution function and Fubini's Theorem gives us

$$\begin{aligned}
\|Tf\|_{L^p}^p = \int_0^\infty p\lambda^{p-1} \mu_{Tf}(\lambda)\, d\lambda &\leq \int_0^\infty p\lambda^{p-1} \mu_{Tf_\lambda^1 + Tf_\lambda^2}(\lambda)\, d\lambda \\
&\leq \int_0^\infty p\lambda^{p-1} \mu_{Tf_\lambda^1}\left(\frac{\lambda}{2}\right) d\lambda + \int_0^\infty p\lambda^{p-1} \mu_{Tf_\lambda^2}\left(\frac{\lambda}{2}\right) d\lambda \\
&\leq \int_0^\infty p\lambda^{p-1-p_1} (2C_{p_1})^{p_1} \int_X |f_\lambda^1|^{p_1}\, d\mu\, d\lambda \\
&= \int_0^\infty p\lambda^{p-1-p_1} (2C_{p_1})^{p_1} \int_{\{|f|>c\lambda\}} |f|^{p_1}\, d\mu\, d\lambda \\
&= \int_0^\infty \int_X p\lambda^{p-1-p_1} (2C_{p_1})^{p_1} |f|^{p_1} \chi_{\{|f|>c\lambda\}}\, d\mu\, d\lambda \\
&= p(2C_{p_1})^{p_1} \int_X \int_0^\infty \lambda^{p-1-p_1} |f|^{p_1} \chi_{\{|f|>c\lambda\}}\, d\lambda\, d\mu \\
&= p(2C_{p_1})^{p_1} \int_X \int_0^{\frac{|f(x)|}{c}} \lambda^{p-1-p_1} |f(x)|^{p_1}\, d\lambda\, d\mu(x) \\
&= p(2C_{p_1})^{p_1} \int_X |f(x)|^{p_1} \int_0^{\frac{|f(x)|}{c}} \lambda^{p-1-p_1}\, d\lambda\, d\mu(x) \\
&= \frac{p(2C_{p_1})^{p_1}}{p-p_1} \int_X |f(x)|^{p_1} \left(\frac{|f(x)|}{c}\right)^{p-p_1} d\mu(x) \\
&= \frac{p(2C_{p_1})^{p_1}}{c^{p-p_1}(p-p_1)} \int_X |f(x)|^p\, d\mu(x).
\end{aligned}$$

As a result

$$\|Tf\|_{L^p}^p \le \frac{p(2C_{p_1})^{p_1}}{c^{p-p_1}(p-p_1)}\|f\|_{L^p}^p, \quad \text{where } c = \frac{1}{2C_\infty}$$

and by that we have proved our result for the first case.

As for the case where $p_2 < \infty$ let us first note here that the weak-L^{p_2} condition of T can be expressed as follows,

$$\mu_{Tf_\lambda^2}\left(\frac{\lambda}{2}\right) \le \left(\frac{2C_{p_2}}{\lambda}\right)^{p_2} \int_X |f_\lambda^2|^{p_2}\, d\mu.$$

As a result

$$\begin{aligned}
\|Tf\|_{L^p}^p = \int_0^\infty p\lambda^{p-1}\mu_{Tf}(\lambda)\, d\lambda &\le \int_0^\infty p\lambda^{p-1}\mu_{Tf_\lambda^1+Tf_\lambda^2}(\lambda)\, d\lambda \\
&\le \int_0^\infty p\lambda^{p-1}\mu_{Tf_\lambda^1}\left(\frac{\lambda}{2}\right) d\lambda + \int_0^\infty p\lambda^{p-1}\mu_{Tf_\lambda^2}\left(\frac{\lambda}{2}\right) d\lambda \\
&= \int_0^\infty p\lambda^{p-1}\left(\frac{2}{\lambda}\right)^{p_1}(C_{p_1})^{p_1}\int_X |f_\lambda^1|^{p_1}\, d\mu\, d\lambda \\
&+ \int_0^\infty p\lambda^{p-1}\left(\frac{2}{\lambda}\right)^{p_2}(C_{p_2})^{p_2}\int_X |f_\lambda^2|^{p_2}\, d\mu\, d\lambda.
\end{aligned}$$

For the first integral on the left in the inequality above we have already deduced an inequality which we can take advantage of, as for the second one following a similar procedure as the first one gives us,

$$\begin{aligned}
\int_0^\infty p\lambda^{p-1}\left(\frac{2}{\lambda}\right)^{p_2}(C_{p_2})^{p_2}\int_X |f_\lambda^2|^{p_2}\, d\mu\, d\lambda &= p(2C_{p_2})^{p_2}\int_0^\infty\int_X \lambda^{p-1-p_2}|f_\lambda^2|^{p_2}\, d\mu\, d\lambda \\
&= p(2C_{p_2})^{p_2}\int_0^\infty\int_{\{|f|\le c\lambda\}} \lambda^{p-1-p_2}|f|^{p_2}\, d\mu\, d\lambda \\
&= p(2C_{p_2})^{p_2}\int_0^\infty\int_X \lambda^{p-1-p_2}|f|^{p_2}\chi_{\{|f|\le c\lambda\}}\, d\mu\, d\lambda \\
&= p(2C_{p_2})^{p_2}\int_X\int_0^\infty \lambda^{p-1-p_2}|f|^{p_2}\chi_{\{|f|\le c\lambda\}}\, d\lambda\, d\mu \\
&= p(2C_{p_2})^{p_2}\int_X\int_{\frac{|f(x)|}{c}}^\infty \lambda^{p-1-p_2}|f(x)|^{p_2}\, d\lambda\, d\mu \\
&= p(2C_{p_2})^{p_2}\int_X |f(x)|^{p_2}\int_{\frac{|f(x)|}{c}}^\infty \lambda^{p-1-p_2}\, d\lambda\, d\mu \\
&= \frac{p(2C_{p_2})^{p_2}}{c^{p-p_2}(p-p_2)}\int_X |f(x)|^{p_2}|f(x)|^{p-p_2}\, d\mu \\
&= \frac{p(2C_{p_2})^{p_2}}{c^{p-p_2}(p-p_2)}\int_X |f(x)|^p\, d\mu.
\end{aligned}$$

Now combining these two results gives us

$$\|Tf\|_{L^p}^p \le \left(\frac{p(2C_{p_1})^{p_1}}{c^{p-p_1}(p-p_1)} + \frac{p(2C_{p_2})^{p_2}}{c^{p-p_2}(p_2-p)}\right)\|f\|_{L^p}^p,$$

and with that we have completed our proof. □

Remark 2.6. The strong-L^p inequalities we deduced above can equivalently be written as

$$\|Tf\|_p \leq 2p^{\frac{1}{p}} \left(\frac{1}{p-p_1} + \frac{1}{p_2 - p} \right)^{\frac{1}{p}} \|f\|_{L^p}$$

where $\theta \in (0,1)$ such that

$$\frac{1}{p} = \frac{\theta}{p_2} + \frac{1-\theta}{p_1}.$$

2.2 Dyadic Maximal Operator

Definition 2.7. Let f be a $L^1_{loc}(\mathbb{R}^n)$ function, the *dyadic maximal operator* of f (denoted by $\mathcal{M}_d f$) is defined as follows,

$$\mathcal{M}_d f(x) := \sup_{D \ni x : D \text{ dyadic}} \frac{\int_D |f|}{|D|},$$

where the dyadic cubes are the cubes formed by the grids $2^{-N}\mathbb{Z}^n$ for $N \in \mathbb{Z}$. We consider the intervals, their cartesian product of which constitutes a dyadic cube to be of the form $[a, b)$. As a result of it every two dyadic cubes are either disjoint or one of them is inside the other.

Now to that prove $\mathcal{M}_d$ is weak-L_1. Having that we will then prove the respective weak-L^1 condition for the Hardy-Littlewood Maximal Operator.

Theorem 2.8 (weak-L_1condition of Dyadic Maximal Operator)**.** *The Dyadic maximal operator is weak-L^1 operator. More specifically*

$$|\{x \in \mathbb{R}^n : \mathcal{M}_d f(x) > \lambda\}| \leq \frac{\|f\|_{L^1}}{\lambda}, \quad \text{for every } \lambda > 0 \text{ and } f \in L^1.$$

Proof. Let $f \in L^1$ and without loss of generality let us assume that $f \geq 0$ and for $m \in \mathbb{Z}$ let $D_m(x)$ be the dyadic cube of the grid $2^{-m}\mathbb{Z}^n$ that contains x and set

$$G_m := \left\{ x \in \mathbb{R}^n : \frac{1}{|D_m(x)|} \int_{D_m(x)} f > \lambda \text{ and } \frac{1}{|D_i(x)|} \int_{D_i(x)} f \leq \lambda, \quad \text{for } i < m \right\}$$

for $\lambda > 0$, i.e. the set of points of $\{\mathcal{M}_d f > \lambda\}$ that the dyadic cube of the grid $2^{-m}\mathbb{Z}^n$ that contains x is the first such that the average over it is greater than λ. Such a cube exists since

$$\frac{1}{|D_i(x)|} \int_{D_i(x)} f \to 0$$

as $i \to -\infty$ for every $x \in \mathbb{R}^n$, since f is an L^1 function.

Let us note here that the sets G_m are pairwise disjoint and can be written as a disjoint union

$$G_m = \bigcup_{i \in I^m} D_i^m$$

where D_i^m stands for a dyadic cube of the grid $2^{-m}\mathbb{Z}^n$ such that

$$\frac{1}{|D_i^m|}\int_{D_i^m} f > \lambda.$$

The reason the D_i^m are pairwise disjoint is that that they are maximal with respect to the condition above.

Consequently

$$\begin{aligned}|\{\mathcal{M}_d f > \lambda\}| &= \sum_{m\in\mathbb{Z}} |G_m| \\ &= \sum_{m\in\mathbb{Z}}\sum_{i\in I^m} |D_i^m| \\ &\le \frac{1}{\lambda}\sum_{m\in\mathbb{Z}}\sum_{i\in I^m}\int_{D_m^i} f \\ &\le \frac{1}{\lambda}\int_{\mathbb{R}^n} f = \frac{1}{\lambda}\|f\|_{L^1}\end{aligned}$$

□

The dyadic cubes we derived above will be of much importance to us and in fact this family of dyadic cubes is called the Calderon-Zygmund Decomposition [5].

Definition 2.9. Let f be an $L^1_{loc}(\mathbb{R}^n)$ and for $m \in \mathbb{Z}$ we define

$$E_m f(x) := \sum_{j\in\mathbb{N}} \frac{1}{D_j^m}\int_{D_j^m} f(y)\,dy \cdot \chi_{D_j}(x), \quad \text{for every } x \in \mathbb{R}^n$$

where the sum is over all the dyadic cubes of the grid $2^{-m}\mathbb{Z}^n$.

Lemma 2.10. *Let (X,μ) be a measure space, $\{T_i\}_{i\in\mathbb{N}}$ be a family of linear operators on $L^p(X)$ and set*

$$T^*f(x) := \sup_{i\in\mathbb{N}} |T_i f(x)|.$$

If T^* is weak-L^p for some $p \in [1,\infty)$ then the set

$$\{f \in L^p(X) : \lim_{i\to\infty} T_i f(x) = f(x), \textit{ for a.e. } x \in X\}$$

is closed in $L^p(X)$.

Proof. Let $\{f_j\}_{j\in\mathbb{N}}$ be a sequence of L^p functions such that $\|f_j - f\|_{L^p} \to 0$ and $T_i f_j(x) \to f_j(x)$ for a.e. $x \in X$ and $j \in \mathbb{N}$. Then for $j \in \mathbb{N}$ we have

$$\begin{aligned}&\mu(\{x \in X : \limsup_{i\to\infty} |T_i f(x) - f(x)| > \lambda\}) \\ &\qquad \le \mu(\{x \in X : \limsup_{i\to\infty} |T_i(f - f_j)(x) - (f - f_j)(x)| > \lambda\}) \\ &\qquad \le \mu(\{x \in X : T^*(f - f_j)(x) > \lambda/2\}) + \mu(\{x \in X : |(f - f_j)(x)| > \lambda/2\})\end{aligned}$$

$$\leq \left(\frac{2C_p}{\lambda}\|f-f_j\|_{L^p}\right)^p + \left(\frac{2}{\lambda}\|f-f_j\|_{L^p}\right)^p.$$

Our initial assumptions guarantees us that the right term in the inequality above tends to zero as $j \to \infty$, so

$$\mu(\{x \in X : \limsup_{i\to\infty} |T_i f(x) - f(x)| > 0\})$$
$$\leq \sum_{k=1}^{\infty} \mu(\{x \in X : \limsup_{i\to\infty} |T_i f(x) - f(x)| > 1/k\}) = 0. \qquad \square$$

Remark 2.11. It is straightforward that for every $k \in \mathbb{N}$, E_k is linear and using the same notation as in Lemma 2.10 we have that E^* is the dyadic maximal operator.

Lemma 2.12. *If f is an L^1_{loc} function then $\lim_{k\to\infty} E_k f(x) = f(x)$ for a.e. $x \in \mathbb{R}^n$.*

Proof. If f is a continuous function with compact support the result clearly holds. Then since $C_c(\mathbb{R}^n)$ are L^p dense and E^* is the dyadic maximal operator and therefore weak-L^1 in Lemma 2.10 we deduce that the result holds true for any L^1 function f. In case f is in L^1_{loc} but not L^1, we just need to consider its restrictions to the dyadic cubes of $\mathbb{Z}^n$ which are L^1. $\square$

Theorem 2.13 (**Calderon-Zygmund Decomposition**)**.** *Let $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$. Then there exists a family of disjoint dyadic cubes $\{D_i\}_{i\in I}$ such that*

1. $\{\mathcal{M}_d f > \lambda\} = \cup_{i\in I} D_i$
2. $f(x) \leq \lambda$, *for almost every $x \notin \{\mathcal{M}_d f > \lambda\}$*
3. $\lambda < \frac{1}{|D_i|}\int_{D_i} f\, dx \leq 2^n\lambda$, *for every dyadic cube D_i of the decomposition.*

Proof. The proof of the first one is just the first part of the proof of the weak-L^1 condition of the dyadic maximal operator (Theorem 2.8), *for the second one* we have that $E_m(x) \leq \lambda$ for every $m \in \mathbb{Z}$ and so by combining it with Lemma 2.12 we have our result.

As for the third and final one, the right direction has already been proved, as for the other part we just have to consider the least dyadic cube D_i^* that contains properly D_i. Then by construction of the D_i

$$\frac{1}{|D_i^*|}\int_{D_i^*} f \leq \lambda$$

and so

$$\frac{1}{|D_i|}\int_{D_i} f \leq 2^n\lambda. \qquad \square$$

So now we have proved that the dyadic maximal operator is weak L^1 and since it is straightforward that it is also weak-L^∞, the Marcinkiewicz Interpolation Theorem ensures us that it is strong L^p for every $1 < p < \infty$.

2.3 Weak-L^1 condition of the Hardy-Littlewood Maximal Operator

Now with the weak-L^1 condition of the dyadic maximal operator in mind, we are going to prove the weak-L^1 condition of the Hardy-Littlewood Maximal Operator. Before we move on with that let us prove here a lemma we will be needing.

Lemma 2.14. *Let Q be a cube of $\mathbb{R}^n$ such that for the length of its sides $l(Q)$ the following inequality holds for some $k \in \mathbb{Z}$,*

$$2^{k-1} \leq l(Q) < 2^k.$$

Then Q can be covered by 2^n dyadic cubes of the grid $2^k\mathbb{Z}^n$.

Proof. We will prove it for the case where $n = 1$. The multidimensional case then stems from the fact that any dyadic cube is the cartesian product of dyadic intervals.

Let us assume to the contrary that we need at least three dyadic intervals R_1, R_2, R_3 of G_k to cover Q where

$$R_1 = [l_1, r_1), \quad R_2 = [l_2, r_2) \quad \text{and} \quad R_3 = [l_3, r_3)$$

with $r_i - l_i = 2^k$ for $i = 1, 2, 3$ and $r_1 \leq l_2$ and $r_2 \leq l_3$.

Since Q has points in common with R_1 and R_3, its convexity guarantees us that it must contain R_2 as a whole which in turn implies that its length is greater than 2^k, a contradiction to our initial assumption about its length. □

Theorem 2.15 (Weak-L^1 condition of the Hardy-Littlewood Maximal Operator)**.** *The Hardy-Littlewood Maximal Operator is a weak-L^1 operator i.e.*

$$|\{\mathcal{M}f > \lambda\}| \leq \frac{C}{\lambda} \int_{\mathbb{R}^n} |f|\, dx, \quad \text{for every } \lambda > 0 \text{ and } f \in L^1$$

for some constant $C > 0$ which depends only on the dimension.

Proof. We will be using the cube centered variant of the Hardy-Littlewood Maximal Operator to prove that for every $\lambda > 0$,

$$|\{\mathcal{M}f > 4^n\lambda\}| \leq 2^n |\{\mathcal{M}_d f > \lambda\}|.$$

Then the weak-L^1 condition of the Hardy-Littlewood Maximal Operator stems directly from the respective one of the dyadic maximal operator (Theorem 2.8).

We will use for a $\lambda > 0$ the Calderon-Zygmund Decomposition at λ i.e.

$$\{\mathcal{M}_d f > \lambda\} = \bigcup_{i \in I} D_i.$$

We define $2D_i$ to be the cube that has the same center as D_i and its sides' length is twice the length of the sides of D_i, we will prove that

$$\{\mathcal{M}f > 4^n\lambda\} \subset \bigcup_{i \in I} 2D_i.$$

So let $x \notin \cup_{i\in I} 2D_i$ and consider any cube Q with x as its center. We now select a $k \in \mathbb{Z}$ such that $2^{k-1} \leq l(Q) < 2^k$, where $l(Q)$ stands for the length of Q. By Lemma 2.14 we have that Q is covered by $m \leq 2^n$ dyadic cubes of the grid $2^k\mathbb{Z}^n$, let call them $R_1, R_2, \dots, R_m$. These cubes are not contained in any of the cubes D_i of the Calderon-Zygmund Decomposition for else $x \in \cup_{i\in I} 2D_i$.

Indeed assuming first that $n = 1$ and

$$R_1 = [l, m) \quad \text{and} \quad R_2 = [m, n)$$

with $m - l = n - m = 2^k$. If some D_i contained for instance R_1 then that would mean that $2D_i$ would contain $2R_1$ where $2R_1$ stems by increasing by 2^{k-1} the length in each side and since the center x is at a distance of at most 2^{k-1} from the sides of Q we have that $2R_1$ contains x and as a result $2D_i$ contains x. As for the case where $n > 1$ it stems from the fact that the dyadic cubes constitute cartesian products of dyadic intervals.

As a result the average over each of these cubes can't be greater than λ and so,

$$\frac{1}{|Q|}\int_Q f\,dx = \frac{1}{|Q|}\sum_{i=1}^{m}\int_{Q\cap R_i} f\,dx \leq \sum_{i=1}^{m}\frac{2^{kn}}{|Q|}\frac{1}{|R_i|}\int_{R_i} f\,dx \leq 2^n m\lambda \leq 4^n\lambda. \qquad \square$$

2.4 Lebesgue's Differentiation Theorem

In this section we will prove Lebesgue's Differentiation Theorem. The proof presented here is from the book by Rubio De Francia and Garcia Cuerva "Weighted Norm Inequalities and Related Topics" [13].

Theorem 2.16 (**Lebesgue's Differentiation Theorem**). *If $f \in L^1_{loc}(\mathbb{R}^n)$ then*

$$\lim_{r\to 0}\frac{\int_{B(x,r)}|f(y) - f(x)|\,dy}{|B(x,r)|} \to 0, \quad \textit{for a.e. } x$$

and as a result

$$\lim_{r\to 0}\frac{\int_{B(x,r)} f\,dx}{|B(x,r)|} \to f(x), \quad \textit{for a.e. } x.$$

Proof. If f is a continuous function, it is straightforward that the result holds, so let us a consider a L^1 function f, not necessarily continuous and set

$$Tf(x) := \limsup_{r\to 0}\frac{\int_{B(x,r)}|f(y) - f(x)|\,dy}{|B(x,r)|}.$$

Then we have that $|Tf| \leq Mf + |f|$.

Due to the fact that T is sub-linear and for h continuous $Th = 0$ we also have that $|Tf| \leq M(f - h) + |f - h|$ for every continous function h .

Considering now for a fixed $\lambda > 0$ the set

$$A_\lambda := \{x \in X : Tf(x) > \lambda\}$$

and passing to the measures gives us

$$\begin{aligned} |A_\lambda| &= |\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \\ &\leq |\{x \in \mathbb{R}^n : M(f-h)(x) > \lambda/2\}| + |\{x \in \mathbb{R}^n : |f(x) - h(x)| > \lambda/2\}| \end{aligned}$$

and since the Hardy-Littlewood Maximal Operator is weak-L^1 we end up with the following inequality

$$|A_\lambda| = |\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq C\frac{\|f-h\|_{L^1}}{\lambda}$$

for some constant $C > 0$.

Now using the fact that the continuous functions of compact support are dense in L^1 we obtain

$$|A_\lambda| = |\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| = 0$$

for every $\lambda > 0$.

As for the second part of the theorem we just have to notice that

$$\left|\frac{\int_{B(x,r)} f(y)\,dy}{|B(x,r)|} - f(x)\right| = \left|\frac{\int_{B(x,r)} f(y) - f(x)\,dy}{|B(x,r)|}\right| \leq \frac{\int_{B(x,r)} |f(y) - f(x)|\,dy}{|B(x,r)|}. \qquad \square$$

2.5 Development of the basic theory using the rearrangement theory

What we going to do here is to investigate the basic theorems we have established so far using rearrangements and by doing so we are going to end up generalizing the Marcinkiewicz Interpolation theorem to some degree. The source for all of these topics presented here is the book of Bennet and Sharpley "Interpolation of Operators"[2].

2.5.1 Another proof of Lebesgue's Differentiation Theorem

We begin with another proof of Lebesgue's Differentiation Theorem. In order to proceed to this direction let us first prove two useful lemmas which we will need. Note also that this time we will be using cubes instead of balls in the proof of the theorem.

Lemma 2.17. *Let A be a subset of $\mathbb{R}^n$ with finite measure and $\mathcal{F}$ be a collection of cubes that covers A. Then there exists a finite pairwise disjoint collection of cubes of $\mathcal{F}$, $Q_1, \dots, Q_m$ such that*

$$\sum_{i=1}^{m} |Q_i| \geq 4^{-n}|A|.$$

Proof. We will prove that if A is compact and the Q_i are open cubes then

$$\sum_{i=1}^{m} |Q_i| \geq 3^{-n}|A|. \tag{2.1}$$

Given that the more general case stems from the facts that we can approximate any cube Q by larger open cube of measure $a|Q|$ for every $a > 1$ and that we can approximate the Lebesgue measure of A by compact subsets of A of measure $b|A|$ for every $b < 1$ and by that is enough to choose them so that

$$\frac{b}{a}3^{-n} = 4^{-n}.$$

Now due to the compactness of A there exists a finite subcover from $\mathcal{F}$ of A so we may also assume that the covering is finite.

1. Now let us first select a cube which has the greatest diameter amongst all of $\mathcal{F}$ and call it Q_1.
2. Next we select a second cube (assuming there exists) Q_2 that has the greatest diameters among all cubes that are disjoint of Q_1.
3. Next, (again assuming there exists) we select a cube Q_3 that has the greatest diameter among all the cubes that are disjoint of Q_1 and Q_2 and so.

This procedure finishes after a finite number of steps since $\mathcal{F}$ is finite. Now let us denote the cubes we elected in the procedure above by $Q_1, Q_2, \dots, Q_m$ depending on which step they were selected.

Now we prove that their respective measures satisfy inequality. For every Q_i ($i = 1, \dots, m$) let $\overline{Q}_i$ to be the cube that has the same center as Q_i and the length of its side is three times the length of the side of Q_i.

By showing that $\overline{Q}_i$ cover A we will have also proved that

$$|A| \leq \sum_{i=1}^{m} |\overline{Q}_i| = \sum_{i=1}^{m} 3^n |Q_i|.$$

So let us assume to the contrary that $\overline{Q_i}$ don't cover A. Then there exists an $x \in A \setminus \cup_{i=1}^{m} \overline{Q_i}$. Since $\mathcal{F}$ is a covering of A there exists a cube Q_x of $\mathcal{F}$ such that $x \in Q_x$.

From construction Q_x can't be larger than Q_1 and Q_x contains x which doesn't belong to $\overline{Q_1}$. As a result Q_x and Q_1 are disjoint. For similar reasons Q_x is disjoint from all Q_i, but this cannot happen by the construction of the cubes Q_i, $i = 1, \dots, m$. □

Theorem 2.18. *If f is an $L^1(\mathbb{R}^n)$ function then*

$$t(\mathcal{M}f)^*(t) \leq 4^n \|f\|_{L^1}, \quad \text{for every } t > 0.$$

Proof. Let us first assume that $f \neq 0$ and that f has compact support. For convenience let us at first use the centered ball variant of the Hardy-Littlewood Maximal Operator to prove the existence of an $r > 0$ such that

$$\frac{C}{|x|^n} \geq \mathcal{M}_{qc}f(x), \quad \text{for } |x| \geq r.$$

Since f if of compact support the points where f doesn't vanish are contained inside a ball $B(0,l)$ for some $l>0$ and so considering any point x with $|x|\geq 2l$ if the average integrals over balls centered at x are non-zero the balls must be larger or equal to the ball $B(x,|x|-l)$ so that they intersect the ball $B(0,l)$ and by extension the points where f doesn't vanish. So

$$\begin{aligned}\mathcal{M}_{qc}f(x) &\leq \frac{\int_{\mathbb{R}} f\,dx}{|B(x,|x|-l)|}\\ &\leq \frac{C}{(|x|-l)^n} \leq \frac{C'}{|x|^n}\end{aligned}$$

for some constant C' depending on f. As a result the cube variant of the Hardy-Littlewood also has this property and also for each $\lambda>0$ the set $E_\lambda := \{\mathcal{M}f>\lambda\}$ is of finite measure.

Now for every $x\in\{\mathcal{M}f>\lambda\}$ there exists a cube Q_x containing x such that

$$\lambda|Q_x| < \int_{Q_x} |f(y)|\,dy.$$

Applying Lemma 2.17 to the cubes we selected above gives us a finite sequence of them $Q_1,Q_2,\ldots,Q_m$ such that

$$\sum_{i=1}^m |Q_i| \geq 4^{-n}|E_\lambda|.$$

Combing these two properties now gives us

$$|E_\lambda| \leq 4^n\sum_{i=1}^m |Q_i| \leq \frac{4^n}{\lambda}\sum_{i=1}^m \int_{Q_i}|f| \leq \frac{4^n}{\lambda}\|f\|_{L^1}.$$

As for the more general case one just needs to use the Monotone Convergence Theorem. □

Note 2.19. Using Theorem 2.18 we can now prove Lebesgue's Differentiation Theorem. We will prove later that the inequality stated in Theorem 2.18 above is equivalent to the fact $\mathcal{M}$ is weak-L^1.

Second proof of Lebesgue's Differentiation Theorem. For an L^1-function f let

$$Tf(x) := \limsup_{|Q|\to 0}\frac{1}{|Q|}\int_Q |f(y)-f(x)|\,dy.$$

Similar to the first proof that we gave at the beginning of the chapter we have that

$$Tf \leq Mf + |f|.$$

Remembering one of the basic properties of the decreasing rearrangement of a function (Theorem 1.11) we then have that

$$(Tf)^*(t) \leq (\mathcal{M}f)^*\left(\frac{t}{2}\right) + f^*\left(\frac{t}{2}\right), \quad \text{for every } t > 0.$$

Now using the inequality that we have proved in the previous theorem (Theorem 2.18) and the fact that

$$f^*\left(\frac{t}{2}\right) \leq f^{**}\left(\frac{t}{2}\right) \leq \left(\frac{2}{t}\right)\|f\|_{L^1}$$

we conclude that

$$Mf \leq \frac{c}{t}\|f\|_{L^1}$$

for some constant c depending only on the dimension. Since the rest of this proof follows the exact same pattern as the first one so we omit it. □

2.5.2 L^p-boundedness of the Hardy-Littlewood Maximal Operator

Lemma 2.20. *Let Ω be an open subset of $\mathbb{R}^n$ of finite measure. Then there exists a disjoint sequence of dyadic cubes $\{D_i\}_{i\in I}$ that covers Ω and such that*

1. $D_i \cap \Omega^c \neq \emptyset$, *for every $i \in I$*
2. $\sum_{i\in I} |D_i| \leq 2^n|\Omega|$.

Proof. Since Ω is open for every $x \in \Omega$ there exists a dyadic cube that contains x and that is contained wholly within Ω. Furthermore since Ω is also of finite not every dyadic cube that contains x is within Ω. As a result there exists a dyadic cube with the smallest diameter that contains x and isn't wholly within Ω. Since these dyadic cubes are maximal with respect to this condition they can't be contained in a bigger one of the same family, so they are pairwise disjoint.

If now $\{D_i\}_{i\in I}$ is the sequence (which can be finite) of these dyadic cubes it is straightforward that it is a covering of A and satisfies the first property. Let us also note here that since these dyadic cubes are maximal with respect to the condition above one of them can't be contained inside an other so they are disjoint with each other.

As for the second one since every $x \in \Omega$ belongs to some D_{i_x} also belongs to some dyadic subcube $\widetilde{D}_{i_x}$ of D_{i_x} of measure $2^{-n}D_{i_x}$, we obtain

$$2^{-n}|D_{i_x}| = |\widetilde{D}_{i_x}| = |\widetilde{D}_{i_x} \cap \Omega| \leq |D_{i_x} \cap \Omega|.$$

By that we then have

$$\sum_{i\in I} |D_i| \leq 2^n \sum_{i\in I} |D_i \cap \Omega| = 2^n|\Omega|.$$

□

Theorem 2.21. *Let (X,μ) be a finite measure space and f an a.e. finite function defined on X. Then*

$$\inf\{\|g\|_{L^1}+t\|h\|_{L^\infty} : f=g+h,\ g\in L^1 \text{ and } h\in L^\infty\} = \int_0^t f^*(s)\,ds = tf^{**}(t), \quad t>0.$$

Proof. The second equality stems directly from the definition of f^{**}. As for the first one let us set a_t to be equal to the infimum on the left side for start.

We will first we prove that

$$\int_0^t f^*(s)\,ds \le a_t.$$

Let us assume that $f\in L^1+L^\infty$ for else the infimum is equal to infinity for else there is nothing to prove.

So there is an L^1 function g and a L^∞ function h such that

$$f=g+h$$

and so

$$f^{**}\le g^{**}+h^{**}.$$

From the definition of f^{**} it now stems that

$$\int_0^\infty f^*(s)\,ds \le \|g\|_{L^1}+t\|h\|_{L^\infty}.$$

Considering the infimum over all such pairs of g,h such that $f=g+h$ we have our result.

Now we prove the reverse inequality. We just have to find $g\in L^1$ and $h\in L^\infty$ such that

$$f=g+h \quad \text{and} \quad \|g\|_{L^1}+t\|h\|_{L^\infty}\le\int_0^t f^*(s)\,ds.$$

We may assume also here that the right side is finite for else we have nothing to prove. As a result of that assumption combined with Hardy's Inequality we have that f is integrable over sets of measure less or equal to t.

Now setting $E:=\{|f|>f^*(t)\}$ and taking into account that f and f^* are equimeasurable we obtain

$$t_0:=\mu(E)\le t,$$

which means that f is integrable over E.

Now if we consider the functions g and h defined by

$$g(x)=\max\{|f(x)|-f^*(t),0\}\cdot\operatorname{sgn}f(x)$$
$$h(x)=\min\{|f(x)|,f^*(t)\}\cdot\operatorname{sgn}f(x)$$

then $f=g+h$ with $g\in L^1$ since

$$\|g\|_{L^1}=\int_E(|f|-f^*(t))\,d\mu=\int_E|f|\,d\mu-\mu(E)f^*(t)\le\int_0^{t_0}f^*(s)\,ds-t_0f^*(t)$$

and $h \in L^\infty$ with $\|h\|_{L^\infty} \leq f^*(t)$.

As a result

$$\|g\|_{L^1} + t\|h\|_{L^\infty} \leq \int_0^{t_0} f^*(s)ds + (t - t_0)f^*(t) \tag{2.2}$$

but f^* is constant and equal to $f^*(t)$ on $[t_0, t]$.

Indeed since $t_0 \leq t$ the monotonicity of f^* gives us $f^*(t) \leq f^*(t_0)$. As for the opposite direction since

$$t_0 = \mu(\{f > f^*(t)\}) \leq t$$

we then have

$$f^*(t_0) = \inf\{\lambda > 0 : \mu(\{|f| > \lambda\}) \leq t_0\} \leq f^*(t).$$

So inequality (2.2) is in fact the one we were seeking. □

Theorem 2.22. *There are constants c_1 and c_2 depending on n such that*

$$c_1(\mathcal{M}f)^*(t) \leq f^{**}(t) \leq c_2(\mathcal{M}f)^*(t), \quad \textit{for every } f \in L^1_{loc}(\mathbb{R}^n) \textit{ and } t > 0.$$

Proof. First we prove the left inequality. Let $t > 0$ and suppose that $f^{**}(t) < \infty$ for else there is nothing to prove. Then by Theorem 2.21, given $\epsilon > 0$ there exist an L^1 function g_t and an L^∞ function h_t such that $f = g_t + h_t$ with

$$\|g_t\| + t\|h_t\| \leq tf^{**}(t) + \epsilon.$$

Using the subadditivity of the maximal operator and Theorem 1.11 we obtain that for every $t > 0$

$$\begin{aligned}(\mathcal{M}f)^*(t) &\leq (\mathcal{M}g_t)^*\left(\frac{t}{2}\right) + (\mathcal{M}h_t)^*\left(\frac{t}{2}\right) \leq \frac{c}{t}\|g_t\|_{L^1} + \|h_t\|_{L^\infty} \\ &\leq \frac{c}{t}\left(\|g_t\| + t\|h\|_{L^\infty}\right) \leq tf^{**}(t) + \epsilon.\end{aligned}$$

Now letting $\epsilon \to 0$ gives us the left inequality.

As for the right inequality, let us assume that $(\mathcal{M}f)^*(t) < \infty$ for else there is nothing to prove. If we set

$$\Omega := \{x \in \mathbb{R}^n : (Mf)(x) > (Mf)^*(t)\}$$

we have that this set is open and since Mf and $(Mf)^*$ are equimeasurable we also have that $|\Omega| \leq t$.

Applying Lemma 2.20 gives us a sequence $\{D_i\}_{i\in I}$ of pairwise disjoint dyadic cubes such that

$$D_i \cap \Omega^c \neq \emptyset \quad \text{and} \quad \sum_{i\in I} |D_i| \leq 2^n|\Omega| \leq 2^n t.$$

Now by setting $F := (\cup_{i\in I} D_i)^c$, $g := \sum_{i\in I} f\chi_{D_i}$ and $h := f\chi_F$ we have that $f = g + h$ and

$$f^{**}(t) \leq g^{**}(t) + h^{**}(t) \leq \frac{1}{t}\|g\|_{L^1(\mathbb{R}^n)} + \|h\|_{L^\infty(\mathbb{R}^n)}.$$

From the construction of Q_i we have that each of them contains a point in Ω^c (i.e. such that Mf at that point is less or equal to $(Mf)^*(t)$) and so

$$\frac{1}{|D_i|}\int_{D_i} f(y)dy \leq (\mathcal{M}f)^*(t)$$

and as a result

$$\|g\|_{L^1(\mathbb{R}^n)} = \sum_{i\in I}\int_{D_i} |f(x)|\,dx \leq \sum_{i\in I}|D_i|(\mathcal{M}f)^*(t) \leq 2^n t(\mathcal{M}f)^*(t).$$

As for h, since F lies within Ω^c,

$$\|h\|_{L^\infty} = \|f\chi_F\|_{L^\infty} \leq \|(\mathcal{M}f)\chi_F\|_{L^\infty} \leq \mathcal{M}f^*(t).$$

Combining now these two results gives us the right inequality. □

Lemma 2.23 (Hardy's Inequalities)**.** *If ψ is a nonnegative measurable function then for every λ such that $-\infty < \lambda < 1$, we have that*

1. $\left\{\int_0^\infty \left(t^\lambda \frac{1}{t}\int_0^t \psi(s)\,ds\right)^q \frac{dt}{t}\right\}^{\frac{1}{q}} \leq \frac{1}{1-\lambda}\left\{\int_0^\infty (t^\lambda\psi(t))^q \frac{dt}{t}\right\}^{\frac{1}{q}}$,

2. $\left\{\int_0^\infty \left(t^{1-\lambda}\int_t^\infty \psi(s)\,\frac{ds}{s}\right)^q \frac{dt}{t}\right\}^{\frac{1}{q}} \leq \frac{1}{1-\lambda}\left\{\int_0^\infty (t^{1-\lambda}\psi(t))^q \frac{dt}{t}\right\}^{\frac{1}{q}}$.

Proof. For the first inequality we express $\psi(s)$ as

$$s^{-\frac{\lambda}{q'}}\psi(s)s^{\frac{\lambda}{q'}}$$

and apply Hölder's inequality to get

$$\frac{1}{t}\int_0^t \psi(s)\,ds \leq \frac{1}{t}\left(\int_0^t s^{-\lambda}\,ds\right)^{\frac{1}{q'}}\left(\int_0^t s^{\frac{\lambda q}{q'}}\psi(s)^q\,ds\right)^{\frac{1}{q}}$$
$$= (1-\lambda)^{-\frac{1}{q'}}t^{-\frac{\lambda}{q'}-\frac{1}{q}}\left(\int_0^t s^{\lambda(q-1)}\psi(s)^q\,ds\right)^{\frac{1}{q}}.$$

From that we obtain

$$\int_0^\infty \left(t^\lambda\frac{1}{t}\int_0^t \psi(s)\,ds\right)^q\frac{dt}{t} \leq (1-\lambda)^{1-q}\int_0^\infty t^{\lambda-2}\int_0^t s^{\lambda(q-1)}\psi(s)^q\,ds\,dt$$
$$= (1-\lambda)^{1-q}\int_0^\infty s^{\lambda(q-1)}\psi(s)^q\int_s^\infty t^{\lambda-2}\,dt\,ds$$

using Tonneli's theorem in the second step and so

$$\int_0^\infty \left(t^\lambda\frac{1}{t}\int_0^t \psi(s)\,ds\right)^q\frac{dt}{t} \leq \frac{1}{(1-\lambda)^q}\int_0^\infty s^{\lambda q}\psi(s)^q\,\frac{ds}{s}.$$

As for the second one we write $\psi(s)/s$ as

$$s^{\frac{\lambda-2}{q'}} \cdot \psi(s)/s \cdot s^{\frac{2-\lambda}{q'}}$$

and then similar to before apply Hölder's inequality and Tonelli's Theorem consequently

$$\begin{aligned} t^{1-\lambda}\int_t^\infty \psi(s)\frac{ds}{s} &\leq t^{1-\lambda}\left(\int_t^\infty s^{\lambda-2}\,ds\right)^{\frac{1}{q'}} \cdot \left(\int_t^\infty \left(\frac{\psi(s)}{s}\right)^q s^{(2-\lambda)(q-1)}\,ds\right)^{\frac{1}{q}} \\ &= t^{1-\lambda}t^{\frac{\lambda-1}{q'}}(1-\lambda)^{-\frac{1}{q'}}\left(\int_t^\infty \psi(s)^q s^{(2-\lambda)(q-1)-q}\,ds\right)^{\frac{1}{q}}. \end{aligned}$$

As a result

$$\begin{aligned} &\int_0^\infty \left(t^{1-\lambda}\int_t^\infty \psi(s)\frac{ds}{s}\right)\frac{dt}{t} \\ &\qquad \leq t^{(1-\lambda)q+(\lambda-1)(q-1)}(1-\lambda)^{-(q-1)}\int_t^\infty \psi(s)^q s^{(2-\lambda)(q-1)-q}\,ds\,\frac{dt}{t} \\ &\qquad = \int_0^\infty t^{-\lambda}(1-\lambda)^{1-q}\int_t^\infty \psi(s)^q s^{(2-\lambda)(q-1)-q}\,ds\,dt \\ &\qquad = (1-\lambda)^{1-q}\int_0^\infty\int_0^s t^{-\lambda}\psi(s)^q s^{(2-\lambda)(q-1)-q}\,dt\,ds \\ &\qquad = (1-\lambda)^{-q}\int_0^\infty s^{1-\lambda}s^{(2-\lambda)(q-1)-q}\psi(s)^q\,ds \\ &\qquad = (1-\lambda)^{-q}\int_0^\infty s^{(1-\lambda)q}\psi(s)^q\,\frac{ds}{s} \end{aligned}$$

and so we have proven both inequalities. □

Theorem 2.24 (Hardy-Littlewood Maximal Theorem). *For every p such that $1 < p < \infty$ the Hardy-Littlewood Maximal Operator is bounded from $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ i.e.*

$$\|\mathcal{M}f\|_{L^p(\mathbb{R}^n)} \leq C\|f\|_{L^p(\mathbb{R}^n)}, \quad \textit{for every } f \in L^p(\mathbb{R}^n)$$

for some constant $C > 0$ depending only on n and p.

Proof. From Theorem 1.12 we have,

$$\|\mathcal{M}f\|_{L^p(\mathbb{R}^n)} = \left(\int_0^\infty ((Mf)^*(t))^p\,dt\right)^{\frac{1}{p}}.$$

Now using Theorem 2.22 we obtain

$$\|\mathcal{M}f\|_{L^p(\mathbb{R}^n)} \leq C\left(\int_0^\infty \left(\frac{1}{t}\int_0^t f^*(s)ds\right)^p dt\right)^{\frac{1}{p}}$$

for some constant $C = C(n,p)$ and so applying the first of Hardy's Inequalities (Theorem 2.23) to the right side above for $\lambda = \frac{1}{p}$ and $q = p$ gives us

$$\|\mathcal{M}f\|_{L^p(\mathbb{R}^n)} \leq cp'\left(f^*(t)^p dt\right) = cp'\|f\|_{L^p}.$$ □

2.5.3 Marcinkiewicz Interpolation Theorem

The weak-L^p condition which we have previously defined as a substitute of the strong -L^p, is a condition that the Hardy-Littlewood Maximal Operator satisfies when $p = 1$ and in conjunction with its L^∞ boundedness it allowed us to prove the operator's strong-L^p boundedness for $p > 1$ using the Marcinkiewicz Interpolation Theorem. Here we are going to list a more general form of this condition, which will also result in a slight generalization of the Marcinkiewicz Interpolation theorem. In order to do that we begin with the introduction of the Lorentz spaces and the Calderon Operator.

Definition 2.25. Let (X, μ) be a σ-finite measure space and $0 < p, q \leq \infty$. The *Lorentz space* $L^{p,q}(X)$ is the set of all measurable a.e. finite functions f such that the quantity

$$\|f\|_{p,q} := \begin{cases} \left\{\int_0^\infty [t^{\frac{1}{p}} f^*(t)]^q \frac{dt}{t}\right\}^{\frac{1}{q}}, & \text{if } q < \infty \\ \sup_{0<t<\infty} t^{\frac{1}{p}} f^*(t), & \text{if } q = \infty \end{cases}$$

is finite.

Remark 2.26. For $q = p$ the spaces $L^{p,p}$ and L^p with their respective norms coincide.

Theorem 2.27. *For every $p \geq 1$ and $q < r$ we have that*

$$L^{p,r} \subset L^{p,s}.$$

In particular there is a constant c depending on p, q, r such that

$$\|f\|_{L^{p,s}} \leq c\|f\|_{L^{p,r}}.$$

Proof. Let us first assume that $p < \infty$ and $f \in L^{p,q}$. We will first prove that the condition holds for $r = \infty$. Indeed

$$t^{\frac{1}{p}} f^*(t) = \left(\frac{p}{q}\int_0^t \left[s^{\frac{1}{p}} f^*(t)\right]^q \frac{ds}{s}\right)^{\frac{1}{q}} \leq \left(\frac{p}{q}\int_0^t \left[s^{\frac{1}{p}} f^*(s)\right]^q \frac{ds}{s}\right)^{\frac{1}{q}} \leq \left(\frac{p}{q}\right)^{\frac{1}{q}} \|f\|_{p,q}$$

where we used the fact that f^* is decreasing in the second step. Consequently

$$\|f\|_{p,\infty} \leq \left(\frac{p}{q}\right)^{\frac{1}{q}} \|f\|_{p,q}.$$

Now for the case where $r < \infty$ we have

$$\|f\|_{p,q} = \left(\int_0^\infty \left[t^{\frac{1}{p}} f^*(t)\right]^{r-q+q} \frac{dt}{t}\right)^{\frac{1}{r}}. \qquad \square$$

Definition 2.28. Let X, Y be two σ-finite measure spaces. An operator T defined on a subspace of the finite a.e. functions on X to the space of the measurable functions on Y is called *quasilinear* if there is a constant $k \geq 1$ such that

$$|T(f + g)(x)| \leq k|Tf(x)| + k|Tg(x)|, \quad \text{for a.e. } x \in X$$

and every f, g inside the domain of T and

$$|T(\lambda f)| = \lambda |Tf|, \quad \text{for every } \lambda > 0$$

and every f belonging to the domain of T.

Definition 2.29. Let $1 \leq p_0 < p_1 \leq \infty$ and $1 \leq q_0 \neq q_1 \leq \infty$ and

$$m := \frac{\frac{1}{q_0} - \frac{1}{q_1}}{\frac{1}{p_0} - \frac{1}{p_1}}.$$

The *Calderon operator* S_σ with respect to $(p_0, q_0; p_1, q_1)$ is defined as

$$(S_\sigma f)(t) = t^{-\frac{1}{q_0}} \int_0^{t^m} s^{\frac{1}{p_0}} f^*(s) \frac{ds}{s} + t^{-\frac{1}{q_1}} \int_{t^m}^\infty s^{\frac{1}{p_1}} f^*(s) \frac{ds}{s}, \quad \text{t>0}.$$

Definition 2.30 (Joint weak $(p_0, q_0; p_1, q_1)$ condition)**.** An operator T is *joint weak type* $(p_0, q_0; p_1, q_1)$ if

$$(Tf)^*(t) \leq C S_\sigma(f^*)(t)$$

where $C > 0$ and S_σ stands for the Calderon Operator defined above.

Definition 2.31. Let $(X, \mu), (Y, \nu)$ be two σ-finite measure spaces, $1 \leq p, q \leq \infty$ and let T is an operator from $L^{p,1}$ to the space of the measurable functions of Y. Then T is called *weak-*(p, q) if it is a bounded operator from $L^{p,1}(X)$ to $L^{q,\infty}$ i.e.

$$\|Tf\|_{q,\infty} \leq C\|f\|_{p,1}, \quad \text{for every } f \in L^{p,1}.$$

Remark 2.32. The following two conditions are equivalent to the above definition

1. $(Tf)^*(t) \leq C t^{-\frac{1}{q}} \|f\|_{p,1}$
2. $\nu(\{|Tf| > \lambda\}) \leq \left[\frac{C}{\lambda}\|f\|_{p,1}\right]^q$ (which is in fact the condition we used before).

Indeed going from our initial definition to 1. and from there to 2. we just need to use the definitions of the (q, ∞) norm and the definition of the decreasing rearrangement respectively. As for the opposite direction again it is enough to consider the definitions we just mentioned.

Lemma 2.33. *Suppose* $1 \leq p_0 < p_1 < \infty$ *and* $1 \leq q_0, q_1 \leq \infty$. *Then for* $i = 0, 1$

$$t^{\frac{1}{q_i}} S_\sigma(f^*)(t) \leq \int_0^\infty s^{\frac{1}{p_i}} f^*(s) \frac{ds}{s}, \quad \textit{for } t > 0 \textit{ and } f \in L^{p_i,1}[0, \infty).$$

Proof. Let us prove the case where $i = 0$.

From the definition of S_σ we have

$$t^{\frac{1}{q_0}} S_\sigma(f^*)(t) = \int_0^{t^m} s^{\frac{1}{p_0}} f^*(s) \frac{ds}{s} + t^{\frac{1}{q_0} - \frac{1}{q_1}} \int_{t^m}^\infty s^{\frac{1}{p_1}} f^*(s) \frac{ds}{s}.$$

Now since $p_0 < p_1$, for $s \geq t^m$

$$t^{\frac{1}{q_0}-\frac{1}{q_1}} = (t^m)^{\frac{1}{p_0}-\frac{1}{p_1}} \leq s^{\frac{1}{p_0}-\frac{1}{p_1}}$$

and so

$$\begin{aligned} t^{\frac{1}{q_0}} S_\sigma(f^*)(t) &\leq \int_0^{t^m} s^{\frac{1}{p_0}} f^*(s)\,\frac{ds}{s} + \int_{t^m}^{\infty} s^{\frac{1}{p_0}-\frac{1}{p_1}+\frac{1}{p_1}} f^*(s)\,\frac{ds}{s} \\ &= \int_0^{\infty} s^{\frac{1}{p_0}} f^*(s)\,\frac{ds}{s}. \end{aligned}$$

The proof for $i = 1$ is similar so we omit it. □

Theorem 2.34. *Suppose* $1 \leq p_0 < p_1 < \infty$ *and* $1 \leq q_0 \neq q_1 \leq \infty$*. A quasilinear operator* T *is of joint weak type* $(p_0, q_0; p_1, q_1)$ *if and only if it is weak* (p_0, q_0) *and* (p_1, q_1)*.*

Proof. Let us suppose first that T *is weak* (p_0, q_0) *and* (p_1, q_1)*.* and $f \in L^{p_0,1} + L^{p_1,1}$. For a fixed $t > 0$ and m defined as in Definition 2.29 let

$$\begin{aligned} f_1(x) &:= \min[|f(x)|, f^*(t^m)] \cdot \operatorname{sgn} f(x); \\ f_0(x) &= f(x) - f_1(x) = [|f(x)| - f^*(t^m)]^+ \cdot \operatorname{sgn} f(x). \end{aligned}$$

For f_1 we have that

$$\begin{aligned} |\{|f_1| > \lambda\}| = |\{\min\{|f|, f(t^m)\} > \lambda\}| &= \begin{cases} 0, & \text{if } \lambda \geq t^m \\ |\{|f| > \lambda\}|, & \text{if } 0 \leq \lambda < t^m \end{cases} \\ &= \begin{cases} 0, & \text{if } \lambda \geq t^m \\ |\,\{f^* > \lambda\}\,|, & \text{if } 0 \leq \lambda < t^m. \end{cases} \end{aligned}$$

So

$$f_1^*(s) = \min[f^*(s), f^*(t^m)]$$

and by that

$$\|f_1\|_{p_1,1} = p_1 t^{\frac{m}{p_1}} f^*(t^m) + \int_{t^m}^{\infty} s^{\frac{1}{p_1}} f^*(s)\,\frac{ds}{s}. \tag{2.3}$$

Following a similar procedure for f_0 we obtain that

$$f_0^*(s) = [f^*(s) - f^*(t^m)]^+$$

and as a result

$$\|f_0\|_{p_0,1} = \int_0^{t^m} s^{\frac{1}{p_0}} f^*(s)\,\frac{ds}{s} - p_0 t^{\frac{m}{p_0}} f^*(t^m). \tag{2.4}$$

Now since T is quasilinear there is a constant k so that

$$\begin{aligned} (Tf)^*(t) &\leq [k(|Tf_0| + |Tf_1|)](t) \\ &\leq k\left[(Tf_0)^*\left(\frac{t}{2}\right) + (Tf_1)^*\left(\frac{t}{2}\right)\right]. \end{aligned}$$

From the weak type hypotheses of T we have

$$(Tf_i)^*\left(\frac{t}{2}\right) \leq \left(\frac{t}{2}\right)^{-\frac{1}{q_i}} C_i\|f_i\|_{p_i,1}, \quad \text{for } i = 0, 1$$

for some constants $C_0, C_1 > 0$.

Combining now inequalities 2.3 and 2.4 with the above inequality gives us

$$\begin{aligned}(Tf)^*(t) &\leq \left(\frac{t}{2}\right)^{-\frac{1}{q_0}} C_0\|f_0\|_{p_0,1} + \left(\frac{t}{2}\right)^{-\frac{1}{q_1}} C_1\|f_1\|_{p_1,1} \\ &\leq \left(\frac{t}{2}\right)^{-\frac{1}{q_0}} C_0 \left\{p_1 t^{\frac{m}{p_1}} f^*(t^m) + \int_{t^m}^{\infty} s^{\frac{1}{p_1}} f^*(s)\,\frac{ds}{s}\right\} \\ &+ \left(\frac{t}{2}\right)^{-\frac{1}{q_1}} C_1 \left\{\int_0^{t^m} s^{\frac{1}{p_0}} f^*(s)\,\frac{ds}{s} - p_0 t^{\frac{m}{p_0}} f^*(t^m)\right\}.\end{aligned}$$

So by setting $C := k \max_{i=1,2} p_i C_i 2^{\frac{1}{q_i}}$ we obtain

$$(Tf)^*(t) \leq CS_\sigma f(t).$$

As for the opposite direction if T is joint weak-type$(p_0, q_0; p_1, q_1)$ i.e. for every f in the domain of f

$$(Tf)^*(t) \leq CS_\sigma f^*(t)$$

for some constant $C > 0$ or equivalently

$$t^{-\frac{1}{q_i}}(Tf)^*(t) \leq Ct^{-\frac{1}{q_i}} S_\sigma f^*(t).$$

Now using Lemma 2.33 we have our result. □

Theorem 2.35 (Marcinkiewicz Interpolation Theorem). *Suppose* $1 \leq p_0 < p_1 < \infty$ *and* $1 \leq q_0 \neq q_1 \leq \infty$ *and for* $\theta \in (0,1)$ *let* p, q *defined as follows*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \quad \textit{and} \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

If X,Y *are* σ*-finite measure spaces and* T *is a quasilinear operator from the space* $(L^{p_0,1} + L^{p_1,1})(X)$ *to the space of measurable functions of* Y*, that is weak-*(p_0, q_0) *and weak-*(p_1, q_1) *with respective constants* C_0 *and* C_1 *then for every* r *such that* $1 \leq r \leq \infty$*,* T *is bounded from* $L^{p,r}$ *to* $L^{p,r}$*.*

In particular there is a constant c *depended on* p_o, q_0, p_1, q_1 *such that*

$$\|Tf\|_{q,r} \leq \frac{c}{\theta(1-\theta)} \max(C_0, C_1)\|f\|_{p,r}.$$

Proof. Let us first note that for

$$m = \frac{\frac{1}{q_0} - \frac{1}{q_1}}{\frac{1}{p_0} - \frac{1}{p_1}}$$

we have

$$\frac{1}{m}\left[\frac{1}{q} - \frac{1}{q_0}\right] = \frac{1}{p} - \frac{1}{p_0} \quad \text{and} \quad \frac{1}{m}\left[\frac{1}{q} - \frac{1}{q_1}\right] = \frac{1}{p} - \frac{1}{p_1}.$$

Now from Theorem 2.34 we have that T is also of joint weak-type $(p_0, q_0; p_1, q_1)$ with a constant $C \le c \cdot \max(C_0, C_1)$ For the case where $r < \infty$ we have

$$\|Tf\|_{q,r} \le C\left\{\int_0^\infty \left[t^{\frac{1}{q}} S_\sigma(f^*(t))\right]^r \frac{dt}{t}\right\}^{\frac{1}{r}}$$

Now applying Minkowski's Inequality we have that

$$\|Tf\|_{q,r} \le C\Bigg[\left\{\int_0^\infty \left[t^{\frac{1}{q}-\frac{1}{q_0}} \int_0^{t^m} s^{\frac{1}{p_0}} f^*(s)\,\frac{ds}{s}\right]^r \frac{dt}{t}\right\}^{\frac{1}{r}} + \left\{\int_0^\infty \left[t^{\frac{1}{q}-\frac{1}{q_1}} \int_{t^m}^\infty s^{\frac{1}{p_1}} f^*(s)\,\frac{ds}{s}\right]^r \frac{dt}{t}\right\}^{\frac{1}{r}}\Bigg].$$

By making the change of variable $u = t^m$ and using the relations for the m above we obtain the inequality

$$\|Tf\|_{q,r} \le C|m|^{-\frac{1}{r}}\Bigg[\left\{\int_0^\infty \left[u^{\frac{1}{p}-\frac{1}{p_0}} \int_0^{u} s^{\frac{1}{p_0}} f^*(s)\,\frac{ds}{s}\right]^r \frac{du}{u}\right\}^{\frac{1}{r}} + \left\{\int_0^\infty \left[u^{\frac{1}{p}-\frac{1}{p_1}} \int_{u}^\infty s^{\frac{1}{p_1}} f^*(s)\,\frac{ds}{s}\right]^r \frac{du}{u}\right\}^{\frac{1}{r}}\Bigg].$$

Now from Hardy's Inequalities 2.23 applied to the inequality above we obtain the following inequality

$$\begin{aligned}\|Tf\|_{q,r} &\le C|m|^{-\frac{1}{r}}\left[c_1\left\{\int_0^\infty \left[u^{\frac{1}{p}} f^*(u)\right]^r \frac{du}{u}\right\}^{\frac{1}{r}} + c_2\left\{\int_0^\infty \left[u^{\frac{1}{p}} f^*(u)\right]^r \frac{du}{u}\right\}^{\frac{1}{r}}\right] \\ &= C|m|^{\frac{-1}{r}}(c_1 + c_2)\|f\|_{p,r},\end{aligned}$$

where

$$\frac{1}{c_1} = \frac{1}{p_0} - \frac{1}{p} = \theta\left(\frac{1}{p_0} - \frac{1}{p_1}\right) \text{ and } \frac{1}{c_2} = \frac{1}{p} - \frac{1}{p_1} = (1-\theta)\left(\frac{1}{p_0} - \frac{1}{p_1}\right).$$

The case where $q = \infty$ is similar so we omit it. □

Remark 2.36. The result still holds for $p_1 = \infty$ if we assume that T satisfies the strong (p_1, q_1) condition.

Indeed assuming that and the fact that the function f_1 of Lemma 2.34 satisfies

$$\|f_1\|_{L^\infty} = f^*(t^m)$$

we obtain that

$$(Tf)^*\left(\frac{t}{2}\right) \leq \left(\frac{t}{2}\right)^{-\frac{1}{q_1}} C_1\|f_1\|_{L^\infty}$$

in place of the one we deduced there. Now following a similar procedure we obtain that T is of joint weak type $(p_0, q_0; p_1, q_1)$. Now by using similar properties the Calderon operator with respect to $(p_0, q_0; p_1, q_1)$ gives us our result.

Note 2.37. The reader who is interested to learn more about the Marcinkiewicz Interpolation theorem can consult [23].

2.6 The L^p norm of the one dimensional uncentered Hardy-Littlewood Maximal Operator on $\mathbb{R}$

As the title implies in this part of the chapter we will determine the exact value of the L^p-norm of the uncentered maximal operator. In particular we will prove that the L^p norm is the unique positive solution of the equation

$$(p-1)x^p - px^{p-1} - 1 = 0.$$

We omit to refer this operator either as the ball or the cube variant of the Hardy-Littlewood maximal operator since both of them coincide in the one dimensional case. The proof presented here is from the paper of Grafakos and Montgomery-Smith "Best Constants for Uncentered Maximal Functions" [16].

Let us start by proving that the equation above has a unique positive solution.

Proof. By setting $g : (0, \infty) \to \mathbb{R}$ with $g(x) := (p-1)x^p - px^{p-1} - 1$ and then differentiating we obtain that

$$g'(x) = p(p-1)(x^{p-1} - x^{p-2}).$$

From this one it stems that the derivative vanishes only at 1 and that it is decreasing on $[0, 1]$ and increasing on $[1, \infty)$. And since the $g(0+) = -1$ and $g(1) = -2$ and the limit of f at infinity is equal to infinity we have that the equation $g(x) = 0$ has only positive solution c_p which in fact is greater than 1. □

So let us state formally the theorem we will prove.

Theorem 2.38. *The L^p norm of the uncentered Hardy-Littlewood Maximal operator $\mathcal{M}_{uc} : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ is equal to the unique positive solution c_p of the equation*

$$(p-1)x^p - px^{p-1} - 1 = 0. \tag{2.5}$$

In order to prove our theorem we are going to need a variation of Riesz's Sunrise Lemma. The proof presented here is the one presented in the book of Hewitt and Stromberg "Real and Abstract Analysis" [19].

Definition 2.39. The left and right maximal operator of an $L^1(\mathbb{R})$ function f are defined by

$$\mathcal{M}_l f(x) := \sup_{a<x} \frac{1}{x-a} \int_a^x |f| \quad \text{and} \quad \mathcal{M}_r f(x) := \sup_{b>x} \frac{1}{b-x} \int_x^b |f|, \quad x \in \mathbb{R}$$

respectively.

Lemma 2.40 (Riesz's Sunrise's Lemma). *Let $f \geq 0$ be a function in $L^1(\mathbb{R})$ and let*

$$M_\lambda^l := \{\mathcal{M}_l f > \lambda\},$$
$$M_\lambda^r := \{\mathcal{M}_r f > \lambda\}.$$

Then the following equalities hold

$$\lambda |M_\lambda^l| = \int_{M_\lambda^l} f\,dx \quad \text{and} \quad \lambda |M_\lambda^r| = \int_{M_\lambda^r} f\,dx. \tag{2.6}$$

Proof. We will prove the equality for the case of M_l since for M_r we use analogous arguments. First let us note that the set

$$M_\lambda^l = \{\mathcal{M}_l f > \lambda\}$$

is open due to the fact that $\int_s^x f\,dy/(x-s)$ is continuous as a function of s for every $x \in \mathbb{R}$. That gives us that for every $x \in M_\lambda^l$ there exists a neighborhood of points of x which is contained in M_λ^l and consequently that M_λ^l is open.

Thus M_λ^l has a unique decomposition as pairwise disjoint open intervals as follows

$$M_\lambda^l = \bigcup_{i=1}^{\infty} (a_i, b_i).$$

Let us note here one of the a_k or b_k might be equal to $-\infty$ or $+\infty$ respectively.

Now we are going to prove that for every $k \in \mathbb{Z}$ and every point $x \in (a_k, b_k)$ there exists a point $s < x$ belonging to (a_k, b_k) such that

$$\frac{1}{x-s} \int_s^x f(y)\,dy > \lambda.$$

If $a_k = -\infty$ it is straightforward that this is true.

So it remains to examine the case where it is $a_k > -\infty$. If it weren't also true in this case then there would be some element $s < a_k$ such that

$$\int_s^x f\,dy > \lambda(x-s)$$

since $\mathcal{M}_l f(x) > \lambda$.

Combining this with the fact that

$$\int_{a_k}^{x} f(y)\,dy \leq \lambda(x - a_k)$$

which is true since $a_k \notin M_\lambda^l$, we obtain that

$$\int_{s}^{a_k} f(y)\,dy = \int_{s}^{x} f(y)\,dy - \int_{a_k}^{x} f(y)\,dy > \lambda(x-s) - \lambda(x-a_k) = \lambda(a_k - s)$$

which implies that $a_k \in M_\lambda^l$ which is a contradiction.

Now let s_x be the infimum of all points $s \in (a_k, b_k)$ such that

$$\frac{1}{x-s}\int_{s}^{x} f > \lambda.$$

We will prove that $s_x = a_k$. Let's assume to the contrary that $s_x > a_k$. Then with a simple limiting process we obtain that $\int_{s_x}^{x} f(y)\,dy \geq \lambda(x - s_x)$. If this inequality was strict then by continuity reasons there would be a $y \in (a_k, s_x)$ such that $\int_y^x f > \lambda(x-y)$ which is a contradiction due to the definition of s_x. Consequently $s_x = a_k$ and

$$\int_{a_k}^{x} f\,dy \geq \lambda(x - a_k).$$

Letting $x \to b_k^-$ we obtain

$$\int_{a_k}^{b_k} f\,dy \geq \lambda(b_k - a_k)$$

in case where $b_k < \infty$.

If either a_k or b_k is equal to $-\infty$ of ∞ then both sides of the above inequality are equal to infinity and so the above inequality is in fact an equality. If not then since $b_k \notin M_\lambda^l$ the reverse inequality holds true and so equality holds in this case also that is

$$\int_{a_k}^{b_k} f = \lambda(b_k - a_k).$$

Now (2.6) for the case of M_λ^l is an immediate consequence of the countable additivity of the Lebesgue measure. □

Remark 2.41. It is true

$$\mathcal{M}_{uc} f = \max\{\mathcal{M}_l f, \mathcal{M}_r f\}.$$

Indeed let a, b such that $a < x < b$. Then we have that

$$\frac{1}{b-a}\int_a^b f\,dy = \frac{1}{b-a}\left(\int_a^x f\,dy + \int_x^b f\,dy\right).$$

By setting $\tau := \frac{x-a}{b-a}$ we obtain the following inequality,

$$\begin{aligned}\frac{1}{b-a}\int_a^b f\,dy &= \frac{\tau}{x-a}\int_a^x f\,dy + \frac{1-\tau}{b-x}\int_x^b f\,dy \\ &\leq \tau\mathcal{M}_l f(x) + (1-\tau)\mathcal{M}_r f(x) \\ &\leq \max\{\mathcal{M}_l f(x), \mathcal{M}_r(x) f\}.\end{aligned}$$

Thus

$$\mathcal{M}_{uc} f(x) \leq \max\{\mathcal{M}_l f(x), \mathcal{M}_r f(x)\}, \quad \text{for every } x \in \mathbb{R}.$$

The opposite direction is direct consequence of the definition of the maximal operators that we treat.

Lemma 2.42. *Let $f \geq 0$ be an $L^1(\mathbb{R})$ function and for $\lambda > 0$ let $A_\lambda := \{Mf > \lambda\}$ and $B_\lambda := \{f > \lambda\}$. Then the following inequality holds*

$$\lambda(|A_\lambda| + |B_\lambda|) \leq \int_{A_\lambda} f\,dx + \int_{B_\lambda} f\,dx. \tag{2.7}$$

Proof. By adding the two equalities (2.6) stated in the Riesz's Sunrise Lemma and using the facts that $A_\lambda = M_\lambda^l \cup M_\lambda^r$ and the elementary property of any measure $|A_\lambda| = |M_\lambda^l| + |M_\lambda^r| - |M_\lambda^l \cap M_\lambda^r|$, we obtain the following equality

$$\begin{aligned}\lambda(|A_\lambda| + |M_\lambda^r \cap M_\lambda^r|) = \lambda(|M_\lambda^l| + |M_\lambda^r|) &= \int_{M_\lambda^l} f(y)\,dy + \int_{M_\lambda^r} f(y)\,dy \\ &= \int_{A_\lambda} f\,dx + \int_{M_\lambda^l \cap M_\lambda^r} f\,dx\end{aligned}$$

where in the last equality we used again the fact that $A_\lambda = M_\lambda^l \cup M_\lambda^r$ and the fact that set function $\nu(A) = \int_A f(x)\,dx$ is a measure on the Lebesgue measurable subsets of $\mathbb{R}$.

Now since $f \leq \lambda$ on $(M_\lambda^r \cap M_\lambda^r) - B_\lambda$,

$$\int_{(M_\lambda^r \cap M_\lambda^l) - B_\lambda} f \leq \lambda|(M_\lambda^l \cap M_\lambda^r) - B_\lambda|. \tag{2.8}$$

Combining this inequality with the previous equality we obtain

$$\lambda(|A_\lambda| + |(M_\lambda^l \cap M_\lambda^r) \cap B_\lambda|) \leq \int_{A_\lambda} f\,dx + \int_{(M_\lambda^l \cap M_\lambda^r) \cap B_\lambda} f\,dx.$$

The set $B_\lambda - (M_\lambda^l \cap M_\lambda^r)$ has zero measure because of the Lebesgue Differentiation Theorem. This last observation leads us to the inequality we sought. □

We are also going to need the following lemma.

Lemma 2.43. *Let f, g be locally integrable functions on $\mathbb{R}$ and $p > 1$, then*

$$\int_0^\infty \lambda^{p-2}\int_{\{g>\lambda\}} f(x)\,dx\,d\lambda = \frac{1}{p-1}\int_{\mathbb{R}} f g^{p-1}\,dx.$$

Its proof is based on applying Fubini's theorem in the same vain as of Theorem 1.12 so we omit it.

Now multiplying inequality 2.7 with λ^{p-2} and integrating over $(0,\infty)$ we obtain by using Lemma 2.43

$$\frac{1}{p}\|\mathcal{M}_{uc}f\|_p^p + \frac{1}{p}\|f\|_p^p \leq \frac{1}{p-1}\|f\|_{L^p}^p + \frac{1}{p-1}\int_{\mathbb{R}} f(x)[\mathcal{M}_{uc}f(x)]^{p-1}\,dx$$

or equivalently

$$(p-1)\|\mathcal{M}_{uc}f\|_p^p - p\int_{\mathbb{R}} f(x)[(\mathcal{M}_{uc}f(x))]^{p-1}\,dx - \|f\|_p^p \leq 0.$$

Now *applying Hölder's inequality* to the integral in the middle above with exponents p and $p/(p-1)$ gives us

$$(p-1)\|\mathcal{M}_{uc}f\|_p^p - p\|f\|_p \cdot \|\mathcal{M}_{uc}f\|_p^{p-1} - \|f\|_p^p \leq 0$$

or that

$$(p-1)\left(\frac{\|\mathcal{M}_{uc}f\|_p}{\|f\|_p}\right)^p - p\left(\frac{\|\mathcal{M}_{uc}f\|_p}{\|f\|_p}\right)^{p-1} - 1 \leq 0.$$

As a consequence

$$\frac{\|\mathcal{M}_{uc}f\|_p}{\|f\|_p} \leq c_p$$

where c_p is described above.

As for the other direction if we trace our steps back we conclude that in order to obtain sharpness of the above stated inequality, we should achieve approximate equalities in inequality (2.8) and in Hölder's inequality which is used in the last step.

Let $f:\mathbb{R}\to[0,\infty]$ with $f(x) := |x|^{-\frac{1}{p}}$. Then we have the following: f is *even symmetrically decreasing and also* $\mathcal{M}f = A_p f$ where A_p stands for the unique positive solution of 2.5.

The first claim is straightforward so we concern ourselves with the second. What we are going to show is that $\mathcal{M}_{uc}f(1) = A_p$. Provided we proved that a change of variables gives us that $\mathcal{M}_{uc}f(x) = \mathcal{M}_{uc}f(1)|x|^{-\frac{1}{p}} = A_p|x|^{-\frac{1}{p}}$ for $x \neq 0$.

As we have shown $\mathcal{M}_{uc}f(1) = \max\{\mathcal{M}_l f(1), \mathcal{M}_r f(1)\}$ and since f is symmetrically decreasing

$$\mathcal{M}_{uc}f(1) = \mathcal{M}_l f(1).$$

So for $a < 1$ let

$$I(a) := \frac{1}{1-a}\int_a^1 f\,dx.$$

Differentiating it with respect to a gives us

$$I'(a) = \frac{-f(a)(1-a) + \int_a^1 f\,dx}{(1-a)^2}.$$

Skipping some calculations we reach the conclusion that the maximum is obtained for $a = -\gamma < 0$ where γ is the solution of the equation

$$-\frac{1}{p}\gamma^{1-\frac{1}{p}} + \left(1 - \frac{1}{p}\right)\gamma^{-\frac{1}{p}} - 1 = 0. \tag{2.9}$$

We then have

$$\mathcal{M}_{uc}f(1) = \frac{1}{1+\gamma}\int_{-\gamma}^{1} f = \frac{1}{1+\gamma}\frac{1+\gamma^{1-\frac{1}{p}}}{1-1/p} = \frac{1}{1+\gamma}\frac{(1-1/p)(\gamma^{1-\frac{1}{p}}+\gamma^{-\frac{1}{p}})}{1-1/p} = \gamma^{-\frac{1}{p}}.$$

where we used the fact that γ is the solution of equation (2.9). Also $\gamma^{-\frac{1}{p}}$ is the unique positive solution of equation (2.5), that is

$$(p-1)\gamma^{-1} - p\gamma^{-1+\frac{1}{p}} - 1 = 0$$

because multiplying by $\gamma^{1-\frac{1}{p}}$ we obtain,

$$(p-1)\gamma^{-\frac{1}{p}} - \gamma^{1-\frac{1}{p}} - p = 0$$

which is true by (2.9).

So f complies with all the criteria we sought except that it is not an L^p function. What we are going to do is estimate f with L^p functions in order to obtain the sharpness we seek. So for $\epsilon > 0$ let

$$f_\epsilon(x) := |x|^{-\frac{1}{p}}\min\{|x|^\epsilon, |x|^{-\epsilon}\}, \quad \text{for } x \neq 0.$$

The functions f_ϵ indeed belong to L^p since

$$\begin{aligned}\int_{\mathbb{R}}|f_\epsilon|^p = 2\int_0^\infty |f_\epsilon|^p = 2\left(\int_0^1 |f_\epsilon|^p + \int_1^\infty |f_\epsilon|^p\right) &= 2\left(\int_0^1 x^{\epsilon p-1} + \int_1^\infty x^{-\epsilon p-1}\right)\\ &= 2\left(\left[\frac{x^{\epsilon p}}{\epsilon p}\right]_0^1 + \left[\frac{x^{-\epsilon p}}{-\epsilon p}\right]_1^\infty\right)\\ &= \frac{4}{\epsilon p} < \infty.\end{aligned}$$

What we are going to show is that

$$\mathcal{M}_{uc}f_\epsilon(x) \geq \left(1 - \frac{1}{p} + \epsilon\right)^{-1}(1+\gamma^{1-\frac{1}{p}+\epsilon})(1+\gamma)^{-1}f_\epsilon(x).$$

Provided that and the fact f_ϵ are L^p we obtain the sharpness by letting $\epsilon \to 0$, using the value of $\mathcal{M}_{uc}f(1)$.

Since f_ϵ is even we only have to prove it for $x > 0$. So if $x > 1/\gamma$, then

$$\begin{aligned}\mathcal{M}_{uc}f_\epsilon(x) &\geq \frac{\int_{-\gamma x}^x f_\epsilon}{x-(-\gamma x)} = \frac{\int_1^x f_\epsilon + 2\int_0^1 f_\epsilon + \int_1^{\gamma x} f_\epsilon}{x+\gamma x}\\ &= \frac{1}{x+\gamma x}\left(\frac{x^{1-\frac{1}{p}+\epsilon}-1}{\left(1-\frac{1}{p}+\epsilon\right)} + 2\frac{1}{\left(1-\frac{1}{p}+\epsilon\right)} + \frac{(\gamma x)^{1-\frac{1}{p}-\epsilon}-1}{\left(1-\frac{1}{p}-\epsilon\right)}\right)\\ &\geq \left(1-\frac{1}{p}+\epsilon\right)^{-1}\left(1+\gamma^{\frac{1}{p'}+\epsilon}\right)(1+\gamma)^{-1}f_\epsilon(x).\end{aligned}$$

If $0 < x \leq \frac{1}{\gamma}$ we follow similar arguments.

Note 2.44. For the best possible weak type inequality regarding the centered Hardy-Littlewood Maximal Operator on $\mathbb{R}$ check [25].

2.7 The Ball variant of the Hardy-Littlewood Maximal Operator defined on Borel measures

Let's take a look at the definition of the Hardy-Littlewood Maximal Operator of any variant i.e.

$$Mf(x) = \sup_{x \in A} \frac{1}{|A|} \int_A f,$$

where A stand for balls or cubes that contain x and might also be centered at x.

As we can see by considering a different measure on $\mathbb{R}^n$ in place of the Lebesgue measure we obtain "a" Hardy-Littlewood Maximal Operator with respect to this new measure.

In this last section of the chapter we are going to investigate the L_p boundedness of that Hardy-Littlewood Maximal Operator defined on balls. The material of this first introductory part is based on the introduction presented in [11].

So if consider a Borel measure μ on $\mathbb{R}^n$, does the L^p boundedness of the Hardy-Littlewood maximal operator with respect to μ still hold. Let us consider the ball variants of the Hardy-Littlewood Maximal Operator which are defined as,

$$\mathcal{M}_c f(x) = \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f \, d\mu$$
$$\mathcal{M} f(x) = \sup_{B \ni x} \frac{1}{\mu(B)} \int_B f \, d\mu$$

using the convention $\frac{0}{0} = 0$ in case there is a a ball of zero measure.

Theorem 2.45 (Weak-L^1 condition for the centered Hardy-Littlewood Maximal Operator)**.** *The Weak-L_1 condition holds for the centered Hardy-Littlewood Maximal Operator without restrictions to the measure.*

For the proof of it we will need the Besicovitch Covering lemma whose proof we postpone until the next chapter and Lindelöf's theorem.

Theorem (**Besicovitch Covering Lemma**)**.** *Let A be a bounded subset of R^n and $\{B_i := B(x_i, r_i)\}_{i \in I}$ a covering of A consisting of balls such that each element of A is the center of at least one of them and their radii are upper bounded (i.e. $\sup_{i \in I} r_i < \infty$) where $r_i > 0$ denotes the radius of the ball B_i)'.*

Then there is a sequence (which can be finite) of the covering that also covers A and that can be distributed into $C(n)$ subsequences, such that the elements of each one of them are pairwise disjoint.

Proof. For $\lambda > 0$ let $E_\lambda := \{\mathcal{M}_c f > \lambda\}$. Then for each $x \in E_\lambda$ there exists a ball with center x $B(x)$ such that

$$\frac{1}{B(x)} \int_{B(x)} f \, d\mu > \lambda.$$

Now by applying the Besicovitch Covering Theorem on that family of balls we obtain C_n families $\mathcal{B}_1, \dots, \mathcal{B}_{C_n}$ each one of them consisting of pairwise disjoint balls such that $\bigcup_{i=1}^{C_n} \mathcal{B}_i \supset E_\lambda$ and so that if $B(x_i)$ an element of them then

$$\frac{1}{\mu(B(x_i))} \int_{B(x_i)} f \, d\mu > \lambda.$$

Consequently

$$\mu(\cup \mathcal{B}_i) \leq \frac{1}{\lambda} \int_{\mathbb{R}^n} |f| \, d\mu, \quad \text{for } i = 1, \dots, C_n$$

and as a result

$$\mu(E_\lambda) \leq \frac{C_n}{\lambda} \|f\|_{L^1}.$$

□

The following proof was extracted from [22].

Theorem 2.46 (Lindelöf's theorem)**.** *For every family $\{G_i\}_{i\in I}$ of open sets in $\mathbb{R}^n$ there exists a countable subfamily $\{G_i\}_{i\in I_0}$ such that*

$$\bigcup_{i\in I} G_i = \bigcup_{i \in I_0} G_i.$$

Proof. Let us consider first the family of all the balls of rational radii and centered around points with rational coordinates. This family is countable and so it can be arranged in a sequence $\{B_m\}_{m\in\mathbb{N}}$. From how these balls were elected we may find for each $m \in \mathbb{N}$ a set G_{i_m} such that

$$B_m \subset G_{i_m}.$$

Now since these balls cover all the sets G_i,i.e.

$$G_i \subset \bigcup_{m\in\mathbb{N}} B_m.$$

the family $\{G_{i_m}\}_{m\in\mathbb{N}}$ is one with the desired properties. □

The proof of this lemma was extracted from [36].

Lemma 2.47. *For every finite sequence of intervals in $\mathbb{R}$ $\{I_j\}_{j=1}^m$, there exist two subsequences $\{I_j\}_{j\in S_1}$ and $\{I_j\}_{j\in S_2}$ such that*

1. *If $I_j, I_k \in S_i$ $(i = 1, 2)$ with $I_j \neq I_k$ we have $I_j \cap I_k = \emptyset$*
2. *$\bigcup_{j=1}^m I_j = \bigcup_{j\in S_1\cup S_2} I_j$.*

Proof. First we will extract a family $\{I_j\}_{j\in P}$ such that

$$\bigcup_{j=1}^{m} I_j = \bigcup_{j\in P} I_j$$

and *every member of it is not contained wholly within the union of the other members* i.e. if I_k is a member of the family $\{I_j\}_{j\in P}$ then

$$I_k \not\subseteq \bigcup_{j\in P: j\neq k} I_j.$$

Beginning with the family $\{I_j\}_{j=1}^m$ itself, if it has this property we stop here. If not we exclude the first interval I_k such that

$$I_k \subset \bigcup_{j=1, j\neq k}^{m} I_j.$$

For this family we have that

$$\bigcup_{j=1, j\neq k}^{m} I_j = \bigcup_{j=1}^{m} I_j.$$

Now if this family's elements satisfy this property we stop here. If not we repeat this process again.

Since the intervals are finite in number this process finishes after a finite number of repetitions and with the end of it we obtain a family $\{I_j\}_{j\in P}$ that satisfies not only the above property but also property (i) of lemma's statement.

Now we order the elements of the family $\{I_j\}_{j\in P}$ as $\{I_{j_l}\}_{l=1}^k$ so that if

$$I_s = (a_s, b_s) \quad \text{and} \quad I_t = (a(t), b(t))$$

are members of the family with $s < t$ then $a_s \leq a_t$.

Using that and the property the intervals of this family have we also obtain that $b_s \leq b_t$ for else we would have $I_s \supset I_t$.

Now we prove that for every two elements I_{j_l} and $I_{j_{l+2}}$ of the sequence

$$I_{j_l} \cap I_{j_{l+2}} = \emptyset.$$

If this didn't hold true then we would have

$$I_{j_l} \cap I_{j_{l+2}} \supset (a_l, b_{l+2}) \supset I_{j_{l+1}}$$

contradicting as before the property this family has. Equivalently we have that the families $\{I_{j_1}, I_{j_3}, \dots\}$ and $\{I_{j_2}, I_{j_4}, \dots\}$ satisfy property (ii). □

The remaining results of this section are based on this paper [14].

Theorem 2.48. *For every $\lambda > 0$ and every non-negative $f \in L^1_{loc}(\mu)(\mathbb{R})$ we have that*

$$\mu(\{\mathcal{M}f > \lambda\}) + \mu(\{f > \lambda\}) \leq \frac{1}{\lambda}\int_{\{\mathcal{M}f>\lambda\}} f\,d\mu + \frac{1}{\lambda}\int_{\{f>\lambda\}} f\,d\mu$$

where $\mathcal{M}$ denotes here the uncentered Hardy-Littlewood maximal operator on $\mathbb{R}$ with respect to μ.

Proof. For $\lambda > 0$ let $E_\lambda := \{\mathcal{M}_B f > \lambda\}$. If $\mu(\{f > \lambda\}) = \infty$ then the right hand side of the inequality is also equal to infinity and so it holds true trivially. So let us assume that $\mu(\{f > \lambda\}) < \infty$. Now for every $x \in E_\lambda$ let I_x be an interval that contains x such that

$$\frac{1}{\mu(I_x)}\int_{I_x} f\,d\mu > \lambda$$

Then by Lindelöf's theorem (Theorem 2.46) there is a countable subcollection $\{I_j\}_{j=1}^{\infty}$ such that

$$\bigcup_{j=1}^{\infty} I_j = \bigcup_{x\in E_\lambda} I_x$$

Let $\mathcal{I}_N := \{I_1, \ldots, I_N\}$ and let

$$F_N := \bigcup_{I\in\mathcal{I}_N} I$$

Then by Lemma 2.47 we split I_N into two subcollections $\mathcal{I}_N^1$ and $\mathcal{I}_N^2$ such that the intervals of each one of these collections are pair-wise disjoint and

$$F_N = \bigcup_{i=1}^{2}\bigcup_{I\in\mathcal{I}_N^i} I.$$

Now if we set $F_N^i := \cup_{I\in\mathcal{I}_N^i} I$ $(i = 1, 2)$ we have that

$$\mu(F_N^i) = \sum_{I\in\mathcal{I}_N^i}\mu(I) < \frac{1}{\lambda}\sum_{I\in\mathcal{I}_N^i}\int_I f\,d\mu = \frac{1}{\lambda}\int_{F_N^i} f\,d\mu, \quad \text{for } i = 1, 2$$

and since $F_n = F_n^1 \cup F_n^2$

$$\begin{aligned}\mu(F_N) + \mu(F_N^1 \cap F_N^2) &= \mu(F_N^1 + \mu(F_N^2)) \\ &< \frac{1}{\lambda}\int_{F_N^1} f\,d\mu + \int_{F_N^2} f\,d\mu \\ &= \frac{1}{\lambda}\int_{F_N} f\,d\mu + \frac{1}{\lambda}\int_{F_N^1\cap F_N^2} f\,d\mu.\end{aligned}$$

Now we will prove that for every μ-measurable set E with $\mu(E) < \infty$

$$\frac{1}{\lambda}\int_E f\,d\mu + \mu(\{f > \lambda\}) \leq \frac{1}{\lambda}\int_{\{f>\lambda\}} f\,d\mu + \mu(E).$$

Indeed

$$\int_E (f-\lambda)\,d\mu = \int_{\{f\le\lambda\}\cap E}(f-\lambda)\,d\mu + \int_{\{f>\lambda\}\cap E}(f-\lambda) \le \int_{\{f>\lambda\}}(f-\lambda)\,d\mu.$$

The inequality to the right above holds since the first integral in the middle is negative.

Now combining these two inequalities gives us

$$\begin{aligned}\mu(F_N)+\mu(F_N^1\cap F_N^2)+\mu(\{f>\lambda\}) &\le \frac{1}{\lambda}\int_{F_N} f\,d\mu + \frac{1}{\lambda}\int_{F_N^1\cap F_N^2} f\,d\mu + \mu\{f>\lambda\}\\ &\le \frac{1}{\lambda}\int_{F_N} f d\mu + \mu(F_N^1\cap F_N^2) + \int_{\{f>\lambda\}} f\,d\mu.\end{aligned}$$

So

$$\mu(F_N)+\mu(\{f>\lambda\}) \le \frac{1}{\lambda}\int_{F_N} f\,d\mu + \frac{1}{\lambda}\int_{\{f>\lambda\}} f\,d\mu.$$

and by letting $N\to\infty$ we have our result. □

Theorem 2.49. *For $1<p<\infty$ let A_p be the unique positive solution of the equation*

$$(p-1)x^p - px^{p-1} - 1 = 0.$$

Then

$$\|\mathcal{M}f\|_{p,\mu} \le A_p\|f\|_{p,\mu}.$$

Proof. Let $f\in L^p(\mathbb{R})$ and without loss of generality let us assume that f is not equal to zero a.e.. Then

$$\begin{aligned}\int_{\mathbb{R}}(\mathcal{M}f)^p\,d\mu + \int_{\mathbb{R}} f^p\,d\mu &= p\int_0^\infty \lambda^{p-1}\mu(\{\mathcal{M}f>\lambda\})\,d\lambda + p\int_0^\infty \lambda^{p-1}\mu(\{f>\lambda\})\,d\lambda\\ &\le p\int_0^\infty \lambda^{p-2}\int_{\{\mathcal{M}f>\lambda\}} f\,d\mu\,d\lambda + p\int_0^\infty \lambda^{p-2}\int_{\{f>\lambda\}} f\,d\mu\,d\lambda\\ &= \frac{p}{p-1}\int_{\mathbb{R}}(\mathcal{M}f)^{p-1}f\,d\mu + \frac{p}{p-1}\int_{\mathbb{R}} f^p\,d\mu.\end{aligned}$$

Thus

$$\int_{\mathbb{R}}(\mathcal{M}f)^p\,d\mu \le \frac{p}{p-1}\int_{\mathbb{R}}(\mathcal{M}f)^{p-1}f\,d\mu + \frac{1}{p-1}\int_{\mathbb{R}} f^p d\mu.$$

Hölder's inequality now gives us

$$\int_{\mathbb{R}}(\mathcal{M}f)^{p-1}f\,d\mu \le \left(\int_{\mathbb{R}}(\mathcal{M}f)^p\,d\mu\right)^{\frac{p-1}{p}}\left(\int_{\mathbb{R}} f^p\,d\mu\right)^{\frac{1}{p}}$$

so

$$(p-1)\|\mathcal{M}f\|_{L^p(\mu)}^p \le p\|\mathcal{M}f\|_{L^p(\mu)}^{p-1}\|f\|_{L^p(\mu)} + \|f\|_{L^p(\mu)}^p.$$

Equivalently

$$(p-1)\left(\frac{\|\mathcal{M}f\|_{L^p(\mu)}}{\|f\|_{L^p(\mu)}}\right)^p - p\left(\frac{\|\mathcal{M}f\|_{L^p(\mu)}}{\|f\|_{L^p(\mu)}}\right)^{p-1} - 1 \le 0.$$ □

Note 2.50. As for the uncentered one this example below shows us that it generally doesn't hold for $n > 1$.

Example 2.51. Let $1 < p < \infty$ and $B_1, B_2, \dots$ be closed balls so that 0 belongs to the boundary of B_i for $i = 1, 2, \dots$ and for each i there exists a point $x_i \in B_i$ with

$$x_i \notin \bigcup_{j \neq i} B_j.$$

and set

$$\mu = \sum_{i=1}^{\infty} \delta_{x_i}{}^{1}$$

where we define $x_0 := 0$.

If $f = \chi_{B_1}$ then $\|f\|_{L^p(\mu)} = \left(\int_{B_1} 1 \, d\mu = \right)^{\frac{1}{p}} = 2^{\frac{1}{p}}$, but

$$\mathcal{M}f(x_i) \geq \frac{1}{\mu(B_i)} \int_{B_i} f \, d\mu = \frac{1}{2}, \quad \text{for } i = 1, 2, \dots$$

and so

$$\|\mathcal{M}f\|_{L^p(\mu)} = \infty.$$

Theorem 2.52. *Let $\lambda > 0$ and $f : \mathbb{R}^n \to [0, \infty]$ be an $L^1_{loc}(\mu)$ function then*

$$\mu(\{\mathcal{M}_c f > \lambda\}) + (C_n - 1)\mu(\{f > \lambda\}) \leq \frac{1}{\lambda} \int_{\{\mathcal{M}_c f > \lambda\}} f \, d\mu + (C_n - 1)\frac{1}{\lambda} \int_{\{f > \lambda\}} f \, d\mu,$$

where C_n stands for the Besicovitch constant and $\mathcal{M}_c$ stands for the centered ball variant of the Hardy-Littlewood Maximal Operator with respect to the measure μ.

Proof. For $\lambda > 0$ let $E_\lambda := \{\mathcal{M}_c f > \lambda\}$. Without loss of generality let us assume that $\mu(E_\lambda) < \infty$ for else the weak-L_1 condition of $\mathcal{M}$ guarantees us that the right side is equal to infinity.

For every $x \in \{\mathcal{M}f > \lambda\}$ let $B(x, r_x)$ be a ball such that

$$\frac{1}{\mu(B(x, r_x))} \int_{B(x, r_x)} f \, d\mu > \lambda.$$

Also

$$\begin{aligned} \int_{B(x,r_x)} f \, d\mu &= \int_{B(x,r_x) \cap (\mathbb{R}^n \setminus E_\lambda)} f \, d\mu + \int_{B(x,r_x) \cap E_\lambda} f \, d\mu \\ &\leq \lambda \mu(B(x, r_x) \cap (\mathbb{R}^n \setminus E_\lambda)) + \int_{B(x,r_x) \cap E_\lambda} f \, d\mu. \end{aligned}$$

[1] δ_{x_i} stands for the Dirac measure at x_i which is defined to be equal to one on every set that contains x and zero for the rest of them.

Combining these two inequalities gives us

$$\int_{B(x,r_x)\cap E_\lambda} f\,d\mu > \lambda\mu(B(x,r_x)\cap E_\lambda)$$

For an $R > 0$ fixed, let $B_R := B(0,R)$ and set $\mathcal{B} := \{B(x,r_x) : x \in B_R \cap E_\lambda\}$. Besicovitch's Covering Theorem gives us subfamilies $\mathcal{B}_1, \ldots, \mathcal{B}_{C_n}$ such that elements of each of them are pairwise disjoint and so that

$$B_R \cap E_\lambda \subset \bigcup_{i=1}^{C_n} \bigcup_{B\in\mathcal{B}_i} B.$$

Now let us set $F_i := \cup_{B\in\mathcal{B}_i} B$ for $i = 1, \ldots, C_n$ and $F := \cup_{i=1}^{C_n} F_i$. Since the elements of each $\mathcal{B}_i$ are pairwise disjoint we obtain that

$$\begin{aligned}\mu(F_i \cap E_\lambda) = \sum_{B\in\mathcal{B}_i} \mu(B\cap E_\lambda) &< \sum_{B\in\mathcal{B}_i} \frac{1}{\lambda}\int_{B\cap E_\lambda} f\,d\mu\\ &= \frac{1}{\lambda}\int_{F_i\cap E_\lambda} f\,d\mu, \quad \text{for } i = 1, \ldots, C_n.\end{aligned}$$ □

Corollary 2.53. *Let $A_{p,n}$ be the unique positive solution of the equation*

$$(p-1)x^p - px^{p-1} - (C_n - 1) = 0$$

then

$$\|\mathcal{M}f\|_{p,\mu} \le A_{p,n}\|f\|_{p,\mu}.$$

Since its proof is similar to that of Corollary 2.49 we omit it.

CHAPTER 3

Covering Lemmas

In this section we will state and prove some covering lemmas that are used in Harmonic Analysis. Some of them depend on the particular measure of the space while others depend only on the geometric properties of the space such as the first covering lemma we concern ourselves with. For the source of the material one can see [8] and [4].

3.1 Besicovitch Covering Lemma

3.1.1 Main variant

Theorem 3.1 (**Besicovitch Covering Lemma**). *Let A be a bounded subset of $\mathbb{R}^n$ and $\{B_i := B(x_i, r_i)\}_{i\in I}$ a covering of A consisting of balls such that each element of A is the center of at least one of them and their radii are upper bounded (i.e. $\sup_{i\in I} r_i < \infty$) where $r_i > 0$ denotes the radius of the ball B_i).*

Then there is a sequence (which can be finite) of the covering that also covers A and that can be distributed into $C(n)$[1] subsequences, such that the elements of each one of those are pairwise disjoint.

Proof. First let us find a sequence of balls of the initial covering that also covers A.

- First we select a ball $B_1 := B(x_1, r_1)$ such that

$$r_1 \geq \frac{3}{4} \sup_{i\in I} r_i.$$

If B_1 covers the whole set A then we are done.

- If not we set $\{B_i\}_{i\in I_1}$ to be the family of balls of B_i whose centers don't lie in B_1 and select a ball $B_2 = B(x_2, r_2)$ out of it such that

$$r_2 \geq \frac{3}{4} \sup_{i\in I_1} r_i.$$

If B_2 and B_1 constitute a covering of A we are finished.

[1] The least number $C(n)$ depending on n is called the Besicovitch Constant. For more information consult [37],[12].

- If not we set $\{B_i\}_{i\in I_2}$ to be the family of balls of $\{B_i\}_{i\in I}$ whose centers don't lie in either B_1 or B_2 and select a ball $B_3 = B(x_3, r_3)$ out of it such as

$$r_3 \geq \frac{3}{4} \sup_{i\in I_2} r_i.$$

If now all three of them cover A the procedure ends. If not we repeat the process.

If the process finishes at a finite number of steps it means that the sets that we elected during the process cover A, thus producing a finite covering of A.

If not the sequence $\{B_i\}_{i\in\mathbb{N}}$ we obtain in this case also constitutes a covering of A. Indeed if we assume the contrary there exists an element $x \in A$ that doesn't belong to any of the B_i that we elected and since $\{B_i\}_{i\in I}$ is a covering of A there exists a ball $B(x, r_x)$ centered at x where $r_x > 0$ such that

$$0 < r_x \leq r_i, \quad \text{for every } i \in \mathbb{N}.$$

But this sequence tends to zero since for $i > j$ we have that $x_i \notin B_j$ and so

$$|x_i - x_j| > r_j = \frac{2}{3}r_j + \frac{1}{3}r_j > \frac{1}{3}r_i + \frac{1}{3}r_j.$$

From that we obtain that the balls $1/3B_i := B(x_i, r_i/3)$ are pairwise disjoint and also since A is bounded these balls are all contained in a *bounded* set B_0. Consequently $r_i \to 0$ thus reaching a contradiction.

So in any case the sequence constitutes a covering of A.

The next claim that we will prove is that for any ball B_m, $m = 1, 2, \ldots$ the number of the previously elected balls that intersect it is bounded by a constant term C_n which depends only on the dimension n, i.e.

$$\#\{1 \leq i \leq m : B_m \cap B_i \neq \emptyset\} < C(n), \quad \text{for every } m \geq 1.$$

To begin with we set $In_m := \{1 \leq i \leq m : B_m \cap B_i \neq \emptyset\}$. We will partition this set into two subsets and then prove that for each one's cardinality has a upper bound depending only on the dimension n.

So we consider the sets,

$$In_m^1 := In_m \cap \{1 \leq i \leq m : r_i \leq 3r_m\} \quad \text{and} \quad In_m^2 := In_m \cap \{1 \leq i \leq m : r_i > 3r_m\}.$$

1. *First we find an upper bound for In_m^1.*

   Let $i \in In_1$, then we have that $1/3B_i$ is contained $B(x_m, 5r_m)$. Indeed since the balls B_i and B_m have non-empty intersection, there is an element y that belongs to both of them and so for every $x \in 1/3B_i$

$$|x - x_m| \leq |x - x_i| + |x_i - x_m| \leq |x - x_i| + |x_i - y| + |y - x_m| \leq \frac{r_i}{3} + r_i + r_m \leq 5r_m.$$

Denoting now by v_n the volume of the unit ball in $\mathbb{R}^n$ we obtain,

$$|5B(x_m, r_m)| \geq \sum_{i \in In_m^1} \left| B\left(x_i, \frac{r_i}{3}\right) \right| = \sum_{i \in In_m^1} v_n r_i^n 3^{-n}$$
$$\geq \sum_{i \in In_m^1} v_n r_m^n 4^{-n} = \# Is_m^1 v_n r_m^n 4^{-n}.$$

Now since the left side is equal to $v_n 5^n r_m^n$. With that we obtain

$$\# Is_m^1 \leq 20^n.$$

Now we consider the set In_m^2.

We first consider two elements $i, j \in In_m^2$ *thus we have* $r_i > 3r_m$, $r_j > 3r_m$ *and* $B_i \cap B_m \neq \emptyset$, $B_j \cap B_m \neq \emptyset$ *and* without loss of generality let us also *assume that* $x_m = 0$.

We are going to show that the angle between the respective x_i, x_j is greater or equal than some positive angle θ_0, i.e. if we set, $\theta := \widehat{x_i x_j}$ then,

$$\theta \geq \theta_0 := \arccos \frac{61}{64}.$$

Provided we proved this claim ,we then consider the unit sphere S^{n-1} of $\mathbb{R}^n$ and a positive number r small enough such that if $x \in S^{n-1}$, the angle between every two points in $B(x, r)$ is smaller than θ_0. Due to the compactness of the unit sphere there is a finite number M_n of these balls that also covers S^{n-1}. Since the rays generated from the x_i passing through the origin form the same angle as the x_i, they can't pass from the same ball. Consequently there are at most M_n of them and so

$$\# In_2^m \leq M_n.$$

Now we prove that there exists such an angle θ_0.

Remembering our assumption $x_m = 0$ we obtain that $0 \notin B_i \cup B_j$. We also have that the intersection of B_i and B_j with B_m is non empty. So

(a) $|x_i| > r_i$

(b) $|x_j| > r_j$

(c) $|x_i| \leq r_i + r_m$

(d) $|x_j| \leq r_j + r_m$

and let us also assume that $|x_i| \leq |x_j|$.

We will prove that if $\cos\theta > \frac{5}{6}$ then $x_i \in B_j$.

- In case $x_i \notin B_j$ and also $|x_i - x_j| > |x_j| > r_j$ we have,

$$\cos\theta = \frac{|x_i|^2 + |x_j|^2 - |x_i - x_j|^2}{2|x_i||x_j|} \leq \frac{|x_i|}{2|x_j|} \leq \frac{1}{2} < \frac{5}{6}.$$

- As for the case where $|x_i - x_j| \leq |x_j|$ and $x_i \notin B_j$ or equivalently $|x_i - x_j| > r_j$ we have that,

$$\begin{aligned}\cos\theta &= \frac{|x_i|^2 + |x_j|^2 - |x_i - x_j|^2}{2|x_i||x_j|} \\ &= \frac{|x_i|}{2|x_j|} + \frac{(|x_j| - |x_i - x_j|)(|x_j| + |x_i - x_j|)}{2|x_i||x_j|} \\ &\leq \frac{|x_i|}{2|x_j|} + \frac{(|x_j| - |x_i - x_j|)2|x_j|}{2|x_i||x_j|} \\ &\leq \frac{1}{2} + \frac{r_j + r_m - r_j}{r_i} \leq \frac{5}{6}.\end{aligned}$$

  So we proved our result.

We will now prove that

$$0 \leq |x_i - x_j| + |x_i| - |x_j| \leq \frac{8}{3}(1 - \cos\theta)|x_j|, \quad \text{if } x_i \in B_j. \tag{3.1}$$

Since $x_i \in B_j$, we have that $i < j$ and so $x_j \notin B_i$ i.e. $|x_i - x_j| > r_i$ and so

$$\begin{aligned}0 \leq \frac{|x_i - x_j| + |x_i| - |x_j|}{|x_j|} &\leq \frac{|x_i - x_j| + |x_i| - |x_j|}{|x_j|}\frac{|x_i - x_j| - |x_i| + |x_j|}{|x_i - x_j|} \\ &= \frac{|x_i - x_j|^2 - (|x_j| - |x_i|)^2}{|x_j||x_i - x_j|} = \frac{-2x_i \cdot x_j + 2|x_i||x_j|}{|x_j||x_i - x_j|} \\ &= \frac{2|x_i|(1 - \cos\theta)}{|x_i - x_j|} \leq \frac{2(r_i + r_m)(1 - \cos\theta)}{r_i} \\ &\leq \frac{8}{3}(1 - \cos\theta).\end{aligned}$$

For last inequality we used the fact that $r_m < 1/3r_i$ since $i \in In_m^2$.

We will now examine the cases $\cos\theta < \frac{5}{6}$ *and* $\cos\theta \geq \frac{5}{6}$ *separately.* In case $\cos\theta < \frac{5}{6}$ it is immediate that $\cos\theta \leq \frac{61}{64}$.

As for the case where $\cos\theta > 5/6$, from what we have proved before we have that $x_i \in B_j$ and since $i < j$, $x_j \notin B_i$. As a result $r_i < |x_i - x_j| \leq r_j$. Combining them with the facts that $r_j \leq \frac{4}{3}r_i$ and $r_j > 3r_m$, we then obtain that

$$|x_i - x_j| + |x_i| - |x_j| \geq r_i + r_i - r_j - r_m \geq \frac{r_j}{2} - r_m \geq \frac{1}{8}(r_j + r_m) \geq \frac{1}{8}|x_j| \tag{3.2}$$

At last combining (3.1) and (3.2) gives us

$$\frac{1}{8}|x_j| \leq \frac{8}{3}(1 - \cos\theta)|x_j|$$

and so $\cos\theta \leq \frac{61}{64}$ in this case also.

For the last step of the proof we will split the sequence into subsequences in a way that the elements of each one of them are pairwise disjoint. Before we do that let us set $C(n) := 20^n + M_n + 1$.

1. First we separate each of the first $C(n)$ terms of the sequence as follows,

$$\{B_1\}, \{B_2\}, \ldots, \{B_{C(n)}\}.$$

2. Next we select the ball $B_{C(n)+1}$. Since it intersects at most $C(n)-1$ of balls that were elected in the previous steps there exists at least one such ball that doesn't intersect $B_{C(n)+1}$ and so we pair it together with it. Let us assume for instance that $B_{C(n)+1} \cap B_1 = \emptyset$, then we obtain the following collection of sets

$$\{B_1, B_{C(n)+1}\}, \{B_2\}, \ldots, \{B_{C(n)}\}.$$

3. Now suppose we have repeated this procedure until the term $m \geq C(n)+1$. At this point we have the following finite collections,

$$\{B_1, B_{C(n)+1}, \ldots\}, \{B_2, \ldots\}, \ldots, \{B_{C(n)}, \ldots\}.$$

   Since B_{m+1} intersects at most $C(n)-1$ of the previous balls there exists at least one collection whose elements are disjoint from B_m.

4. Proceeding inductively we obtain the desired subsequences. □

Remark 3.2. Note that we have only used the boundedness of the set above in order to prove that the sequence of the radii tends to zero. Now we will prove that a similar result holds even when the set A is unbounded.

Theorem 3.3 (Generalization of the Besicovitch Covering Lemma for Unbounded Sets)**.** *Let A be a subset of $\mathbb{R}^n$ and $\{B_i\}_{i\in I}$ be a covering of A compromised of balls such that each element of A is the center of at least one of them and such that their radii are bounded from above (i.e. $\sup_{i\in I} r_i < \infty$ where r_i denotes the radius of the ball B_i).*

Then there is a sequence (finite or infinite) of these balls that covers A which can distributed into $\tilde{C}_n$ subsequences such that any two balls in each of these subsequences are pairwise disjoint.

Proof. Since we have already shown the result above for bounded sets we assume that A is unbounded. We set $r_0 := sup_{i\in I} r_i$ and consider the following subsets of A,

$$A_m := A \cap \{x \in \mathbb{R}^n : 2r_0(m-1) \leq |x| < 2r_0 m\}.$$

Each of A_m is bounded and so applying the Besicovitch Covering Lemma we obtain C_n subfamilies $\mathcal{B}_m^1, \ldots, \mathcal{B}_m^{C_n}$ that consist of pairwise disjoint balls of the initial cover of A such that

$$A_m \subset \bigcup_{i=1}^{C_n} \bigcup \mathcal{B}_m^i.$$

Now for each $j = 1, \ldots, C_n$ each ball of the families $\mathcal{B}_m^j$ is disjoint to any of the balls of $\mathcal{B}_{m+2}^j$ for every m. So by considering the following families of balls

$$\mathcal{B}_{odd}^1 := \bigcup_{m:odd} \mathcal{B}_m^1, \ldots, \mathcal{B}_{odd}^{C_n} := \bigcup_{m:odd} \mathcal{B}_m^{C_n}$$

$$\mathcal{B}^1_{even} := \bigcup_{m:even} \mathcal{B}^1_m, \ldots, \mathcal{B}^{C_n}_{even} := \bigcup_{m:even} B^{C_n}_m$$

we obtain $\tilde{C}_n := 2C_n$ families of balls from $\{B_i\}_{i\in I}$ that have the desired properties. □

3.1.2 Other variants

Theorem 3.4. *Let $A \subset \mathbb{R}^n$ be a bounded set and for every element x of A let r_x be a positive number that is less than one. Then there exists a countable collection of pairwise disjoint balls $\{B(x_i, r_{x_i})\}_{i\in I}$ where the x_i are points of A such that*

$$A \subset \bigcup_{i\in I} B(x_i, 3r_{x_i}).$$

Proof. Let us first extract a sequence out of $\{B_i\}_{i\in I}$ that will also constitute a covering of A.

1. We first restrict ourselves to the balls their radii lie in $(1/2, 1]$ and extract a maximal family $\mathcal{F}_1$ of disjoint balls out of them. Since A is bounded $\mathcal{F}_1$ cannot contain an infinite number of pairwise disjoint balls of such radii. If $\mathcal{F}_1$ covers A the procedure stops.

2. If not we proceed to balls whose radii lie in $(\frac{1}{4}, \frac{1}{2}]$ and extract a maximal family of disjoint balls that are also disjoint from $\cup\mathcal{F}_1$. For the same reason mentioned in the previous steps $\mathcal{F}_2$ is finite. If now $\mathcal{F}_1 \cup \mathcal{F}_2$ covers the whole set. If now $\mathcal{F}_1$ and $\mathcal{F}_2$ constitute a covering of A and the procedure ends.

3. Suppose now we have constructed the families $\mathcal{F}_1, \ldots, \mathcal{F}_m$ where $m \geq 2$. If $\cup_{i=1}^m \mathcal{F}_i$ constitutes a covering of A the proof is complete. If not we extract a (possibly empty) maximal disjoint family $\mathcal{F}_{m+1}$ such that $\cup\mathcal{F}_{m+1} \cap (\cup_{j=1}^m \mathcal{F}_|) = \emptyset$ and proceed accordingly.

This procedure has two possible outcomes. The first is that it ends after a finite number of repetitions which means that there is an $m \in \mathbb{N}$ such that $\mathcal{F}_1 \cup \mathcal{F}_2 \ldots \mathcal{F}_m$ covers A and also consists of pairwise disjoint sets. Thus the desired result is true in this case.

Otherwise we obtain a sequence

$$B(x_1, r_{x_1}), B(x_2, r_{x_2}), \ldots$$

of pairwise disjoint balls and *it remains to show that* $A \subset \cup_{i=1}^m B(x_i, 3r_{x_i})$. Let $x \in A$, if now $B(x, r_x)$ is one of the balls we elected then we are finished. If not then since $r_x \in (1/2^k, 1/2^{k-1}]$ for some $k \in \mathbb{N}$, there exists a $B(x_i, r_{x_i})$ in $\mathcal{F}_k$ such that $B(x, r_x) \cap B(x_i, r_{x_i}) \neq \emptyset$ due to the maximality of $\mathcal{F}_k$. Also since both r_x and r_{x_i} belong to the same range $(1/2^k, 1/2^{k-1}]$ we have that

$$2r_{x_i} \geq r_x.$$

and so

$$|x - x_i| \leq r_x + r_{x_i} \leq 3r_{x_i}.$$

which gives

$$x \in B(x_i, 3r_{x_i}).$$

Thus the family $B(x_i, r_{x_i}) = \cup_m F_m$ constitutes a cover of A. □

For the next covering lemma we first need to introduce two new notions.

Definition 3.5. A *Radon measure* on $\mathbb{R}^n$ is a Borel measure which is finite on the bounded Borel measurable subsets of $\mathbb{R}^n$.

Definition 3.6. Let A be subset of $\mathbb{R}^n$. A fine Besicovitch covering $\mathcal{F}$ of A is a family of balls that covers A and such that for every $\epsilon > 0$ and $x \in A$ there exists a ball $B(x, r_x)$ of $\mathcal{F}$ with $r_x < \epsilon$.

Theorem 3.7 (Measure-Theoretical Covering Theorem)**.** *Let A be a bounded subset of $\mathbb{R}^n$ and $\mathcal{F}$ a fine Besicovitch covering of A. If μ be a Radon measure on $\mathbb{R}^n$ and μ_e its associated outer measure defined as*

$$\mu_e(E) := \inf\{\mu(\mathcal{O}),\ \mathcal{O} \text{ open and } E \subset \mathcal{O}\}, \quad \text{for every } E \subset \mathbb{R}^n.$$

then there exists a countable collection $\{B_i\}_{i\in I}$ of disjoint balls of $\mathcal{F}$, such that

$$\mu_e\left(A \backslash \bigcup_{i\in I} B_i\right) = 0,$$

i.e. $\{B_i\}_{i\in I}$ forms a measure-theoretical covering of A.

Proof. Without loss of generality let us assume that $\mu_e(A) > 0$ for else there is nothing to prove. Also since A is bounded and $\mathcal{F}$ contains balls centered around any point of A of arbitrary small radius we may also assume without loss of the generality that A and all the elements of $\mathcal{F}$ are contained in some larger ball B_0.

Applying the Besicovitch Covering Theorem to the set A with respect to the covering $\mathcal{F}$, we obtain C_n subsequences of elements of $\mathcal{F}$, $\{\mathcal{B}_1, \ldots, \mathcal{B}_{C_n}\}$ that are consisted of pairwise disjoint balls and such that

$$A \subset \bigcup_{i=1}^{C_n} \bigcup_{j\in\mathcal{I}_i} \mathcal{B}_j^i$$

where $\mathcal{B}_i = \{B_j^i\}_{j\in\mathcal{I}_i}$ for $i = 1, \ldots, C_n$, where I_i is either a initial interval of $\mathbb{N}$ or $I_i = \mathbb{N}$ for every $i \in \mathcal{I}$.

As a result

$$\mu_e\left(A \cap \bigcup_{i=1}^{C_n} \bigcup_{j\in\mathcal{I}_i} B_j^i\right) = \mu_e(A) > 0,$$

so there exists some $i_0 \in \{1, \ldots, C_n\}$ such that

$$\mu_e\left(A \bigcap \bigcup_{j \in \mathcal{I}_{i_0}} B_j^{i_0}\right) \geq \frac{1}{C_n}\mu_e(A).$$

The set $A \cap \cup_{j \in \mathcal{I}_0} B_j^{i_0}$ is of finite measure since the $B_j^{i_0}$ are pairwise disjoint and are all contained in the ball B_0. Consequently

$$\mu_e\left(A \cap \bigcup_{j \in \mathcal{I}_0} B_j^{i_0}\right) \leq \sum_{j \in \mathcal{I}_{i_0}} \mu(B_j^{i_0}) \leq \mu(B_0) < \infty.$$

Thus there exists an $j_0 \in \mathbb{N}$ such that

$$\mu_e\left(A \cap \bigcup_{j=1}^{j_0} B_j^{i_0}\right) \geq \frac{1}{2C_n}\mu_c(A)$$

and since this finite covering is a μ-measurable set we have that from Caratheodory's criterion that

$$\begin{aligned}\mu_e(A) = \mu_e\left(A \bigcap \bigcup_{j=1}^{j_0} B_j^{i_0}\right) + \mu_e\left(A \backslash \bigcup_{j=1}^{j_0} B_j^{i_0}\right) \\ \geq \frac{1}{2C_n}\mu_e(A) + \mu_e\left(A \backslash \bigcup_{j=1}^{j_0} B_j^{i_0}\right).\end{aligned}$$

so by setting $k := 1 - \frac{1}{2C_n} \in (0,1)$ we have that

$$\mu_e\left(A \backslash \bigcup_{j=1}^{j_0} B_j^{i_0}\right) \leq k\mu_e(A).$$

If now the set $A_1 := A \backslash \bigcup_{j=1}^{j_0} B_j^{i_0}$ has zero measure zero we are finished. If not let $\mathcal{F}_1$ denote the family of balls of $\mathcal{F}$ that do not intersect any of the finite balls that have already been constructed. Since $\mathcal{F}_1$ is a fine Besicovitch of A_1 we obtain a finite collection of balls $\{B_i\}_{i=1}^m$ such that

$$\mu_e\left(A_1 \backslash \bigcup_{i=1}^{m} B_i\right) \leq k\mu(A_1) \leq k^2\mu(A).$$

If $A_2 = A_1 \backslash \bigcup_{i=1}^m B_i$ has now measure equal to zero we are finished by considering the family. Otherwise we proceed to the next step accordingly.

If at some repetition the remaining set has zero measure then the proof is complete by selecting all the balls that were chosen in the previous steps. Otherwise we obtain a sequence of balls $\{B_1, B_2, \dots\}$ out of $\mathcal{F}$ such that if s_l is the number of balls that we elected until the l-step then

$$\mu_e\left(A \setminus \bigcup_{i=1}^{s_l} B_i\right) \leq k^l \mu(A).$$

So by letting $l \to \infty$ we get our result. □

3.2 Vitali Covering Lemma

A fine Vitali Covering $\mathcal{F}$ of a subset A of $\mathbb{R}^n$ is a collection of closed cubes satisfying the following condition. For every $\epsilon > 0$ and $x \in A$ there exists a cube Q of $\mathcal{F}$ that contains x and such that $\operatorname{diam}(Q) < \epsilon$. Let us note here that the cubes Q are not necessarily centered around x. This is a most general form of the fine covering which is weaker that the one stated in Definition 3.6.

Note 3.8. A useful result that we will needing during the proof of this particular lemma is the following. If Q is a cube the sides of which have length $l > 0$ then for its diameter we have that

$$\operatorname{diam}(Q) = \sqrt{n}l.$$

Theorem 3.9. *Let A be a bounded, Lebesgue measurable subset in $\mathbb{R}^n$ and $\mathcal{F}$ be a fine Vitali covering of A. Then there is a sequence (finite or infinite) $\{Q_i\}_{i\in I}$ of cubes of $\mathcal{F}$ with pairwise disjoint interiors such that*

$$\left|A \setminus \bigcup_{i\in I} Q_i\right| = 0$$

Proof. Without loss of generality let us assume that A together with the cubes of $\mathcal{F}$ are included in a larger cube Q. We proceed to the following inductive procedure.

1. Let Q_0 be a cube of $\mathcal{F}$. If it covers the whole set then we are finished.

2. If not then we denote by $\mathcal{F}_1$ the family of cubes of $\mathcal{F}$ that have disjoint interiors from Q_0 and denote the supremum of their diameters as d_1.

   Now we select a cube Q_1 from $\mathcal{F}_1$ such that $\operatorname{diam}(Q_1) > \frac{1}{2}d_1$. If $Q_0 \cup Q_1$ covers the set A we are finished.

3. If not we denote by $\mathcal{F}_3$ the family of cubes belonging to $\mathcal{F}$ such us their interior are disjoint from Q_0 and Q_1 and denoting d_3 the supremum of the diameters of the cubes of $\mathcal{F}_3$ we select a cube $Q_3 \in \mathcal{F}_3$ such that

   $$\operatorname{diam}(Q_3) > \frac{1}{2}d_3.$$

   If now all three of these cubes cover A then the procedure ends. If not we proceed accordingly.

If the procedure above ends at some step it means that we have reached a finite family of cubes of $\mathcal{F}$ that covers A and the proof ends. Otherwise we obtain a sequence of cubes let us call it $\{Q_i\}_{i=1}^{\infty}$.

Now we just have to prove that the sequence of cubes $\{Q_i\}_{i=1}^{\infty}$ satisfies the statement of the theorem.

We first observe that

$$\lim_{i\to\infty} \text{diam}(Q_i) = 0.$$

Indeed by Note 3.8we have that for every cube Q

$$\text{diam(Q)} = \sqrt{n} l(Q)$$

where $l(Q)$ denotes the length of the sides of Q and since

$$\sum_{i=1}^{\infty} \left(\frac{(\text{diam} Q_i)}{\sqrt{n}}\right)^n = \sum_{i=1}^{\infty} |Q_i| < \infty$$

the result follows immediately from the convergence of the series.

We now prove that $\{Q_i\}_{i=1}^{\infty}$ forms a measure theoretical covering of A. Let us assume the contrary that is

$$\left| A \setminus \bigcup_{i=1}^{\infty} Q_i \right| \geq 2\epsilon$$

for some $\epsilon > 0$.

Now for each Q_i we consider the cube Q_i' that has the same center as Q_i with faces parallel to the faces of Q_i and such that

$$\text{diam} Q_i' = 5 \ \text{diam} Q_i.$$

Since

$$\sum_{i=1}^{\infty} |Q_i'| < \infty$$

there exists some i_ϵ such that

$$\left| \bigcup_{i=i_\epsilon+1}^{\infty} Q_i' \right| \leq \sum_{i=i_{\epsilon+1}}^{\infty} |Q_i'| \leq \epsilon$$

Therefore we get

$$\left| (A \setminus \bigcup_{i=1}^{i_\epsilon} Q_i) \setminus \bigcup_{i=i_\epsilon+1}^{\infty} Q_i' \right| \geq \left| A \setminus \bigcup_{i=1}^{\infty} Q_i \right| - \left| \bigcup_{i=i_\epsilon+1}^{\infty} Q_i' \right| \geq \epsilon$$

and so there exists an x such that

$$x \in \left(A \setminus \bigcup_{i=1}^{\infty} Q_i \right) \setminus \bigcup_{i=i_\epsilon+1}^{\infty} Q_i'.$$

Also x has positive distance 2σ from the union of the first i_ϵ closed cubes since they are finite in number.

From the definition of the fine Vitali covering $\mathcal{F}$ we can find for a $0 < \delta < \sigma$ a cube Q_δ that has x at its interior and such that $\text{diam}(Q_\delta) < \delta$. By construction Q_δ doesn't intersect any of the first i_ϵ cubes and as a result $Q_\delta \in \mathcal{F}_{i_{\epsilon+1}}$.

Our next claim is that

$$Q_\delta \cap Q_m^o \neq \emptyset, \quad \text{for some } m = i_\epsilon + 1, i_\epsilon + 2, \dots .$$

If this isn't true then we should have that

$$0 < \delta = \text{diam}(Q_\delta) \leq d_m, \quad \text{for every } m \geq i_\epsilon + 1$$

where $d_m \to 0$ as $m \to \infty$ and with that we reached a contradiction.

Let $m_0 \geq i_\epsilon + 1$ be the least natural number such that $Q_\delta \cap Q_{m_0}^o \neq \emptyset$, so we must have

$$Q_\delta \notin \mathcal{F}_{m_0+1}, \quad Q_\delta \in \mathcal{F}_{m_0}$$

and by that we also have $\delta \leq d_{m_0}$.

Now since x doesn't belong to Q'_{m_0} we claim that

$$d_{m_0} \geq \delta = \text{diam}(Q_\delta) > 2\ \text{diam}(Q_{m_0}) > d_{m_0}.$$

To prove the second inequality we first concentrate on the case $n = 1$. Consequently Q_{m_0} is of the form

$$Q_{m_0} = \left[x - \frac{l}{2}, x + \frac{l}{2}\right]$$

where $l > 0$ and

$$Q'_{m_0} = \left[x - \frac{5}{2}l, x + \frac{5}{2}l\right].$$

Since Q_δ intersects Q_{m_0} and contains x where $x \in Q'_{m_0}$ its diameter must be greater than $2l = 2\ \text{diam}(Q_{m_0})$.

Now for the case where $n > 1$ we just have to remember that the closed cube Q_{m_0} and Q'_{m_0} are written as

$$Q_{m_0} = \Pi_{i=1}^n Q_{m_0,i} \quad \text{and} \quad Q'_{m_0} = \Pi_{i=1}^n Q'_{m_0,i}$$

where $Q_{m_0,i}$ and $Q'_{m_0,i}$ are closes intervals. Passing to these intervals and by extension to the one dimensional case which we have already proved and in conjunction with Note 3.8 we reach

$$\text{diam}(Q_\delta) = 2\sqrt{n}l(Q_\delta) > 2\sqrt{n}l(Q_{m_0}) = 2\ \text{diam}(Q_{m_0})$$

where $l(Q_\delta)$ and $l(Q_{m_0})$ denote the length of the sides of Q_δ and Q_{m_0} respectively. By proving this last inequality we reached a contradiction. □

CHAPTER 4

On the boundedness of the Hardy-Littlewood Maximal Operator (A_p condition)

The purpose of this chapter is to investigate the conditions under which a weight function w (non-negative measurable function) must comply with in order for the Hardy-Littlewood Maximal Operator to be weak-L^p_w . The condition we will come upon is called A_p [28]. The content of this chapter is influenced primary from the books of Duandikoetxea "Fourier Analysis" [9] and Grafakos "Classical Fourier Analysis [15].

4.1 The A_p condition ($p < \infty$)

Theorem 4.1. *Let $1 \leq p < \infty$. A necessary and sufficient condition for the Hardy-Littlewood Maximal Operator*[1] *to be weak-$L^p_w(\mathbb{R}^n)$ (where w is a weight function) i.e. there exists a constant $C > 0$ such that*

$$w(\{x : \mathcal{M}_q f(x) > \lambda\})^2 \leq \frac{C}{\lambda^p} \int_{\mathbb{R}^n} |f(x)|^p w(x)\, dx, \quad \textit{for every } f \in L^p_w(\mathbb{R}^n)$$

is for w to be locally integrable and satisfy the condition below.

- *If $p = 1$ then $\frac{w(Q)}{|Q|} \leq Cw(x)$, for every cube Q and a.e. $x \in Q$;*
- *if $p > 1$ then $\frac{w(Q)}{|Q|} \left(\int_Q w^{1-p'}\right)^{p-1} \leq C$, for every cube Q*

where in both cases C stands for a positive constant.

For each $p \in [1, \infty)$ the condition described above is called the "A_p condition".

Proof. First we are going to prove the necessity of the A_p condition and then its sufficiency.

[1]Throughout the course of the proof we will be using the uncentered cube variant of it, only in a small part we will also be using the cube centered variant of the Hardy-Littlewood Operator also and always with respect to the Lebesgue measure.

[2]For a measurable subset of $\mathbb{R}^n$ we set $w(A) := \int_A w(t)\, dt$.

- So suppose that the Hardy-Littlewood Maximal Operator is weak-$L_p(w)$ where w is weight and let us consider a function $f \in L^p$ that does not vanish a.e. and a cube Q such that $f(Q) := \int_Q f > 0$. Then substituting $f\chi_Q$ to the weak-L^p condition gives

$$w(\{\mathcal{M}_q f\chi_Q > \lambda\}) \leq \frac{C}{\lambda^p}\int_Q |f|^p w\,dx, \quad \text{for every } \lambda > 0.$$

If λ is moreover less than $|f(Q)|/|Q|$ where $f(Q) := \int_Q f$ then we obtain

$$w(Q) \leq \frac{C}{\lambda^p}\int_Q |f|^p w(x)\,dx.$$

Since this holds for every $\lambda < |f(Q)|/|Q|$, a limiting argument guarantees that it is also true for $|f(Q)|/|Q|$ i.e.

$$w(Q)\left(\frac{|f(Q)|}{|Q|}\right)^p \leq C\int_Q |f(x)|^p w(x)\,dx. \tag{4.1}$$

Now considering a measurable subset S of Q and substituting its characteristic function above, we obtain

$$w(Q)\left(\frac{|S|}{|Q|}\right)^p \leq Cw(S).$$

From the condition above we conclude the following two consequences.

1. *If w is not locally integrable then w is equal to ∞ a.e..*
   Indeed let Q be a cube such that $w(Q) = \infty$. Then for every cube $Q^{'}$ that contains Q we also have that $w(Q^{'}) = \infty$ and from the result above we get that $w(S) = \infty$ for every bounded measurable subset of $\mathbb{R}^n$. Thus Lebesgue's Differentiation theorem gives us that $w = \infty$ a.e.. As a result w must be locally integrable.
2. *If w vanishes on a set of positive measure then w vanishes a.e..*
   Let S be such that $|S| > 0$ and $w|_S = 0$. Then for every cube that contains S we have $w(Q) = 0$ and since w is non negative we get that w is equal to zero a.e.. Thus for the rest of the proof we can assume that $w > 0$ almost everywhere.

Now we will examine the cases where $p = 1$ and $p > 1$ separately.

- If $p = 1$ then the inequality we reached is

$$\frac{w(Q)}{|Q|} \leq C\frac{w(S)}{|S|}, \quad \text{for every cube } Q \text{ and } S \subset Q \text{ such that } |S| > 0.$$

If we denote by s the essential infimum of f on Q, then we have that for every $\epsilon > 0$ there is a S_ϵ subset of Q of positive measure such that $w \leq s + \epsilon$ for almost every $x \in S_\epsilon$, thus

$$\frac{w(Q)}{|Q|} \leq C(s + \epsilon).$$

Since this holds true for every $\epsilon > 0$ a limiting argument gives us

$$\frac{w(Q)}{|Q|} \leq Cs.$$

As a result

$$\frac{w(Q)}{|Q|} \leq Cw(x), \quad \text{for every cube } Q \text{ and a.e. } x \in Q \tag{A_1}$$

where $C > 0$ is a constant.

– In case $p > 1$ we consider the sequence $\{f_n\}_{n\in\mathbb{N}}$ with $f_n := \min\left\{w^{1-p'}, n\right\}$ where $1/p + 1/p' = 1$ and replace each of its terms to inequality (4.1). With that we obtain

$$w(Q)\left(\frac{f_n(Q)}{|Q|}\right)^p \leq C\int_Q |f_n(x)|^p|w(x)\,dx$$

or equivalently

$$w(Q)\left(\frac{1}{|Q|}\int_Q \min\left\{n, w^{1-p'}\right\}\right)^p\left(\int_Q\left(\min\left\{n, w^{1-p'}\right\}\right)^p\right)^{-1} \leq C.$$

Since this sequence is increasing (i.e. $f_{n+1} \geq f_n$) and nonnegative the monotone convergence theorem gives us

$$w(Q)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1} \leq C.$$

The reason we didn't just replace the function $w^{1-p'}$ to the relation above from the beginning is that we didn't know whether it was locally integrable or not beforehand. This last relation shows us that in fact it is.
So we have proved

$$w(Q)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1} \leq C, \quad \text{for every cube } Q \tag{A_p}$$

where $C > 0$ depends on w.

- *Now let's prove the sufficiency of the condition above.* Again we will examine the cases where $p = 1$ or not separately.

  – For $p = 1$ the A_1 condition can be expressed as

$$\frac{w(Q)}{|Q|} \leq Cw(x), \quad \text{for every cube } Q \text{ and for almost every } x \in Q$$

  for some constant $C > 0$.
  We prove its necessity in two steps.

1. First we will show that the A_1 condition is equivalent to

$$\mathcal{M}_q w(x) \leq C w(x), \quad \text{for almost every } x \in \mathbb{R}^n$$

   and also with the same constant C
2. and then prove the following inequality

$$\int_{\{x:\mathcal{M}_q f(x)>\lambda\}} w(x)\,dx \leq \frac{C}{\lambda}\int_{\mathbb{R}^n} |f(x)|\mathcal{M}_q w(x)\,dx, \quad \text{for every } \lambda{>}0.$$

Provided we proved these two statements the weak-$L^1(w)$ condition for the Hardy-Littlewood Maximal Operator then follows immediately.

1. Let x be a point such that

$$Mw(x) > Cw(x).$$

   Then there exists a cube Q comprised of rational vertices that contains x such that

$$\frac{1}{|Q|}\int_Q w > Cw(x).$$

   But the set of such points in each cube has zero measure and since the collection of all such cubes is countable, the set of all those points has measure zero. Equivalently

$$Mw(x) \leq Cw(x), \quad \text{for a.e. x.}$$

   The opposite direction is an immediate result of the definitions.
2. For the second one we are going to use the Calderon-Zygmund Decomposition. So let us assume that $f \in L^1(\mathbb{R}^n)$ and without loss of generality we may also assume that $f \geq 0$ and for $\lambda > 0$ we consider the respective Calderon-Zygmund Decomposition $\{D_i\}_{i\in I}$ at λ.
   Then similar to Theorem 2.18 we have

$$\{M_{qc} f > 4^n\lambda\} \subset \bigcup_{i\in I} 2D_i$$

   where $2D_i$ stands for the cube that has the same center as D_i and its sides have two times the length of D_i.
   So

$$\begin{aligned}
\int_{\{M_{qc}f>4^n\lambda\}} w(x)\,dx &\leq \sum_{i\in I}\int_{2D_i} w(x)\,dx \\
&= \sum_{i\in I} 2^n|D_i|\frac{1}{|2D_i|}\int_{2D_i} w(x)\,dx \\
&\leq \frac{2^n}{\lambda}\sum_{i\in I}\int_{D_i} f(y)\left(\frac{1}{|2D_i|}\int_{2D_i} w(x)\,dx\right)dy \\
&\leq \frac{2^n C}{\lambda}\int_{\mathbb{R}^n} f(y) M_q w(y)\,dy.
\end{aligned}$$

   Now using the fact that the different variants of the Hardy-Littlewood Maximal Operator are pointwise equivalent (Remark 2.3) we have our result.

- To prove the sufficiency for $p > 1$ we will first prove

$$w(Q)\left(\frac{f(Q)}{|Q|}\right)^p \leq C \int_Q |f|^p w\, dx$$

for every $f \in L^p_w$.

Indeed Hölder's inequality gives us

$$\begin{aligned}\left(\frac{\int_Q |f|\, dx}{|Q|}\right)^p &= \left(\frac{1}{|Q|}\int_Q |f| w^{\frac{1}{p}} w^{-\frac{1}{p}}\, dx\right)^p \\ &\leq \left(\frac{1}{|Q|}\int_Q |f|^p w\, dx\right)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1} \\ &\leq C\left(\frac{1}{|Q|}\int_Q |f|^p w\, dx\right)\left(\frac{|Q|}{w(Q)}\right)\end{aligned}$$

and so we obtain that $f \in L^1_{loc}$. Since the result we want to prove is local without loss of generality we may also assume that $f \in L^1$.

Now for a $\lambda > 0$ we consider the Calderon-Zygmund Decomposition at $4^{-n}\lambda$, i.e.

$$\{\mathcal{M}_d f > 4^{-n}\lambda\} = \bigcup_{i\in I} D_i.$$

Then again using a similar argument as in Theorem 2.15 but with a constant 3 since the cubes are not necessarily centered around the points so we need to dilate further the cubes we have

$$\{\mathcal{M}_q f > \lambda\} \subset \bigcup_{i\in I} 3D_i$$

and so

$$\begin{aligned}w(\{\mathcal{M}_q f > \lambda\}) &\leq \sum_{i\in I} w(3D_i) \\ &= \sum_{i\in I} w(3D_i)\left(\frac{|D_i|}{|3D_i|}\right)^p\left(\frac{|3D_i|}{|D_i|}\right)^p \\ &\leq C3^{np}\sum_{i\in I} w(D_i) \\ &= C3^{np}\sum_{i\in I} w(D_i)\left(\frac{f(D_i)}{|D_i|}\right)^p\left(\frac{|D_i|}{|f(D_i)|}\right)^p \\ &\leq C^2 3^{np}\sum_{i\in I}\left(\frac{|D_i|}{f(D_i)}\right)^p\int_{D_i}|f|^p w\, dx \\ &\leq C3^{np}\left(\frac{4^n}{\lambda}\right)^p\int_{\mathbb{R}^n}|f|^p w\, dx.\end{aligned}$$

□

Note 4.2. The constant $C > 0$ appearing in the A_p condition must be greater or equal to 1.

Indeed as we were trying to prove the necessity of the A_p condition we reached the inequality

$$w(Q)\left(\frac{|S|}{|Q|}\right)^p \leq Cw(S)$$

for every $S \subset Q$ measurable.

Now by replacing Q in place of S we have our result.

Our next purpose is to show a characteristic property that all the A_p classes have. A useful tool that we will use is the so called "Reverse Hölder's Inequality" which we present right below but before that let us prove a lemma which we will be needing.

Lemma 4.3. *Let w be an A_p weight, then for every $a \in (0,1)$ there exists a number $b \in (0,1)$ such that if Q is a cube and $S \subset Q$ with*

$$|S| \leq a|Q|$$

then

$$w(S) \leq bw(Q).$$

Proof. As we were trying to prove the necessity of the A_p condition we stumbled upon the inequality,

$$\left(\frac{|S|}{|Q|}\right)^p \leq C\frac{w(S)}{w(Q)}, \quad \text{for every cube } Q \text{ and measurable } A \subset Q.$$

By setting $S := Q \setminus A$ the above inequality can be equivalently written as

$$\left(1 - \frac{|S|}{|Q|}\right)^p \leq C\left(1 - \frac{w(S)}{w(Q)}\right).$$

So it is enough to set $b := 1 - (1-a)^p/C$ which belongs to $(0,1)$ since $C \geq 1$. □

Theorem 4.4 (Reverse Hölder's Inequality)**.** *Let w be an A_p weight, $p > 1$. Then there are constants C and γ such that*

$$\left(\frac{1}{|Q|}\int_Q (w(x))^{1+\gamma}\,dx\right)^{\frac{1}{1+\gamma}} \leq C\left(\frac{1}{|Q|}\int_Q w\,dx\right)$$

for every cube Q.

Proof. For a fixed cube Q and an $a \in (0,1)$ let

$$a_0 := \frac{1}{|Q|}\int_Q w$$

and consider an increasing sequence

$$a_0 < a_1 < a_2 < \dots$$

where

$$a_{k+1} := (2^{-n} a^{-1}) a_k.$$

and for each of the a_k we consider the Calderon-Zygmund Decomposition of Q at a_k [3].

Then we obtain a sequence of dyadic cubes $\{D_i^k\}$ for each k such that

1. $a_k < \frac{1}{|D_i^k|} \int_Q w \leq 2^n a_k$
2. $w(x) \leq a_k$ for a.e. $x \notin U_k := \cup D_i^k$.

Also from the maximality of these cubes with respect to the condition

$$\frac{1}{|Q'|} \int_{Q'} w > a_k$$

we have that each cube D_l^{k+1} is contained at some cube D_m^k for every $k \in \mathbb{N}$.

So

$$\begin{aligned} 2^n a_k \geq \frac{1}{|D_{k,l}|} \int_{D_{k,l}} w &\geq \frac{1}{|D_{k,l}|} \int_{D_{k,l} \cap U_{k+1}} w \\ &= \frac{1}{|D_{k,l}|} \sum_{j : D_{k+1,j} \subset Q_{k,l}} |D_{k+1,j}| \frac{1}{|D_{k+1,j}|} w \\ &> \frac{|D_{k,l} \cap U_{k+1}|}{|D_{k,l}|} a_{k+1} \\ &= \frac{|D_{k,l} \cap U_{k+1}]|}{|D_{k,l}|} 2^n a^{-1} a_k. \end{aligned}$$

Consequently $|D_{k,l} \cap U_{k+1}| \leq a |D_{k,l}|$ and so from Lemma 4.3

$$\frac{w(D_{k,l} \cap U_{k+1})}{w(D_{k,l})} \leq b$$

and so

$$w(U_{k+1}) \leq b w(U_k).$$

As a result $w(U_k) \leq b^k w(U_0)$. Also since $|U_{k+1}| \leq a|U_k|$ implies that $|U_k| \to 0$, we get

$$Q \approx {}^{4} (Q \setminus U_0) \bigcup \left(\bigcup_{i=1}^{\infty} U_k \setminus U_{k+1} \right).$$

[3]The proof in this case is similar to the one we presented with respect to the grids $2^{-N}\mathbb{Z}^n$

[4]We mean that their set difference has zero measure

For every $\gamma > 0$ we have

$$\begin{aligned}\int_Q w^{1+\gamma} &= \int_{Q\setminus U_0} w^\gamma(t)w(t)\,dt + \sum_{i=0}^{\infty}\int_{U_i\setminus U_{i+1}} w^\gamma(t)w(t)\,dt \\ &\le a_0^\gamma w(Q\setminus U_0) + \sum_{i=0}^{\infty} a_{i+1}^\gamma w(U_i) \\ &= a_0^\gamma w(Q\setminus U_0) + \sum_{i=0}^{\infty}((2^n a^{-1})^{i+1}a_0)^\gamma b^i w(U_0) \\ &\le a_0^\gamma\left(1 + (2^n a^{-1})^\gamma \sum_{i=0}^{\infty}(2^n a^{-1})^{\gamma i} b^i\right) w(Q).\end{aligned}$$

Thus by choosing $\gamma > 0$ small enough so that $(2^n a^{-1})^\gamma b < 1$ we obtain

$$\frac{1}{|Q|}\int_Q w^{1+\gamma} \le \left(\frac{1}{|Q|}\int_Q w\right)^\gamma \left(1 + \frac{(2^n a)^\gamma}{1-(2^n a)^\gamma b}\right)\int_Q w.$$

Now since the constants C, γ appearing here are independent of the cube Q the proof is complete. □

Remark 4.5. One may notice that if we replace the Lebesgue measure with another Radon measure μ that satisfies the doubling condition

$$\mu(3Q) \le C\mu(Q), \quad \text{for every cube } Q$$

where $C > 0$ is constant, the same result still holds assuming that if

$$\mu(A) \le \mu(Q)$$

then

$$w(A) \le w(Q).$$

Later on we will need this particular remark. For more information check [15] and [8].

Lemma 4.6. *A weight w is A_p for some $p > 1$ if and only if $w^{1-p'}$ is an $A_{p'}$ weight.*

Proof. The $A_{p'}$ condition of $w^{1-p'}$ can be expressed as

$$\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)\left(\frac{1}{|Q|}\int_Q (w^{1-p'})^{1-p}\right)^{p'-1} \le C$$

or equivalently

$$\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)\left(\frac{1}{|Q|}\int_Q w\right)^{p'-1} \le C$$

which is just the A_p condition of w raised to $p'-1$. □

Theorem 4.7. *For every $p>1$ the class A_p is compromised of all the A_q weights for $1\le q<p$ i.e.*

$$A_p = \bigcup_{1\le q<p} A_q.$$

Proof. If w is an A_1 weight then

$$\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1} \le \sup_Q w^{-1} = (\inf_Q w)^{-1} \le C\frac{|Q|}{w(Q)}$$

and so $w\in A_p$.

As for the case where w is an A_q weight where $q\in(1,p)$ Hölder's inequality gives us

$$\begin{aligned}\left(\frac{1}{|Q|}\int_Q w^{1-p'}\right)^{p-1} &\le \left(\frac{1}{|Q|}\left(\int_Q w^{1-q'}\right)^{\frac{p'-1}{q'-1}} |Q|^{\frac{q'-p'}{q'-1}}\right)^{p-1}\\ &= \left(\left(\int_Q w^{1-q'}\right)^{\frac{p'-1}{q'-1}} |Q|^{\frac{1-p'}{q'-1}}\right)^{p-1}.\end{aligned}$$

But the right side of the inequality above is just the left side of the A_q condition of w which is turn is less or equal to $C|Q|/w(Q)$ for some constant $C>0$.

As for the opposite direction as we have already shown if w is an A_p weight then $w^{1-p'}$ is an $A_{p'}$ weight and so by using the Reverse Hölder's Inequality (Theorem 4.4) we obtain constants $\gamma, C>0$ such that

$$\left(\frac{1}{|Q|}\int_Q w^{(1-p')(1+\gamma)}\right)^{\frac{1}{1+\gamma}} \le \frac{C}{|Q|}\int_Q w^{1-p'}$$

for every cube Q. Combining it with the A_p condition of w we reach

$$\left(\frac{1}{|Q|}\int_Q w^{(1-p')(1+\gamma)}\right)^{\frac{1}{1+\gamma}} \le C\left(\frac{w(Q)}{|Q|}\right)^{\frac{1}{p-1}}.$$

Now considering a $q>1$ such that $q'-1 := (p'-1)(1+\gamma)$, we then have $q<p$ and the above inequality can equivalently expressed as

$$\left(\frac{1}{|Q|}\int_Q w^{1-q'}\right)^{\frac{p'-1}{q'-1}} \le C\left(\frac{w(Q)}{|Q|}\right)^{\frac{1}{p-1}}$$

which in turn is equivalent for w to be an A_q weight. □

4.2 The A_∞ condition

As we have seen, for $p>1$ the A_p condition is the union of all classes A_q with $q<p$. It turns out we can find a condition independent of p that all A_p weights satisfy. This is what we are going to concern ourselves with, so let us start with a definition.

Definition 4.8. We define as A_∞ the class of all weights for which there exists some constant $C > 0$ such that

$$\frac{1}{|Q|}\int_Q w \leq C \exp\cdot\left(\frac{1}{|Q|}\int_Q \log w\right), \quad \text{for every cube } Q.$$

The definition of A_∞ is a natural extension of the A_p condition for $p < \infty$ for if w is an A_p weight then

$$\frac{w(Q)}{|Q|}\cdot\exp\left(\frac{1}{|Q|}\int_Q \log w^{-1}\right) \leq \frac{w(Q)}{|Q|}\left(\frac{1}{|Q|}\int_Q w^{-\frac{1}{p-1}}\right)^{p-1} \leq C.$$

so letting $p \to \infty$ gives us the A_∞ condition. It turns out that *ONLY* the A_p weights satisfy this condition, in order to prove that let us first show some properties that we will need.

Theorem 4.9. *The following hold true for an A_∞ weight where*

$$|w|_{A_\infty} := \sup_{Q:cube}\frac{w(Q)}{|Q|}\cdot\exp\left(\frac{1}{|Q|}\int_Q \log w^{-1}\right).$$

1. $|w|_{A_\infty} \geq 1$.
2. $|w|_{A_\infty}$ *can equivalently be expressed as*

$$|w|_{A_\infty} = \sup_{Q:cube}\sup_{\log f\in L^1(Q)}\left\{\frac{w(Q)}{\int_Q fw}\cdot\exp\left(\frac{1}{|Q|}\int_Q \log|f|\right)\right\}.$$

3. *w is a doubling measure. In particular for every $\lambda > 1$ and every cube Q*

$$w(\lambda Q) \leq 2^{\lambda^n}|w|_{A_\infty}^{\lambda^n} w(Q).$$

Proof. 1. It is a straightforward consequence of Jensen's inequality.

2. Replacing $f := w^{-1}$ to the expression inside the supremum is taken over gives us that its supremum is greater or equal to $|w|_{A_\infty}$.

   As for the other direction, Jensen's inequality gives us

$$\exp\left\{\frac{1}{|Q|}\int_Q \log(|f|w)\right\} \leq \frac{1}{|Q|}\int_Q |f|w$$

   or equivalently by using the fact that $\log(ab) = \log(a) + \log(b)$ for $a, b > 0$ we obtain

$$\frac{w(Q)}{\int_Q |f|w}\cdot\exp\left(\frac{1}{|Q|}\int_Q \log|f|\right) \leq \frac{w(Q)}{|Q|}\cdot\exp\left(-\frac{1}{|Q|}\int_Q \log w\right).$$

3. For a fixed cube Q and $\lambda > 1$ let f defined as

$$f(x) = \begin{cases} c, & \text{if } x \in Q \\ 1, & \text{if } x \in \mathbb{R}^n \setminus Q \end{cases}$$

where $c > 0$ is such that $c^{1/\lambda^n} = 2[w]_{A_\infty}$. Substituting it in the expression inside the supremum of 2. gives us

$$\frac{w(\lambda Q)}{w(\lambda Q \setminus Q) + cw(Q)} \cdot \exp\left\{\frac{\log c}{\lambda^n}\right\} \leq [w]_{A_\infty}$$

or equivalently

$$\frac{w(\lambda Q)}{w(\lambda Q \setminus Q) + 2^{\lambda^n}[w]_{A_\infty}^{\lambda^n} w(Q)} 2[w]_{A_\infty} \leq [w]_{A_\infty}$$

which in turn gives us

$$2w(\lambda Q) \leq w(\lambda Q \setminus Q) + 2^{\lambda^n}[w]_{A_\infty}^{\lambda^n} w(Q)$$

or

$$w(\lambda Q) \leq (2^{\lambda^n}[w]_{A_\infty}^{\lambda^n} - 1)w(Q).$$ □

Note 4.10. For the rest of this chapter we will use the following notation. For a locally integrable function we set

$$< f >_Q := \frac{1}{|Q|}\int_Q f.$$

Theorem 4.11. *A weight w is an A_∞ weight if and only if it satisfies one of the following.*

1. *There are constants $\gamma, \delta \in (0, 1)$ such that for every cube Q*

$$|\{x \in Q : w(x) \leq \gamma < w >_Q\}| \leq \delta|Q|.$$

2. *There exist constants a, b such that if Q is a cube and $A \subset Q$ measurable with*

$$|A| \leq a|Q|$$

then

$$w(A) \leq bw(Q).$$

3. *w satisfies the reverse Hölder's inequality i.e. there exist $\epsilon, C > 0$ such that*

$$\left(\frac{1}{|Q|}\int_Q w^{1+\epsilon}\right)^{\frac{1}{1+\epsilon}} \leq \frac{C}{|Q|}\int_Q w, \quad \text{for every cube } Q.$$

4. *There are constants $\epsilon, C > 0$ such that for every cube Q and $A \subset Q$ measurable*

$$\frac{w(A)}{w(Q)} \leq C\left(\frac{|A|}{|Q|}\right)^\epsilon.$$

5. *There are constants $a, b \in (0,1)$ such that if Q is a cube and $A \subset Q$ with*
$$w(A) \le a w(Q)$$
then
$$|A| \le b|Q|.$$

6. *w is an A_p weight for some $p < \infty$.*

Proof. Throughout this proof we will use the symbol $< w >_Q$ to denote average integral $w(Q)/Q$.

- $(A_\infty) \implies (1)$
Since $|\lambda w|_{A_\infty} = |w|_{A_\infty}$ for any $\lambda > 0$ we may assume that $\int_Q \log w = 0$ and so $< w >_Q \le |w|_{A_\infty}$. Then
$$\begin{aligned} |\{x \in Q : w(x) \le \gamma w < w >_Q\}| &\le |\{x \in Q : w(x) \le \gamma |w|_{A_\infty}\}| \\ &= |\{x \in Q : \log(1 + w(x)^{-1}) \ge \log(1 + (\gamma |w|_{A_\infty})^{-1})\}| \\ &\le \frac{1}{\log(1 + (\gamma |w|_{A_\infty})^{-1})} \int_Q \log \frac{1+w}{w} \\ &= \frac{1}{\log(1 + (\gamma |w|_{A_\infty})^{-1})} \int_Q \log(1+w) \\ &\le \frac{1}{\log(1 + (\gamma |w|_{A_\infty})^{-1})} \int_Q w \\ &\le \frac{|w|_{A_\infty}|Q|}{\log(1 + (\gamma |w|_{A_\infty})^{-1})} = \frac{1}{2}|Q| \end{aligned}$$
So setting $\gamma := |w|_{A_\infty}^{-1}(e^{2|w|_{A_\infty}} - 1)^{-1}$ we obtain the inequality we sought with $\delta := \frac{1}{2}$.

- $(1) \implies (2)$
Let Q be a cube and $A \subset Q$ with $w(A) \ge b w(Q)$ with b to be defined later on. By setting $S := Q \setminus A$ we obtain $w(S) \le (1-b) w(Q)$. Now decomposing S as $S_1 \cup S_2$ where
$$S_1 = \{x \in S : w(x) \ge \gamma < w >_Q\} \quad \text{and} \quad S_2 = \{x \in S : w(x) \le \gamma < w >_Q\}.$$
From what we already proved $|S_2| \le \delta |Q|$. As for S_1
$$|S_1| \le \frac{1}{\gamma < w >_Q} \int_S w = \frac{|Q|}{\gamma} \frac{w(S)}{w(Q)} \le \frac{1-b}{\gamma}|Q|.$$
As a result
$$|S| = |S_1| + |S_2| \le \frac{1-b}{\gamma}|Q| + \delta|Q| = \left(\delta + \frac{1-b}{\gamma}\right)|Q|.$$
Setting now $a := \frac{1-\delta}{2}$ and $b := 1 - \frac{(1-\delta)\gamma}{2}$ gives the result we sought.

- (2) $\implies$ (3)

  The proof of it is just the remark of Reverse Hölder's inequality (Remark 4.5).

- (3) $\implies$ (4)

  Applying first Hölder's Inequality and then the Reverse one gives us

$$\begin{aligned}\int_A w &\le \left(\int_A w^{1+\epsilon}\right)^{\frac{1}{1+\epsilon}} |A|^{\frac{\epsilon}{1+\epsilon}} \\ &\le \left(\frac{1}{|Q|}\int_Q w^{1+\epsilon}\right)^{\frac{1}{1+\epsilon}} |Q|^{\frac{1}{1+\epsilon}}|A|^{\frac{\epsilon}{1+\epsilon}} \\ &\le \frac{C}{|Q|}\int_Q w \cdot |Q|^{\frac{1}{1+\epsilon}}|A|^{\frac{\epsilon}{1+\epsilon}}\end{aligned}$$

  and so

$$\frac{w(A)}{w(Q)} \le C\left(\frac{|A|}{|Q|}\right)^{\epsilon_0}$$

  where $\epsilon_0 := \frac{\epsilon}{1+\epsilon}$.

- (4) $\implies$ (5)

  Let us first consider a small enough so that $b := Ca^\epsilon < 1$. Then if Q is a cube and $A \subset Q$ with

$$|A| \le a|Q|$$

  then

$$w(A) \le bw(Q).$$

  Replacing A with $Q \setminus A$ the condition above can be equivalently stated as,
  if $A \subset Q$ with

$$|A| \ge (1-a)|Q|$$

  then

$$w(A) \ge (1-b)w(Q)$$

  or equivalently,
  if $A \subset Q$ satisfies

$$w(A) \le (1-b)w(Q)$$

  then

$$|A| \le (1-a)|Q|.$$

- (5) $\implies$ (6)

  Implication (5) can equivalently be expressed as follows, if $w(A) \le aw(Q)$ then

$$\int_A w^{-1}(t)w(t)\,dt \le b\int_Q w^{-1}(t)w(t)\,dt.$$

  or equivalently, if

$$w(A) \le aw(Q)$$

then

$$\int_A w^{-1}\,d\mu \le \int_Q w^{-1}\,d\mu$$

where μ is defined as

$$\mu(\mathcal{O}) := \int_{\mathcal{O}} w\,dx, \quad \text{for every } \mathcal{O} \text{ Borel measurable subset of } \mathbb{R}^n.$$

And so by applying Remark 4.5 we obtain

$$\left(\frac{1}{w(Q)}\int_Q w^{-1-\gamma}\,d\mu\right)^{\frac{1}{1+\gamma}} \le \frac{C}{w(Q)}\int_Q w^{-1}d\mu$$

for every cube Q, where $C > 0$.

So if we set $p := 1 + 1/\gamma$ we obtain that $p' = \gamma + 1$ and the above inequality can equivalently be expressed as

$$\left(\frac{1}{w(Q)}\int_Q w^{1-p'}\right) \le \frac{C|Q|}{w(Q)}$$

which in turn shows us that w is an A_p weight.

- (6) $\implies$ (A_∞) is already proved. □

CHAPTER 5

Bellman functions on Harmonic Analysis

In this last chapter we will study two Bellman functions, that of the dyadic maximal operator and Carleson's Embedding theorem but we are not going to concern ourselves on how they came up. The reader who wants to learn more can start by looking [31],[30] and [39]. Before that we introduce a generalization of the notion of the family dyadic cubes we introduced in chapter 2 which will now be named as tree and present a theorem that associates the dyadic maximal operator with the Hardy operator.

5.1 A Theorem regarding the Dyadic Maximal Operator

5.1.1 Primary Notions

Definition 5.1. Let (X, μ) be a nonatomic probability space. A *tree* $\mathcal{T}$ is a collection of measurable subsets of X such that

1. $X \in \mathcal{T}$.
2. For every element $A \in \mathcal{T}$ there exists a sub-collection $\mathcal{C}(A)$ of $\mathcal{T}$ such that
$$A = \bigcup_{B \in \mathcal{C}(A)} B$$
such that the B are pairwise disjoint.
3. The tree $\mathcal{T}$ satisfies
$$\mathcal{T} = \bigcup_{n \in \mathbb{N}} \mathcal{T}_n,$$
where $\mathcal{T}_0 := \{X\}$ and for $m \geq 0$
$$\mathcal{T}_{m+1} := \bigcup_{A \in \mathcal{T}_m} \mathcal{C}(A).$$
4. Moreover $\lim_{m \to \infty} \sup_{I \in \mathcal{T}_m} \mu(I) = 0$.

Example 5.2. If we consider the measure space $(0,1]^n$ embedded with the Lebesgue measure, the dyadic cubes of $(0,1]^n$ which are described in Chapter 2 constitute a tree on that space.

Definition 5.3. Let (X,μ) be a non atomic measure space and $\mathcal{T}$ be a tree of it. We define the *dyadic maximal operator of a measurable function $f : X \to \mathbb{R}$ with respect to* $\mathcal{T}$ as follows,

$$\mathcal{M}_{\mathcal{T}} f(x) := \sup_{I\in\mathcal{T}: x\in I} \frac{\int_I |f|\, d\mu}{\mu(I)}, \quad x \in X.$$

5.1.2 The Geometry of the Dyadic Maximal Operator

This subsection shares the same name and content as in the paper of Nikolidakis [32]. The main theorem of this subsection is the following.

Theorem 5.4. *Let $(X,\mu,\mathcal{T})$ be a nonatomic probability space equipped with a tree structure, $g : [0,\infty) \to [0,\infty)$ be an decreasing, right continuous function, $h : [0,\infty) \to [0,\infty)$ be an integrable function and $G : [0,\infty) \to [0,\infty)$ is an increasing function. Then for every $k \in (0,1]$ the following holds*

$$\sup\left\{\int_K G\left[(\mathcal{M}_{\mathcal{T}}\phi)^*(t)\right] h(t)\,dt, \phi \geq 0, \phi^* = g, K \subset (0,1] \text{ measurable with } |K| = k\right\}$$
$$= \int_0^k G\left(\frac{1}{t}\int_0^t g(u)\,du\right) h(t)\,dt.$$

For the proof we will need the following lemmas.

Lemma 5.5. *Let (X,μ) be a nonatomic probability space. Then for every I in the tree $\mathcal{T}$ and every $0 < a < 1$ there exists a family $\mathcal{F}_a$ of pairwise disjoint subsets of I belonging to the tree $\mathcal{T}$ such that*

$$\sum_{J\in\mathcal{F}_a} \mu(J) = (1-a)\mu(I).$$

Proof. We will examine the case where $I = X$ since for any other member of the tree we can follow the same procedure. From (iv) of the definition there exists an m_1 such that for every $I \in \mathcal{T}_{m_1}$,

$$\mu(I) \leq \frac{1-a}{2}.$$

We consider a family $F_1 \subset \mathcal{T}_{m_1}$ maximal so as

$$\sum_{I\in\mathcal{T}_{m_1}} \mu(I) \leq 1-a.$$

Now if equality occurs in the above inequality we are finished.

Otherwise we have

$$\sum_{I\in\mathcal{T}_{m_1}} \mu(I) \geq \frac{1-a}{2}$$

for else we could add another set of $\mathcal{T}_{m_1}$ to the collection $\mathcal{F}_1$ which contradicts the maximal property of $\mathcal{F}_1$. Now we set $\epsilon_1 := 1 - a - \sum_{I\in\mathcal{T}_{m_1}} \mu(I)$. By property 4. of definition of a tree there exists $m_2 > m_1$ such that for every $I \in \mathcal{T}_{m_2}$

$$\mu(I) \le \frac{\epsilon_1}{2}.$$

We then find a maximal family $\mathcal{F}_2$ of sets in $\mathcal{T}_2$ disjoint from the ones in $\mathcal{F}_1$ such that

$$\sum_{I\in F_2} \mu(I) \le \epsilon_1.$$

If now equality occurs then $F_1 \cup F_2$ is a family with the desired property. If not then as before

$$\sum_{I\in\mathcal{F}_2} \mu(I) > \frac{\epsilon_1}{2}$$

and so

$$0 < \epsilon_2 := \epsilon_1 - \sum_{I\in\mathcal{F}_2} \mu(I) \le \frac{\epsilon_1}{2} < \frac{1-a}{4}$$

and by extension

$$0 < (1-a) - \sum_{I\in\mathcal{F}_1\cup\mathcal{F}_2} \mu(I) < \frac{1-a}{4}$$

and proceed accordingly.

So this procedure has two possible outcomes. To finish after a finite number of steps or not. If the first occurs that means the sets we have elected until that point form a family with the desired property. If not then at each step we obtained a family $\mathcal{F}_i$ such that

$$0 < (1-a) - \sum_{I\in\mathcal{F}_1\cup\cdots\cup\mathcal{F}_i} \mu(I) < \frac{1-a}{2^i}.$$

Indeed as we have shown the result holds for $i = 1, 2$, so let us assume that it holds for some $i \ge 2$. Then for the family $\mathcal{F}_{m+1}$ we have

$$0 < \epsilon_{i+1} = (1-a) - \sum_{I\in\mathcal{F}_1\cup\cdots\cup\mathcal{F}_i} \le \frac{\epsilon_i}{2} < \frac{1-a}{2^{i+1}}$$

where in the last step we used our inductive hypothesis.

So the sets that were chosen during the process also form a family with the desired property. □

Lemma 5.6. *Let (X, μ) be a nonatomic measure space and $(A_i)_{i\in I}$ be a countable, measurable partition of X such that $\mu(A_i) > 0$ for every $i \in I$. Then if $\phi : X \to \mathbb{R}^+$ is an integrable function such that $\int_X \phi\, d\mu = f$, there exists a rearrangement h of ϕ so that*

$$\frac{1}{\mu(A_i)} \int_{A_i} h\, d\mu = f, \quad \textit{for every } i \in I.$$

Proof. First by setting $g := \phi^*$, we will find a measurable set $B_1 \subset [0,1]$ such that

$$|B_1| = \mu(A_1) \quad \text{and} \quad \frac{1}{|B_1|}\int_{B_1} g(u)\,du = f.$$

In order to achieve that let us first notice that

$$\frac{1}{\mu(A_1)}\int_0^{\mu(A_1)} g(u)\,du \geq f \geq \frac{1}{\mu(A_1)}\int_{1-\mu(A_1)}^1 g(u)\,du. \tag{5.1}$$

For the proof of this inequality remember (Theorem 1.12) that

$$\int_X \phi\,d\mu = \int_0^1 g(u)\,du.$$

Now since $x \mapsto 1/x\int_0^x g$ is decreasing on $(0,1]$ the first inequality stems directly from it. Similarly the second stems from the fact that $x \mapsto 1/x\int_{1-x}^1 g$ is decreasing.

Now as a result of inequality (5.1) there exists an $r \in [0, 1-\mu(A_1)]$ such that,

$$\frac{1}{\mu(A_1)}\int_r^{r+\mu(A_1)} g(u)\,du = f.$$

If we set $B_1 := [r, r+\mu(A_1)]$ then

$$|B_1| = \mu(A_1) \quad \text{and} \quad \frac{1}{|B_1|}\int_{B_1} g(u)\,du = f.$$

Now since X is nonatomic there exists a measurable function $h_1 : A_1 \to [0,\infty)$ such that $(h_1)^* = (g/B_1)^*$, thus

$$\frac{1}{\mu(A_1)}\int_{A_1} h_1\,d\mu = f.$$

Finally we repeat the same procedure on $X \setminus A_1$ and so on. By doing so we extend the function h_1 successively on X in a way that so that it satisfies the desired properties. □

Lemma 5.7. *For every $\phi : X \to [0,\infty)$ integrable,*

$$(\mathcal{M}_\mathcal{T}\phi)^*(t) \leq \frac{1}{t}\int_0^t \phi^*(u)\,du, \quad \textit{for every } t > 0.$$

Proof. Let $\phi : X \to [0,\infty)$ be an integrable function and $t \in (0,1)$. Since the case where ϕ^* is constant is straightforward let us assume that ϕ^* is not. In case ϕ^* is constant on some interval $(0,a)$ with $0 < a < 1$ the result still holds in this interval since $\phi^*(x) = \phi^*(0) = \text{esssup}\phi$ for every $x \in (0,a)$. So we only have to examine the case where $x \notin (0,a)$[1]. Now let $A := \frac{1}{t}\int_0^t \phi^*(u)\,du$. Then since ϕ^* is decreasing we have that

$$A \geq \int_0^1 \phi^*(u)\,du.$$

[1]We will use similar arguments in the sequel without mentioning it specifically.

Setting $E := \{\mathcal{M}_{\mathcal{T}}\phi > A\}$ gives us

$$\mu(E) \leq \frac{1}{A}\int_E \phi \, d\mu.$$

As a consequence

$$A = \frac{1}{t}\int_0^t \phi^*(u)\, du \leq \frac{1}{\mu(E)}\int_E \phi\, d\mu \leq \frac{1}{\mu(E)}\int_0^{\mu(E)} \phi^*(u)\, du.$$

And due to the fact that ϕ^* is non-increasing we conclude that $\mu(E) < t$.

Now since $(\mathcal{M}_{\mathcal{T}}\phi)^*$ and $\mathcal{M}_{\mathcal{T}}\phi$ are equimeasurable we have that

$$\mu(E) = \mu(\{(\mathcal{M}_{\mathcal{T}}\phi)^* > A\}).$$

Since $(\mathcal{M}_{\mathcal{T}}\phi)^*$ is decreasing the set $\{(\mathcal{M}_{\mathcal{T}}\phi)^* > A\}$ is an interval of the form $[0, \gamma)$ and since $\mu(E) < t$ we obtain that $\gamma < t$. So $t \notin \{(\mathcal{M}_{\mathcal{T}}\phi)^* > A\}$ and so

$$(\mathcal{M}_{\mathcal{T}}\phi)^*(t) \leq A = \frac{1}{t}\int_0^t \phi^*(u)\, du. \qquad \square$$

Proof of theorem. We will find for every $a \in (0, 1]$ a measurable function $\phi_a : X \to [0, \infty)$ with $(\phi_a)^* = g$ so that

$$\limsup_{a\to 0^+} \int_0^k G[(\mathcal{M}_{\mathcal{T}}\phi_a)^*]\, du \geq \int_0^k G\left(\frac{1}{t}\int_0^t g(u)\, du\right) h(t)\, dt.$$

Provided that and Lemma 5.7 we have a complete proof of the theorem.

So let $a \in (0, 1)$, from Lemma 5.5 we obtain for every $I \in \mathcal{T}$a family of disjoint subsets of I, I_a such that

$$\sum_{J\in I_a} \mu(J) = (1 - a)\mu(I).$$

Now let S_a be the smallest subset of $\mathcal{T}$ such that $X \in S_a$ and for every $I \in S_a$, $I_a \subset S_a$.

Indeed for this construction we just need to consider the intersection of all the families that have the above property.

Next for each $I \in S_a$ let $A_I := I \setminus \cup_{J\in I_a} J$. Then from the choice of I_a we have that $\mu(A_i) = a\mu(I)$. Also we have that S_a can be written as follows

$$S_a = \bigcup_{m=0}^{\infty} S_a^m$$

where $S_a^0 = \{X\}$ and $S_a^{m+1} = \bigcup_{I\in S_a^m} I_a$.

To prove that notice first that S_a must contain each of the sets S_a^m and for the other direction their union of S_a^m is a family which shares the same property as S_a. So since S_a is the least which has this property we obtain the desired equality.

Now for every $I \in S_a$ we define $r(I)$ to be the unique integer m such that $I \in S_a^m$ and for this m we define

$$\gamma(I) := \frac{1}{a(1-a)^m} \int_{(1-a)^{m+1}}^{(1-a)^m} g(u)\, du$$

and

$$b_m(I) := \sum_{S_a \ni J \subset I : r(J) = r(I)+m} \mu(J).$$

Then inductively we have that

$$b_m(I) := (1-a)^m \mu(I).$$

For $m = 1$ it's obvious from the definition of S_a. As for $m > 1$ we just need to notice that for $m \geq 1$ and $I \in S_a$

$$b_{m+1}(I) = \sum_{S_a \ni J \subset I : r(J) = r(I)+m+1} \mu(J) = \sum_{S_a \ni J \subset I : r(J) = r(I)+m} (1-a)\mu(J) = (1-a) b_m(I).$$

We also have that for every $I \in S_a$

$$I \approx \bigcup_{S_a \ni J \subset I} A_J.$$ [2]

Finally for every m we define the measurable subset of X, $S_m^a := \bigcup_{I \in S_a^m} I$.

Now since

$$\begin{aligned}
\mu(S_m^a \backslash S_{m+1}^a) &= \mu(S_m^a) - \mu(S_{m+1}^a) \\
&= b_m(X) - b_{m+1}(X) \\
&= (1-a)^m - (1-a)^{m+1} \\
&= a(1-a)^m
\end{aligned}$$

and X is nonatomic there exists a measurable function $t_m^a : S_m^a \backslash S_{m+1}^a \to [0, \infty)$ such that

$$[t_m^a]^* = \left(g/[(1-a)^{m+1}, (1-a)^m\right)^*.$$

Now define $t_a : X \to [0, \infty)$ as follows

$$t_a := \sum_{m=0}^{\infty} t_a^m \chi_{S_m^a \backslash S_{m+1}^a}$$

[2] here we mean that their symmetric difference has measure zero.

Then we do have that

$$t_a^* = g$$

and since t_m^a is equimeasurable with $g|_{((1-a)^{m+1},(1-a)^m)}$ we have that

$$\int_{S_m^a \setminus S_{m+1}^a} t_m^a \, d\mu = \int_{(1-a)^{m+1}}^{(1-a)^m} g(u)\, du$$

so

$$\frac{1}{\mu(S_m^a \setminus S_{m+1}^a)} \int_{S_m^a \setminus S_{m+1}^a} t_a \, d\mu = \gamma_m.$$

Using Lemma 5.6 and the fact that $S_m^a \setminus S_{m+1}^a \approx \cup_{I \in S_a^m} A_I$ we can find a rearrangement of t_a^m, ϕ_a^m such that

$$\frac{1}{\mu(A_i} \int_{A_I} \phi_a^m \, d\mu = \gamma_m, \quad \text{for every } I \in S_m^a.$$

Now define ϕ_a by

$$\phi_a := \sum_{i=0}^{\infty} \phi_a^m \chi_{S_m^a \setminus S_{m+1}^a}.$$

Then

$$\phi_a^* = g$$

and for $I \in S_a^m$ we have that

$$\begin{aligned}
\frac{1}{\mu(I)} \int_I \phi_a \, d\mu &= \frac{1}{\mu(I)} \sum_{S_a \ni J \subset I} \int_{A_J} \phi_a \, d\mu = \frac{1}{\mu(I)} \sum_{l \geq 0} \sum_{S \ni J \subset I : r(J) = r(I) + l} \int_{A_J} \phi_a \, d\mu \\
&= \frac{1}{\mu(I)} \sum_{l \geq 0} \sum_{S \ni J \subset I : r(J) = r(I) + l} \gamma_{m+l} a_J \\
&= \frac{1}{\mu(I)} \sum_{l \geq 0} \sum_{S \ni J \subset I : r(J) = r(I) + l} a\mu(J) \frac{1}{a(1-a)^{m+l}} \int_{(1-a)^{m+l+1}}^{(1-a)^{m+l}} g(u)\, du \\
&= \frac{1}{\mu(I)} \sum_{l \geq 0} \frac{1}{(1-a)^{m+l}} \int_{(1-a)^{m+l+1}}^{(1-a)^{m+l}} g(u)\, du \sum_{S \ni J \subset I : r(J) = r(I) + l} \mu(J) \\
&= \frac{1}{\mu(I)} \sum_{l \geq 0} \frac{1}{(1-a)^{m+l}} \int_{(1-a)^{m+l+1}}^{(1-a)^{m+l}} g(u)\, du \cdot b_l(I) \\
&= \frac{1}{(1-a)^m} \sum_{l \geq 0} \int_{(1-a)^{m+l+1}}^{(1-a)^{m+l}} g(u)\, du = \frac{1}{(1-a)^m} \int_0^{(1-a)^m} g(u)\, du
\end{aligned}$$

So if $x \in S_m^a \setminus S_{m+1}^a$ there exists $I \in S_a^m$ such that $x \in I$ and so

$$\mathcal{M}_{\mathcal{T}}(\phi_a)(x) \geq \frac{1}{\mu(I)} \int_I \phi_a \, d\mu = \frac{1}{(1-a)^m} \int_0^{(1-a)^m} g(u)\, du =: \theta_m.$$

Also since $\mu(S_m^a) = (1-a)^m$, for $m \geq 0$ we obtain that

$$(\mathcal{M}_\mathcal{T}\phi_a)^*(t) \geq \theta_m, \quad \text{if } t < (1-a)^m.$$

Now for every $a \in (0,1)$ we select m_a such that $(1-a)^{m_a+1} \leq k < (1-a)^{m_a}$ and so we have that $\lim_{a\to 0^+}(1-a)^{m_a} = k$.

Now let us set $\nu(A) = \int_A h(t)\,dt$ for every $A \subset (0,1]$ and first consider the case where $\limsup_{a\to 0^+} \int_0^k G[(\mathcal{M}_\mathcal{T}\phi_a)^*]d\nu(t)$ is infinite the proof is complete. As for the other case, i.e. if

$$\limsup_{a\to 0^+} \int_0^k G[(\mathcal{M}_\mathcal{T}\phi_a)^*]\,d\nu(t) < \infty$$

we have that

$$\int_0^{(1-a)^{m_a}} G[\mathcal{M}_{(\mathcal{T}}\phi_a)]\,d\nu \geq \sum_{l\geq 0} \int_{(1-a)^{m_a+1+l}}^{(1-a)^{m_a+l}} G(\theta_m)\,d\nu$$

$$= \sum_{l\geq 0} G\left(\frac{1}{(1-a)^{m_a+l}} \int_0^{(1-a)^{m_a+l}} g(u)\,du\right) v\left([(1-a)^{m_a+l+1}, (1-a)^{m_a+l}]\right) \quad (5.2)$$

Consequently

$$\limsup_{a\to 0^+} \int_0^{(1-a)^{m_a}} G[(\mathcal{M}_\mathcal{T}\phi_a)^*]\,d\nu \geq \int_0^k G\left(\frac{1}{t}\int_0^t g(u)\,du\right)\,d\nu(t).$$

Also

$$\limsup_{a\to 0^+} G[(\mathcal{M}_\mathcal{T}\phi_a)^*(k)] < \infty$$

since otherwise

$$\limsup_{a\to 0^+} \int_0^k G[(\mathcal{M}_\mathcal{T}\phi_a)^*]\,d\nu(t) = \infty$$

which is a contradiction.

So the right hand side of the inequality

$$\int_k^{(1-a)^{m_a}} G[(\mathcal{M}_\mathcal{T}\phi_a)]^* \leq \left(\int_k^{(1-a)^{m_a}} h(u)\,du\right) G[(\mathcal{M}_\mathcal{T}\phi_a)^*(k)] \quad (5.3)$$

tends to zero as $a \to 0^+$ and so

$$\lim_{a\to 0^+} \int_k^{(1-a)^{m_a}} G[(\mathcal{M}_\mathcal{T}\phi_a)^*(t)]\,d\nu(t) = 0.$$

With that the proof is complete. □

Generalization of the theorem

With analogous reasoning we can prove the following.

Theorem 5.8. *Let $g : (0,1] \to [0,\infty)$ be a right continuous non-increasing function, G_1,G_2 be non-decreasing and non-negative functions defined on $[0,\infty)$ and $k \in (0,1]$. Then the following equality holds true*

$$\sup\left\{\int_K G_1(\mathcal{M}_\mathcal{T}\phi)G_2(\phi)\,d\mu : \phi^* = g \text{ and } \mu(K) = k\right\} = \int_0^k G_1\left(\frac{1}{t}\int_0^t g\right) G_2(g(t))\,dt.$$

5.2 The Bellman function of the Dyadic Maximal Operator

In this section we are going to concern ourselves with finding the exact value of the Bellman function associated with the Dyadic Maximal Operator which is defined as

$$B_p^\mathcal{T}(f,F,L) = \sup\left\{\int_X \{\mathcal{M}_\mathcal{T}\phi, L\}^p : \phi \geq 0, \int_X \phi = f, \int_X \phi^p = F\right\}$$

for all f, F, L such that $0 \leq f^p \leq F$ and $L \geq f$. The papers where one can find the following results are [26] and [34].

Let us first consider for every $p > 1$ the function

$$H_p(z) := -(p-1)z^p + pz^{p-1}$$

defined on $[1, p/(p-1)]$. Now since

$$H_p^{'}(z) = p(p-1)(z^{p-2} - z^{p-1})$$

we obtain that H^p is strictly decreasing on the interval $[1, p/(p-1)]$ and it map onto $[0,1]$. By denoting as $\omega_p : [0,1] \to [1, p/(p-1)]$ its inverse function we are going to prove the following theorem.

Theorem 5.9 (Bellman function of the Dyadic Maximal Operator)**.** *Let $(X,\mu,\mathcal{T})$ be a non-atomic measure space with a tree structure and $p > 1$, then the corresponding Bellman function is given by,*

$$B_p^\mathcal{T}(F,f,L) = \begin{cases} F\omega_p\left(\frac{pL^{p-1}f-(p-1)L^p}{F}\right)^p, & \text{if } f \leq L < \frac{p}{p-1}f \\ L^p + \left(\frac{p}{p-1}\right)^p (F - f^p), & \text{if } L \geq \frac{p}{p-1}f. \end{cases}$$

In order to continue let us first introduce for a non-increasing function $g : [0,1] \to [0,\infty]$ the following

$$v_g(L) := \int_0^1 \max\left(\frac{1}{t}\int_0^t g, L\right)^p dt;$$
$$u_g(L) := \int_0^1 g(t) \max\left(\frac{1}{t}\int_0^t g, L\right)^{p-1} dt$$

for every $L \geq f$.

Lemma 5.10. *The following identity holds between u and v*

$$v_g(L) = L^p - \frac{p}{p-1} f L^{p-1} + \frac{p}{p-1} u_g(L), \quad \text{for every } L \geq f.$$

Proof. First let us split v_g as follows

$$\begin{aligned}
v_g(L) &= \int_0^\infty p\lambda^{p-1} \left| \left\{ t \in (0,1] : \max\left(\frac{1}{t} \int_0^t g, L \right) > \lambda \right\} \right| d\lambda \\
&= \int_0^L p\lambda^{p-1} \left| \left\{ t \in (0,1] : \max\left(\frac{1}{t} \int_0^t g, L \right) > \lambda \right\} \right| d\lambda \\
&+ \int_L^\infty p\lambda^{p-1} \left| \left\{ t \in (0,1] : \max\left(\frac{1}{t} \int_0^t g, L \right) > \lambda \right\} \right| d\lambda \\
&= L^p + \int_L^\infty p\lambda^{p-1} \left| \left\{ t \in (0,1] : \max\left(\frac{1}{t} \int_0^t g, L \right) > \lambda \right\} \right| d\lambda.
\end{aligned}$$

Now for each $\lambda > L$ we consider the unique $\beta(\lambda) \in (0,1]$ such that

$$\frac{1}{\beta(\lambda)} \int_0^{\beta(\lambda)} g(u)\, du = \lambda.$$

Note that we can assume that $g(0+) = \infty$. In the opposite case the proof is analogous. Then we have that

$$v_g(L) = L^p + \int_L^\infty p\lambda^{p-1} |A_\lambda|\, d\lambda$$

where

$$A_\lambda = \left\{ t \in [0,1] : \frac{1}{t} \int_0^t g > \lambda \right\} = (0, \beta(\lambda)).$$

So

$$\begin{aligned}
v_g(\lambda) &= L^p + \int_L^\infty p\lambda^{p-1} \beta(\lambda)\, d\lambda \\
&= L^p + \int_L^\infty p\lambda^{p-1} \left(\frac{1}{\lambda} \int_0^{\beta(\lambda)} g(u)\, du \right) d\lambda \\
&= L^p + \int_L^\infty p\lambda^{p-2} \left(\int_{\left\{u : \frac{1}{u} \int_0^u g > \lambda\right\}} g(u)\, du \right) d\lambda \\
&= L^p + \int_L^\infty p\lambda^{p-2} \left(\int_{\left\{u : \max\left(\frac{1}{u} \int_0^u g, L\right) > \lambda\right\}} g(u)\, du \right) d\lambda \\
&= L^p + \int_0^1 g(u) \frac{p}{p-1} [\lambda^{p-1}]_L^{\max\left(\frac{1}{u}\int_0^u g, L\right)}\, du \\
&= L^p - \frac{p}{p-1} L^{p-1} f + \frac{p}{p-1} u_g(L).
\end{aligned}$$

□

Lemma 5.11. *For every f and F such that $0 < f^p < F$ and $L \geq f$ we have that*

$$\int_X \max(\mathcal{M}_\mathcal{T}\phi, L)^p \leq \begin{cases} F\omega_p\left(\frac{pL^{p-1}f-(p-1)L^p}{F}\right)^p, & \text{if } L < \frac{p}{p-1}f \\ L^p + \left(\frac{p}{p-1}\right)^p (F - f^p), & \text{if } L \geq \frac{p}{p-1}f. \end{cases}$$

Proof. Let us first denote I to be the integral to the left in the above inequality. Then

$$\begin{aligned} I &= \int_0^\infty p\lambda^{p-1}\mu(\{\max(\mathcal{M}_\mathcal{T}\phi, L) > \lambda\})\,d\lambda \\ &= \int_0^L + \int_L^\infty p\lambda^{p-1}\mu(\{\max(\mathcal{M}_\mathcal{T}\phi, L) > \lambda\})\,d\lambda := II + III. \end{aligned}$$

For II we have that

$$II = L^p.$$

As for III, by using the weak-L^1 inequality of the dyadic maximal operator we obtain

$$\begin{aligned} III &= \int_L^\infty p\lambda^{p-1}\mu(\{\mathcal{M}_\mathcal{T}\phi > \lambda\})\,d\lambda \\ &\leq \int_L^\infty p\lambda^{p-1}\left(\frac{1}{\lambda}\int_{\{\mathcal{M}_\mathcal{T}\phi > \lambda\}} \phi\,d\mu\right)d\lambda \\ &= \int_L^\infty p\lambda^{p-2}\left(\int_{\{\max(\mathcal{M}_\mathcal{T}\phi,L) > \lambda\}} \phi\,d\mu\right)d\lambda \\ &= \int_X \phi(x)\left(\int_L^{\max(\mathcal{M}_\mathcal{T}\phi,L)} p\lambda^{p-2}d\lambda\right)d\mu(x) \\ &= \int_X \phi(x)\left[\frac{p}{p-1}\lambda^{p-1}\right]_L^{\max(\mathcal{M}_\mathcal{T}\phi(x),L)} d\mu(x) \\ &= \frac{p}{p-1}\int_X \phi(x)\max(\mathcal{M}_\mathcal{T}\phi(x), L)^{p-1}\,d\mu(x) - \frac{p}{p-1}L^{p-1}f. \end{aligned}$$

Next applying Hölder's Inequality gives us

$$\begin{aligned} III &\leq \frac{p}{p-1}\left(\int_X \phi^p\,d\mu\right)^{\frac{1}{p}} \cdot \left(\int_X \max(\mathcal{M}_\mathcal{T}\phi, L)^p\right)^{\frac{p-1}{p}} - \frac{p}{p-1}L^{p-1}f \\ &= \frac{p}{p-1}F^{\frac{1}{p}}I^{\frac{p-1}{p}} - \frac{p}{p-1}L^{p-1}f. \end{aligned}$$

As a consequence

$$I \leq \frac{p}{p-1}F^{\frac{1}{p}}I^{\frac{p-1}{p}} + L^p - \frac{p}{p-1}L^{p-1}f.$$

Equivalently setting $w := \left(\frac{I}{F}\right)^{\frac{1}{p}}$ gives us

$$pw^{p-1} - (p-1)w^p \geq \frac{pL^{p-1}f - (p-1)L^p}{F} =: b.$$

Now we study right hand side of the last mentioned inequality.

If we consider the function $h : [f, \infty] \to \mathbb{R}$ defined as

$$h(t) := pt^{p-1}f - (p-1)t^p,$$

then h is strictly decreasing thus the right hand side is less than $f^p/F \leq 1$.

Now we will examine separately the cases where $b \geq 0$ and $b < 0$.

1. First we examine the case where $b \geq 0$, which as result gives us that $b \in [0, 1]$ since $L \leq p/(p-1)F$. If $w \leq 1$ then we have that $\frac{I}{F} \leq 1$ and since $\omega_p(b) \geq 1$ we have that $I \leq F[\omega_p(b)]^p$. If $w > 1$ then since H_p is strictly decreasing we have by the last inequality that we reached

$$\begin{aligned} H_p(w) \geq b &\implies w \leq \omega_p(b) \\ &\implies \frac{I}{F} \leq [\omega_p(w)]^p \\ &\implies F \leq F\omega_p\left(\frac{pL^{p-1}f - (p-1)L^p}{F}\right)^p. \end{aligned}$$

2. For the remaining case where $b < 0$ or equivalently $L > \frac{p}{p-1}f := L_0$ as we have already seen

$$I = \int_X \max(\mathcal{M}_\mathcal{T}\phi, L)^p \, d\mu = L^p + III.$$

Since

$$III \leq \int_L^\infty p\lambda^{p-2}\left(\int_{\{\mathcal{M}_\mathcal{T}\phi > \lambda\}} \phi \, d\mu\right) d\lambda$$

and because of the fact that $L > L_0$, we have

$$III \leq \int_{L_0}^\infty p\lambda^{p-2}\left(\int_{\{\mathcal{M}_\mathcal{T}\phi > \lambda\}} \phi \, d\mu\right) d\,\lambda = \int_X \max(\mathcal{M}_\mathcal{T}\phi, L_0)^p \, d\mu - L_0^p.$$

And since we have already established the case for L_0 we have the following

$$\begin{aligned} \int_X (\max \mathcal{M}_\mathcal{T}\phi, L_0)^p \, d\mu &\leq F\omega_p\left(\frac{pL_0^{p-1}f - (p-1)L_0^p}{F}\right)^p \\ &= F[\omega_p(0)]^p = F\left(\frac{p}{p-1}\right)^p. \end{aligned}$$

Thus we reach to the following inequality

$$I \leq L^p + F\left(\frac{p}{p-1}\right)^p - L_0^p = L^p\left(\frac{p}{p-1}\right)^p (F - f^p). \qquad \square$$

So now in order to prove the theorem we just need to prove that the inequality appearing in Lemma 5.11 is sharp.

Theorem 5.12. *The inequality*

$$\int_X \max(\mathcal{M}_\mathcal{T}\phi, L)^p \leq \begin{cases} F\omega_p\left(\frac{pL^{p-1}f-(p-1)L^p}{F}\right)^p, & \text{if } L < \frac{p}{p-1}f \\ L^p + \left(\frac{p}{p-1}\right)^p (F - f^p), & \text{if } L \geq \frac{p}{p-1}f \end{cases}$$

is sharp.

In order to prove that we just have to approximate simultaneously equalities in the in-between inequalities that we came upon during the proof which we list down below. For now let us assume that $L < \frac{p}{p-1}f$.

The first inequality we came upon is the following

$$\int_0^1 \max(\mathcal{M}_\mathcal{T}\phi, L)^p \, d\mu \leq L^p - \frac{p}{p-1}L^{p-1}f + \frac{p}{p-1}\int_X \phi \max(\mathcal{M}_\mathcal{T}\phi, L)^{p-1} \, d\mu \quad (5.4)$$

and the second is the following Hölder's inequality

$$\int_X \phi(x) \max(\mathcal{M}_\mathcal{T}\phi(x), L)^{p-1} \, d\mu(x) \leq \left(\int_X \phi^p \, d\mu\right)^{\frac{1}{p}} \cdot \left(\int_X \max(\mathcal{M}_\mathcal{T}\phi, L)^p\right)^{\frac{p-1}{p}}. \quad (5.5)$$

For the first inequality using Theorem 5.8 for $k = 1$ and

$$G_1(t) := \max(t, L)^p \quad \text{and} \quad G_2(t) := 1$$

we obtain that

$$v_g(L) = \sup_{\phi^*=g} \int_X \max(\mathcal{M}_\mathcal{T}\phi, L)^p \, d\mu$$

and for

$$G_1(t) := \max(t, L)^{p-1} \quad \text{and} \quad G_2(t) := t$$

we obtain

$$u_g(L) = \int_X \phi \max(\mathcal{M}_\mathcal{T}\phi, L)^{p-1} \, d\mu$$

where u and v are defined as in Lemma 5.10. That lemma also guarantees us that this inequality is sharp.

As for the second inequality since it is a Hölder's inequality in order for it to be in fact an equality, g must be so that

$$\max\left(\frac{1}{t}\int_0^t g, L\right) = cg(t)$$

where $c := \omega_p\left(\frac{pL^{p-1}f-(p-1)L^p}{F}\right)$.

In other words we prove the following.

Lemma 5.13. *There exists a continuous function $g : (0,1] \to [0,\infty)$ such that*

$$\int_0^1 g\,dt = f, \qquad \int_0^1 g^p\,dt = F$$

and

$$\max\left(\frac{1}{t}\int_0^t g, L\right) = cg(t), \quad \textit{for every } t \in (0,1]$$

where $c := \omega_p\left(\frac{pL^{p-1}f-(p-1)L^p}{F}\right)$.

Proof. Let us first assume that $L < \frac{p}{p-1}f$.

We define g in the following way

$$g(t) = \begin{cases} Kt^{-1+\frac{1}{c}}, & \text{if } t \in [0,\gamma] \\ \frac{L}{c}, & \text{if } t \in [\gamma, 1] \end{cases}$$

where γ and K are constants yet to be defined so that $\frac{1}{\gamma}\int_0^\gamma g(u)du = L$ or equivalently

$$Kc\gamma^{-1+\frac{1}{c}} = L. \tag{5.6}$$

Then g is decreasing, continuous and satisfies the third property.

Now we determine the value of γ. We should have that

$$\begin{aligned}\int_0^1 g^p(u)du = F &\Leftrightarrow \frac{K^p[t^{-p+\frac{p}{c}+1}]_0^\gamma}{-p+\frac{p}{c}+1} + \frac{L^p}{c^p}(1-\gamma) = F \\ &\Leftrightarrow \frac{K^p c^p \gamma^{-p+\frac{p}{c}+1}}{c^p(-p+\frac{p}{c}+1)} + \frac{L^p}{c^p}(1-\gamma) = F \\ &\Leftrightarrow \frac{c^p K^p \gamma^{-p+\frac{p}{c}+1}}{-(p-1)c^p + pc^{p-1}} + \frac{L^p}{c^p}(1-\gamma) = F.\end{aligned}$$

Using the first equality of this proof gives us

$$\frac{L^p\gamma}{-(p-1)c^p + pc^{p-1}} + \frac{L^p}{c^p}(1-\gamma) = F.$$

Also from the definition of c stems that

$$-(p-1)c^p + pc^{p-1} = \frac{pL^{p-1}f-(p-1)L^p}{F} = b,$$

so our equation takes the following form

$$\frac{FL^p\gamma}{pL^{p-1}f-(p-1)L^p} + \frac{L^p}{c^p}(1-\gamma) = F$$

and as a result

$$\gamma = \frac{F - \frac{L^p}{c^p}}{L^p\left(\frac{1}{b} - \frac{1}{c^p}\right)}.$$

Now we show that γ *indeed belongs to* $[0,1]$. First to show that $\gamma \geq 0$. Using the facts that

$$L^p \leq \int_X \max\left(\mathcal{M}_{\mathcal{T}}\phi, L\right)^p \, d\mu$$

and

$$\int_X \max\left(\mathcal{M}_{\mathcal{T}}\phi, L\right)^p \, d\mu \leq [\omega_p(b)]^p = c^p F$$

we obtain

$$F - \frac{L^p}{c^p} \geq 0.$$

Also since c satisfies $-(p-1)c^p + pc^{p-1} = b$ we have

$$p(c^p - c^{p-1}) = c^p - b > 0.$$

As a result

$$\frac{1}{b} - \frac{1}{c^p} > 0.$$

Now let us prove that $\gamma \leq 1$, equivalently

$$F - \frac{L^p}{c^p} \leq \frac{L^p}{b} - \frac{L^p}{c^p}$$

or

$$F \cdot b \leq L^p.$$

Taking into consideration the definition of b the above inequality is equivalent to

$$F \cdot \frac{pL^{p-1}f - (p-1)L^p}{F} \leq L^p$$

which is in turn equivalent to the condition

$$L^{p-1}f \leq L^p$$

which is true by definition.

Now let us prove that g satisfies the other two conditions.

For the first one i.e. $\int_0^1 g\,dt = f$ we have the following

$$\begin{aligned}
\int_0^1 g\,dt &= \int_0^\gamma g\,dt + \int_\gamma^1 g\,dt \\
&= \int_0^\gamma K t^{-1+\frac{1}{c}}\,dt + \frac{L}{c}(1-\gamma) \\
&= Kc\gamma^{\frac{1}{c}} + \frac{L}{c}(1-\gamma) \\
&= L\gamma + \frac{L}{c}(1-\gamma).
\end{aligned}$$

So we need to prove that

$$L\gamma + \frac{L}{c}(1-\gamma) = f$$

or equivalently

$$\gamma = \frac{f - \frac{L}{c}}{L\left(1 - \frac{1}{c}\right)}.$$

Thus we have to prove that

$$\frac{f - \frac{L}{c}}{L\left(1 - \frac{1}{c}\right)} = \frac{F - \frac{L^p}{c^p}}{L^p\left(\frac{1}{b} - \frac{1}{c^p}\right)}$$

or equivalently

$$\frac{fc - L}{c - 1} = \frac{Fc^p - L^p}{L^{p-1}(\frac{c^p}{b} - 1)}.$$

Solving it with respect to b we obtain

$$b = \frac{c^{p-1}(fc - L)L^{p-1}}{F(c^p - c^{p-1}) - L^p + fL^{p-1}}$$

and since

$$c^p - c^{p-1} = \frac{-b + c^p}{p}$$

the equation above takes the following form

$$b = \frac{c^{p-1}(fc - L)L^{p-1}}{\frac{F}{p}(-b + c^p) - L^p + fL^{p-1}}.$$

Now let us simplify the divisor of the quantity mentioned right above.

$$\begin{aligned}
\frac{F}{p}(-b + c^p) - L^p + fL^{p-1} &= \frac{F}{p}\left(\frac{pL^{p-1}f - (p-1)L^p}{F} + c^p\right) - L^p + fL^{p-1} \\
&= -L^{p-1}f + \frac{p-1}{p}L^p + \frac{F}{p}c^p - L^p + fL^{p-1} \\
&= \frac{F}{p}c^p - \frac{L^p}{p} = \frac{Fc^p - L^p}{p}
\end{aligned}$$

and so the equation above takes the following form

$$b = \frac{pc^{p-1}(fc - L)L^{p-1}}{Fc^p - L^p}.$$

Equivalently

$$\begin{aligned}
&\frac{pc^p f}{L} - pc^{p-1} = b\left(\frac{Fc^p}{L^p} - 1\right) \\
\Leftrightarrow &\frac{pc^p f}{L} - b - (p-1)c^p = b\left(F\frac{c^p}{L^p} - 1\right) \\
\Leftrightarrow &\frac{pf}{L} - (p-1) = b\frac{F}{L^p}
\end{aligned}$$

where we used the fact that $-(p-1)c^p - pc^{p-1} = b$ at the second step. This last equation is equivalent to

$$b = \frac{pL^{p-1}f - (p-1)L^p}{F}$$

which is the definition of b.

Now we examine the case where $L \geq \frac{p}{p-1}f$.

What we are going to do is to construct a sequence of $\{g_n\}_{n\in\mathbb{N}}$, so that for every $n \in \mathbb{N}$, $g_n : (0,1] \to [0,\infty)$ is a decreasing and continuous function such that

$$\int_0^1 g_n\,dt = f, \quad \int_0^1 g_n^p\,dt = F$$

and

$$\lim_{n\to\infty} \int_0^1 \max\left(\frac{1}{t}\int_0^t g_n\,du, L\right)^p dt \geq L^p + \left(\frac{p}{p-1}\right)^p (F - f^p).$$

Let us define

$$g_n(t) = \begin{cases} k_n t^{-1+\frac{1}{c_n}}, & \text{if } t \in (0,\gamma_n] \\ \frac{L_n}{c}, & \text{if } t \in [\gamma_n, 1] \end{cases}$$

where L_n is a sequence such that $L_n \to L_0^-$ and γ_n, c_n, b_n, k_n are defined as follows

- $\gamma_n = \frac{F - \frac{L_n^p}{c_n}}{L_n^p\left(\frac{1}{b_n} - \frac{1}{c_n^p}\right)}$
- $b_n = \frac{pL_n^{p-1}f - (p-1)L_n^p}{F}$
- $c_n = \omega_p(b_n)$
- and k_n is such that $k_n c_n \gamma_n^{-1+\frac{1}{c_n}} = L_n$.

Then since $L_n \to L_0^-$, we have that $b_n \to 0$, $c_n \to \frac{p}{p-1}$ and $\gamma_n \to \frac{f - L_0\frac{p-1}{p}}{L_0\left(1 - \frac{p-1}{p}\right)} = 0$. As we have shown in the first case (i.e. when $L < \frac{p}{p-1}f$)

$$\int_0^1 \max\left(\frac{1}{t}\int_0^t g_n, L_n\right)^p dt = [\omega_p(b_n)]^p F \to \left(\frac{p}{p-1}\right)^p F$$

and so

$$\begin{aligned}\int_0^1 \max\left(\frac{1}{t}\int_0^t g_n, L\right)^p dt &= L^p + \int_L^\infty p\lambda^{p-2}\left(\int_{\{u:\frac{1}{u}\int_0^u g_n > \lambda\}} g_n(t)\,dt\right) d\lambda \\ &= L^p + \int_{L_0}^\infty g_n(t)\,dt - \int_{L_0}^L p\lambda^{p-2}\left(\int_{\{u:\frac{1}{u}\int_0^u g_n > \lambda\}} g_n(t)\,dt\right) \\ &= L^p - L_0^p + \int_0^1 \max\left(\frac{1}{t}\int_0^t g_n, L_0\right)^p dt \\ &- \int_{L_0}^L p\lambda^{p-2}\left(\int_{\{u:\frac{1}{u}\int_0^u g_n > \lambda\}} g_n(t)\,dt\right) d\lambda\end{aligned}$$

and since

$$\max\left(\frac{1}{t}\int_0^t g_n, L_n\right) = \omega_p(b_n)g_n(t)$$

we obtain that

$$\int_0^1 \max\left(\frac{1}{t}\int_0^t g_n, L_0\right)^p dt \geq \int_0^1 \max\left(\frac{1}{t}\int_0^t g_n, L_n\right)^p dt$$
$$= [\omega_p(b_n)]^p \int_0^1 g_n^p(u)\, du = F[\omega_p(b_n)]^p$$

where we used the fact that $L_n < L_0$ for every $n > m$ where $m \in \mathbb{N}$. So

$$\lim_{n\to\infty} \int_0^1 \max\left(\frac{1}{t}\int_0^t g_n, L_0\right)^p dt = F\left(\frac{p}{p-1}\right)^p.$$

Finally by setting

$$a_n(L) := \int_{L_0}^{L} p\lambda^{p-2}\left(\int_{\left\{u:\frac{1}{u}\int_0^u g_n>\lambda\right\}} g_n(u)\, du\right) d\lambda$$

we have that

$$\begin{aligned} a_n(L) &\leq \int_{L_0}^{L} p\lambda^{p-2}\left(\int_{\left\{u:\frac{1}{u}\int_0^u g_n>L_0\right\}} g_n\right) d\lambda \\ &= \left(\int_{\left\{u:\frac{1}{u}\int_0^u g_n>L_0\right\}} g_n(u)\, du\right)\int_{L_0}^{L} p\lambda^{p-2}\, d\lambda \\ &= \tau_L \cdot \int_{\left\{u:\frac{1}{u}\int_0^u g_n>L_0\right\}} g_n(u)\, du \end{aligned}$$

and since

$$\left\{t\in(0,1] : \frac{1}{t}\int_0^t g_n \geq L_0\right\} \subset \left\{t\in(0,1] : \frac{1}{t}\int_0^t g_n \geq L_n\right\}$$

we also have that

$$\left|\left\{t\in(0,1] : \frac{1}{t}\int_0^t g_n \geq L_0\right\}\right| \leq \left|\left\{t\in(0,1] : \frac{1}{t}\int_0^t g_n \geq L_n\right\}\right| = \gamma_n$$

where the last equality holds true since γ_n is the unique element such that

$$\frac{1}{\gamma_n}\int_0^{\gamma_n} g_n = L_n.$$

Now since $\gamma_n \to 0$, $a_n(L) \to 0$ we obtain

$$\lim_{n\to\infty}\int_0^1 \max\left(\frac{1}{t}\int_0^t g_n, L\right)^p dt \geq L^p - L_0^p + \left(\frac{p}{p-1}\right)^p F = L^p + \left(\frac{p}{p-1}\right)^p (F - f^p)$$

and the proof is complete. □

Note 5.14. This result was first shown here [26] using a different method.

5.3 The Bellman function of the Dyadic Maximal Operator related to Carleson's Imbedding Theorem

The Bellman function of the Dyadic Maximal Operator related to Carleson's Imbedding Theorem which we are going to concern ourselves with is defined as follows,

$$\mathcal{B}_p^{\mathcal{T}}(f,F,k) = \sup\Bigg\{ \sum_{I\in\mathcal{T}} \lambda_I (<\phi>_I)^p : \phi\geq 0, \int_X \phi = f, \int_X \phi^p = F \text{ and } \lambda_I \geq 0$$
$$\text{such that } \sum_{J\in\mathcal{T}:J\subset I} \lambda_J \leq \mu(I) \text{ for every } I \text{ in } \mathcal{T} \text{ and } \sum_{I\in\mathcal{T}} \lambda_I = k \Bigg\}.$$

I refer to the paper of Nikolidakis [33] as the source of the context for the rest of this section.

Theorem 5.15. *The following identity holds true,*

$$\mathcal{B}_p^{\mathcal{T}}(f,F,k) = \sup\Bigg\{ \left(F - \frac{(f-B)^p}{(1-k)^{p-1}}\right)\omega_p\left(\frac{B^p}{k^{p-1}\left(F-\frac{(f-B)^p}{(1-k)^{p-1}}\right)}\right)^p,$$
$$B\in[0,F] \text{ such that } h_k(B)\leq F\Bigg\}.$$

Remark 5.16. Using this expression of the Bellman function, the author in [26] proved that

$$\mathcal{B}_p^{\mathcal{T}}(f,F,k) = \left[F\omega_{p,k}\left(\frac{f^p}{F}\right)^p - (1-k)f^p\right]\left[\frac{1-(1-k)\omega_{p,k}\left(\frac{f^p}{F}\right)^{p-1}}{k}\right]$$

where $\omega_{p,k}(U)$, $U\in[0,1]$ is defined as the unique positive solution on the interval $[1, 1+k/(p-1)]$ of the equation

$$-(p-1)z^p + (p-1+k)z^{p-1} = U\left[1+(1-k)\left(\frac{p-1}{z}-p\right)\right].$$

Lemma 5.17. *Let $\phi : X \to [0,\infty)$ be a measurable function such that $\int_X \phi d\mu = f$, $\int_X \phi^p d\mu = F$ where $0 < f^p \leq F$. Also let K be a measurable subset of X such that $\mu(K) = k \in (0,1)$. Then the following inequality holds true,*

$$\int_K (\mathcal{M}_{\mathcal{T}}\phi)^p \, d\mu \leq \int_0^k [\phi^*(u)]^p \, du \cdot \omega_p\left(\frac{\left(\int_0^k \phi^*(u)du\right)^p}{k^{p-1}\int_0^k [\phi^*(u)]^p \, du}\right)^p.$$

Proof. First by using Hardy's inequality we obtain that

$$\int_K (\mathcal{M}_{\mathcal{T}}\phi)^p \leq \int_0^k [(\mathcal{M}_{\mathcal{T}}\phi)^*]^p(t)\, dt.$$

For the integral to right above we obtain using Lemma 5.7 that

$$\int_0^k [(\mathcal{M}_{\mathcal{T}}\phi)^*]^p \, dt \le \int_0^k \left(\frac{1}{t} \int_0^t \phi^*(u) \, du \right)^p dt.$$

Now if we set $f_k := \int_0^k \phi^*(u)du$, then

$$\begin{aligned} \int_0^k \left(\frac{1}{t} \int_0^t \phi^*(u) \, du \right)^p dt &= \int_0^\infty p\lambda^{p-1} \left| \left\{ t \in (0,k] : \frac{1}{t} \int_0^t \phi^* \ge \lambda \right\} \right| d\lambda \\ &= \int_0^{f_k} + \int_{f_k}^\infty p\lambda^{p-1} \left| \left\{ t \in (0,k] : \frac{1}{t} \int_0^t \phi^* \ge \lambda \right\} \right| d\lambda. \end{aligned}$$

Now since for every $\lambda \in [0, f_k]$,

$$\frac{1}{t} \int_0^t \phi^* \ge \lambda, \quad \text{for every } t \in (0,k].$$

Thus

$$\int_0^{f_k} p\lambda^{p-1} \left| \left\{ t \in (0,k] : \frac{1}{t} \int_0^t \phi^* \ge \lambda \right\} \right| d\lambda = k(f_k)^p = \frac{1}{k^{p-1}} \left(\int_0^k \phi^* \right)^p.$$

As for the second integral of the above equality, when $\lambda > f_k$ there exists an $a(\lambda)$ such that

$$\frac{1}{a(\lambda)} \int_0^{a(\lambda)} \phi^*(u) \, du = \lambda$$

and so

$$\left\{ t \in (0,k] : \frac{1}{t} \int_0^k \phi^*(u) \, du > \lambda \right\} = (0, a(\lambda)).$$

As a result its measure is equal to $a(\lambda)$ and so,

$$\begin{aligned} \int_{f_k}^\infty p\lambda^{p-1} \left| \left\{ t \in (0,k] : \frac{1}{t} \int_0^t \phi^* \ge \lambda \right\} \right| d\lambda &= \int_{f_k}^\infty p\lambda^{p-1} a(\lambda) \, d\lambda \\ &= \int_{f_k}^\infty p\lambda^{p-1} \frac{1}{\lambda} \left(\int_0^{a(\lambda)} \phi^*(u) \, du \right) d\lambda \\ &= \int_{f_k}^\infty p\lambda^{p-2} \left(\int_{\left\{ t \in (0,k] : \frac{1}{t} \int_0^t \phi^* > \lambda \right\}} \phi^*(u) \, du \right) d\lambda \\ &= \int_0^k \frac{p}{p-1} \phi^*(t) \left[\lambda^{p-1} \right]_{f_k}^{\frac{1}{t} \int_0^t \phi^*} dt \end{aligned}$$

where in the last step we used Fubini's theorem.

So we reached the following equality,

$$\int_0^k \left(\frac{1}{t} \int_0^t \phi^*(u) \, du \right)^p = -\frac{1}{p-1} \frac{1}{k^{p-1}} \left(\int_0^k \phi^* \right)^p + \frac{p}{p-1} \int_0^k \phi^*(t) \left(\frac{1}{t} \int_0^t \phi^* \right)^{p-1} dt.$$

Now applying Holder's inequality to the second integral to the right gives us the following inequality

$$\int_0^k \left(\frac{1}{t}\int_0^t \phi^*\right)^p dt \leq -\frac{1}{p-1}\frac{1}{k^{p-1}}\left(\int_0^k \phi^*\right)^p + \frac{p}{p-1}\left(\int_0^k [\phi^*]^p\right)^{\frac{1}{p}}\left[\int_0^k \left(\frac{1}{t}\int_0^t \phi^*\right)^p dt\right]^{\frac{p-1}{p}}. \quad (5.7)$$

Let us consider now the following quantities

$$J(k) = \int_0^k \left(\frac{1}{t}\int_0^t \phi^*\right)^p dt \quad , \quad A(k) = \int_0^k [\phi^*]^p \quad \text{and} \quad B(k) = \int_0^k \phi^*.$$

Then inequality (5.7) can equivalently be expressed as

$$J(k) \leq -\frac{1}{p-1}\frac{1}{k^{p-1}}[B(k)]^p + \frac{p}{p-1}[A(k)]^{\frac{1}{p}}[J(k)]^{\frac{p-1}{p}}$$

or

$$\frac{J(k)}{A(k)} \leq -\frac{1}{p-1}\left(\frac{[B(k)]^p}{k^{p-1}A(k)}\right) + \frac{p}{p-1}\left[\frac{J(k)}{A(k)}\right]^{\frac{p-1}{p}}.$$

Now by setting $\Lambda(k) := \left[\frac{J(k)}{A(k)}\right]^{\frac{1}{p}}$ we obtain that

$$\Lambda(k)^p \leq -\frac{1}{p-1}\left(\frac{[B(k)]^p}{k^{p-1}A(k)}\right) + \frac{p}{p-1}\Lambda(k)^{p-1}$$

which can equivalently be written as

$$H_p(\Lambda(k)) \geq \frac{\left(\int_0^k \phi^*\right)^p}{k^{p-1}\int_0^k [\phi^*]^p}.$$

Applying now the inverse function of H_p, ω_p we obtain

$$\Lambda(k) \leq \omega_p\left(\frac{\left(\int_0^k \phi\right)^p}{k^{p-1}\int_0^k [\phi^*]^p}\right)$$

and so

$$J(k) \leq \int_0^k [\phi^*]^p \omega_p\left(\frac{\left(\int_0^k \phi^*\right)^p}{k^{p-1}\int_0^k [\phi^*]^p}\right)^p. \qquad \square$$

Proof of theorem. For the quantities A, B defined in the above proof we have the following,

1. $B^p \leq k^{p-1}A$

2. $A \le F$ and $B \le f$
3. $(f-B)^p \le (1-k)^{p-1}(F-A)$.

Consequently

$$\mathcal{B}_p^{\mathcal{T}}(f,F,k) \le \sup\left\{ A\omega_p\left(\frac{B^p}{k^{p-1}A}\right) : A, B \text{ satisfy (1)-(3)} \right\}$$

and so by 3.

$$\mathcal{B}_p^{\mathcal{T}}(f,F,k) \le \sup\left\{ \left(F - \frac{(f-B)^P}{(1-k)^{p-1}}\right)\omega_p\left(\frac{B^p}{k^{p-1}\left(F-\frac{(f-B)^p}{(1-k)^{p-1}}\right)}\right) : \right.$$
$$\left. B \in [0,f] \text{ and } h_k(B) \le F \right\}.$$

As for the other direction let $k \in (0,1]$ and B as above and let

$$A = F - \frac{(f-B)^p}{(1-k)^{p-1}} \tag{5.8}$$

and let us also fix a $\delta > 0$.

Using Lemma 5.5 we elect a family of pairwise disjoint elements of $\mathcal{T}$, $\{I_1, I_2, \dots\}$ such that

$$\sum_{i\in I} \mu(I_i) = k \quad \text{and} \quad \frac{B^p}{k^{p-1}} \le A.$$

Now restricting ourselves for each j to the subtree $\mathcal{T}(I_j)$ of $\mathcal{T}$ defined as

$$\mathcal{T}(I_j) := \{J \subset I_i : J \in \mathcal{T}\}$$

we obtain a function ϕ_j such that

$$\int_{\mathcal{I}_j} \phi_j^p \, d\mu = \frac{A}{k}\mu(\mathcal{I}_i), \quad \int_{\mathcal{I}_i} \phi_j \, d\mu = \frac{B}{k}\mu(\mathcal{I}_i)$$

and

$$\int_{I_i} \left(\mathcal{M}_{\mathcal{T}(I_i)}(\phi_j)\right)^p \, d\mu \ge \delta\frac{A}{k}\omega_p\left(\frac{B^p}{k^{p-1}A}\right)^p \mu(I_i), \tag{5.9}$$

using the value $B(f,F,f)$ of the Bellman function of the Dyadic Maximal Operator which we evaluated in Theorem 5.9. Now we set $K := \cup_j I_j$ and consider a function ψ such that

$$\int_{X\setminus K} \psi \, d\mu = f - B \quad \text{and} \quad \int_{X\setminus K} \psi^p \, d\mu = F - A.$$

Let us note here that the value of A (5.8) gives us that

$$\psi = \frac{f-B}{1-k} = \left(\frac{F-A}{1-k}\right)^{\frac{1}{p}}.$$

Now setting $\phi := \sum_j \phi_j \chi_{I_j} + \psi \chi_{X \setminus K}$ gives us

$$\int_X \phi \, d\mu = f, \quad \int_X \phi^p \, d\mu = F$$

and

$$\begin{aligned}\int_K (\mathcal{M}_{\mathcal{T}}\phi)^p \, d\mu &\geq \delta A \omega_p \left(\frac{B^p}{k^{p-1} A} \right)^p \\ &= \delta \left(F - \frac{(f-B)^p}{(1-k)^{p-1}} \right) \omega_p \left(\frac{B^p}{k^{p-1} \left(F - \frac{(f-B)^p}{(1-k)^{p-1}} \right)} \right).\end{aligned}$$

Since this holds for every $\delta \in (0,1)$ our proof is complete. $\square$

Elementary Measure Theory

Maximal Operators and measure theory are linked in a direct way. The definition of the maximal operators needs the notion of the measure so we are going to list here the basic theory of measures based on the books "Real and Complex Analysis"[35] by Rudin, "Measure Theory"[4] by Bogachev and "Real Analysis: Measures, Integrals and Applications" [22] by Makarov and Podkorytov.

.1 Basic notions and theorems

First we are going to introduce the notion of the σ-algebra which is generally speaking a collection of subsets of an initially defined fixed set that satisfy certain conditions.

Definition .18. Let X be a nonempty set. A *σ-algebra* S of X is a collection of subsets on X such that

1. $X \in S$.
2. If $A \in S$ then $A^c \in S$.
3. If $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of sets of S then $\cup_{n=1}^{\infty} A_n \in S$.

The elements of S are called measurable subsets of X.

Remark .19. Using the axioms we establish the following,

- $\emptyset \in S$.
- if $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of sets of S then $\cap_{n=1}^{\infty} A_n \in S$.

In fact we can exchange the first and third axioms with the two conditions above.

Remark .20. A family S of subsets on a set X which satisfies the first two axioms in addition to the following, if $\{A_n\}_{n=1}^{k}$ is a finite sequence of sets of S then $\cup_{n=1}^{k} A_n \in S$, is simply called an algebra.

Definition .21. Let X be a nonempty set. A *measure* on X is a mapping $\mu : S \to [0, \infty]$[3], where S is a σ-algebra of X, so that if $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint sets of S then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

And now we define what a measure space is.

Definition .22. A *measure space* is an ordered triad (X, S, μ) where X is a non-empty set, S a σ-algebra on X and μ a measure on X defined on S.

Note .23. Due to how cumbersome the above definition is we sometimes might write only "a measure space X" or "a measure space (X, μ)" provided it is clear in these two cases which is the measure and which are the measurable sets.

Now let's define one of the primary notions of measure theory, that of a measurable function.

Definition .24. Let (X, μ) be a measure space. A function $f : X \to Y$ where Y is a topological space, is called measurable if the inverse image of the open subsets of Y are measurable subsets of X (i.e. $f^{-1}(A)$ is measurable for every $A \subset Y$ open).

Theorem .25. *Let (X, μ) be a measure space, Y, Z two topological spaces and $f : X \to Y$ a measurable function. If $\Phi : Y \to Z$ is a continuous function then the composition of f and Φ (i.e. $\Phi \circ f$) is measurable.*

Proof. Let V be a open set of Z. Since Φ is continuous $\Phi^{-1}(V)$ is open and since f is continuous $f^{-1}\left(\Phi^{-1}(V)\right)$ is open or equivalently $(\Phi \circ f)^{-1}(V)$ is open. □

Corollary .26. *Let (X, μ) be a measure space and $f, g : X \to \mathbb{R}$ be measurable functions. Then for every $\lambda \in \mathbb{R}$ the following functions are measurable,*

1. $f + g$.
2. λf.
3. fg.
4. $\frac{f}{g}$ *provided* $g \neq 0$.

Theorem .27. *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions defined on a measure space X taking values on $[-\infty, \infty]$. Then the following functions are measurable,*

$$f = \sup_{n\in\mathbb{N}} f_n \quad \textit{and} \quad \tilde{f} := \limsup_{n\in\mathbb{N}} f_n.$$

The same results hold true with infimum in place of supremum.

[3]Here we define $[0, \infty]$ to be the set $[0, \infty) \cup \{\infty\}$

Proof. For every $a \in \mathbb{R}$ and $n \in \mathbb{N}$ we have that $f_n^{-1}((a,\infty])$ is measurable since f_n is measurable and so $f^{-1}((a,\infty])$ is also measurable since

$$f^{-1}((a,\infty]) = f_n^{-1}((a,\infty]).$$

Now since for each $b \in \mathbb{R}$ we have that

$$[-\infty, b) = \bigcup_{c<b} [-\infty, c] = \bigcup_{c<b} (c,\infty]^c$$

we also have that $f^{-1}([-\infty,b))$ are measurable. Moreover using the fact that for every $a, b \in \mathbb{R}$ with $a < b$ we have that

$$(a,b) = [-\infty, b) \cap (a, \infty]$$

we obtain that $f^{-1}((a,b))$ is measurable.

Since every open set of $[-\infty,\infty]$ is a countable union of sets of the above form we obtain that f is measurable.

As for $\tilde{f}$ we just have to notice that $\tilde{f} = \inf_{n\in\mathbb{N}} \sup_{i\geq n} f_i$. □

Having defined the notion of the measure we now define the notion of the integral with respect to a measure. We are going to define it through the use of the notion of simple functions which we list below.

Definition .28. Let (X,μ) be a measure space. A function $s : X \to [0,\infty)$ is called *simple* if it is of the form

$$s = \sum_{i=1}^{n} a_i \chi_{A_i}$$

where a_i $(i = 1,\ldots,n)$ are non negative numbers and A_i $(i = 1,\ldots,n)$ are measurable subsets of X.

Remark .29. It stems directly from the fact that the A_i above are measurable that simple functions are measurable.

Now we are going to define the notion of the integral through these three steps

1. First we define the integral of a simple function.
2. Then we define the integral of an extended measurable[4] non negative function.
3. And lastly we are going to define the integral of an extended real valued function[5].

[4] the range of f lies $[0,\infty]$

[5] the values of f can be any real number or $-\infty$, ∞ and we denote that set as $[-\infty,\infty]$

Definition .30 (Integral of a simple function)**.** Let (X, μ) be a measure space. The integral of a simple function $s : X \to [0, \infty)$(denoted by $\int_X s\, d\mu$) is defined as follows, if

$$s = \sum_{i=1}^{n} a_i \chi_{A_i}$$

where $a_i \geq 0$ for $i = 1, \ldots, n$ and A_i for $i = 1 \ldots, n$ are measurable, then

$$\int_X s\, d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

Note .31. The value of the integral is independent of the different forms a simple function can have.

Definition .32 (Integral of an extended non negative function)**.** Let (X, μ) be a measure space and $f : X \to [0, \infty]$ be a measurable function on that measure space. We define the integral of f over X as follows

$$\int_X f\, d\mu := \sup \left\{ \int_X s\, d\mu : s \text{ is a simple function such that } 0 \leq s \leq f \right\}.$$

Remark .33. Note that the value of the integral may equal ∞.

Now for the last part towards the definition of the integral we first need to define two more notions, the positive and the negative parts of a measurable function. Without further delay let (X, μ) be a measure space and $f : X \to [-\infty, \infty]$ be a measurable function. We define the functions $f^+, f^- : X \to [0, \infty]$ as follows

$$f^+ := \max\{f, 0\} \quad \text{and} \quad f^- := \max\{-f, 0\}.$$

They are called the positive and negative parts of f respectively.

Now we are ready to define the integral of an extended real valued function. Before we do that let us first establish some basic rules between the real numbers and $\infty, -\infty$.

1. $0 \cdot \infty = 0$.
2. $a \cdot \infty = \infty \cdot a = \infty$ for every positive number a.
3. $a \cdot \infty = \infty \cdot a = -\infty$ for every negative number a.
4. $a + \infty = \infty + a = \infty$ for every $a \in \mathbb{R}$.

We set the same for $-\infty$ as well as the following three conditions

1. $\infty + \infty = \infty$.
2. $-\infty - \infty = -\infty$.
3. $\infty \cdot \infty = -\infty \cdot -\infty = \infty$.

4. $-\infty \cdot \infty = \infty \cdot -\infty = -\infty$.

Definition .34 (Integral of a extended real valued function)**.** Let (X, μ) be a measure space and $f : X \to [-\infty, \infty]$ be a measurable function. We define the integral of X over f (denoted again as $\int_X f \, d\mu$) as follows

$$\int_X f \, d\mu := \int_X f^+ \, d\mu - \int_X f^- \, d\mu$$

provided one of the integral on the right hand side is finite.

Now we are going to list some of the basic theorems of the measure theory that we will use throughout this thesis.

Theorem .35. *Let (X, μ) be a measure space and $f : X \to [0, \infty]$ be a measurable function. Then there is a sequence of simple functions that converges pointwise to f.*

Proof. For every natural number n we define a simple function $s_n : X \to [0, \infty)$ as follows,

$$s_n(x) = \begin{cases} 0, & \text{if } 0 \leq f(x) < \frac{1}{2^n} \\ \frac{1}{2^n}, & \text{if } \frac{1}{2^n} \ \leq f(x) < \frac{2}{2^n} \\ \vdots & \\ \frac{j-1}{2^n}, & \text{if } \frac{j-1}{2^n} \leq f(x) < \frac{j}{2^n} \\ \frac{n-1}{2^n}, & \text{if } \frac{n(2^n-1)}{2^n} \leq f(x) < n \\ n, & \text{if } f(x) \geq n. \end{cases}$$

Then this sequence of functions converges pointwise to f. □

Now we are going to prove three of the most basic theorems of measure theory,

- Monotone Convergence Theorem.
- Fatou's Lemma.
- Dominated Convergence Theorem.

Theorem .36 (Monotone Convergence Theorem)**.** *Let (X, μ) be a measure space and $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of non negative measurable functions such that*

$$f_n \leq f_{n+1}$$

for every $n \in \mathbb{N}$, that converges pointwise to a function f.

Then f is measurable and

$$\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Proof. Let us notice first that the function f above is measurable from . Also since the sequence of integrals $\int_X f_1\,d\mu$, $\int_X f_2\,d\mu$, $\int_X f_3\,d\mu, \ldots$ is increasing there exists a $a \in [0, \infty]$ such that

$$\int_X f_n\,d\mu \to a.$$

Now since $f_n \leq f$ and by extension $\int_X f_n\,d\mu \leq \int_X f\,d\mu$ we have that

$$a \leq \int_X f\,d\mu.$$

As for the reverse inequality let us consider a simple function s such that $0 \leq s \leq f$ and number $c \in (0,1)$. Then for every $n \in \mathbb{N}$ we define the set E_n as follows

$$E_n := \{x : f_n(x) \geq cs(x)\}.$$

These sets are measurable, form an increasing sequence and also their union is the whole measure space itself. In order to see that just consider the cases where $f = 0$ or not. In the first case if we consider a point x we have that $x \in E_1$. In the second case since $cs(x) \leq f(x)$ we have due to the pointwise convergence of $\{f_n\}_{n\in\mathbb{N}}$ to f that there must be some n such that $x \in E_n$.

Now since

$$\int_X f_n\,d\mu \geq \int_{E_n} f\,d\mu \geq c\int_{E_n} s\,d\mu,$$

by letting $n \to \infty$ we have that

$$a \geq c\int_X s\,d\mu$$

and since this inequality holds true for every $c \in (0,1)$ we have that

$$a \geq \int_X s\,d\mu$$

for every simple function s such that $0 \leq s \leq f$. □

Theorem .37 (Fatou's Lemma)**.** *Let (X, μ) be a measure space and $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions. Then*

$$\int_X \liminf_{n\to\infty} f_n\,d\mu \leq \liminf_{n\to\infty} \int_X f_n\,d\mu.$$

Proof. Let us first set $g_n := \inf_{i\geq n} f_i$. Then we do have that $g_n \leq f_n$ and so

$$\int_X g_n\,d\mu \leq \int_X f_n\,d\mu.$$

Also the sequence $\{g_n\}_{n\in\mathbb{N}}$ is increasing and $g_n \to \liminf_{n\to\infty} f_n$ pointwise. So by using the monotone convergence theorem we have that the left side of the inequality above tends to the integral over X of $\liminf_{n\to\infty} f_n$. We have thus proved the inequality. □

Theorem .38 (Dominated Convergence Theorem)**.** *Let (X, μ) be a measure space and and $\{f_n\}_{n\in\mathbb{N}}\}$ be a sequence of measurable functions such that*

$$f_n(x) \to f(x), \quad \text{for a.e. } x \in X$$

and

$$|f_n| \leq |g|, \quad \text{for some measurable function } g$$

for which, $\int_X |g|\, d\mu < \infty$. Then

$$\int_X |f_n - f|\, d\mu \to 0$$

and as consequence

$$\int_X f_n\, d\mu \to \int_X f\, d\mu.$$

Proof. First let us notice that since $|f_n| \leq |g|$ and $g \in L^1(X)$ we have that $f \in L^1(X)$. Now let us apply Fatou's Lemma to $2g - |f_n - f|$. Then we have that

$$\begin{aligned} \int_X 2g\, d\mu &\leq \int_X \liminf_{n\to\infty} (2g - |f_n - f|)\, d\mu \\ &= \int_X 2g\, d\mu + \liminf_{n\to-\infty} - \int_X |f_n - f|\, d\mu \\ &= \int_X 2g\, d\mu - \limsup_{n\to\infty} \int_X |f_n - f|\, d\mu. \end{aligned}$$

Now since $\int_X |g|\, d\mu < \infty$ we have

$$\limsup_{n\to\infty} \int_X |f_n - f|\, d\mu \leq 0. \qquad \square$$

.2 Extension of Measure

Definition .39. Let X be a nonempty set. A family of sets $\mathcal{P}$ is called a *semiring* if

1. $\emptyset \in \mathcal{P}$.
2. If A, B belong to $\mathcal{P}$ then $A \cap B \in \mathcal{P}$.
3. If A, B belong to $\mathcal{P}$ then

$$A \setminus B = \bigcup_{i=1}^{m} Q_i$$

where $m \in \mathbb{N}$ and Q_i are pairwise disjoint members of $\mathcal{P}$.

Note .40. In this and the following section by the notion "measure" we may also mean set functions defined on semirings with values in $[0, \infty]$ that have the same properties as the measures we defined in the first section.

Theorem .41. *Let $\{P_n\}_{n\in\mathbb{N}}$ be a sequence of elements of a semiring $\mathcal{P}$. Then for every $n \in \mathbb{N}$ there exists a finite sequence $\{Q_i\}_{i\in I}$ of pairwise disjoint elements of $\mathcal{P}$ such that*

$$\bigcup_{i=1}^{n} P_i = \bigcup_{i=1}^{m} Q_i.$$

The same result still holds for the union $\cup_{n=1}^{\infty} P_n$.

Lemma .42. *Let P, $P_1, \ldots, P_n$ where $n \in \mathbb{N}$ be a sequence of elements of a semiring $\mathcal{P}$. Then for every $n \in \mathbb{N}$ there exist a finite sequence $\{Q_i\}_{i=1}^{m}$ of pairwise disjoint elements of $\mathcal{P}$ such that*

$$P \setminus \bigcup_{i=1}^{n} P_i = \bigcup_{i=1}^{m} Q_i.$$

Proof. For $n = 1$ the result stems directly from the definition of the semiring. Suppose now that it holds for all natural numbers $1, \ldots, n$ where $n \geq 2$. Then we have that

$$P \setminus \bigcup_{i=1}^{n+1} P_i = \left(P \setminus \bigcup_{i=1}^{n} P_i \right) \setminus P_{n+1} = \left(\bigcup_{i=1}^{m} Q_i \right) \setminus P_{n+1} = \bigcup_{i=1}^{m} (Q_i \setminus P_{n+1})$$

where Q_i are pairwise disjoint.

Now using the third property of the definition of a semiring we obtain for every $j = 1, \ldots, m$ a finite sequence of elements $Q_1^j, \ldots, Q_{k_j}^j$ such that

$$Q_j \setminus P_{n+1} = \bigcup_{i=1}^{k_j} Q_i^j.$$

As a consequence $P \setminus \cup_{i=1}^{n+1} P_i$ can be expressed as follows,

$$P \setminus \bigcup_{i=1}^{n+1} P_i = \bigcup_{j=1}^{m} \bigcup_{i=1}^{k_j} Q_i^j.$$

□

proof of theorem. From lemma above we prove that for for every $n \in \mathbb{N}$,

$$P_n \setminus \bigcup_{i=1}^{n-1} P_i = \bigcup_{j=1}^{k_n} Q_i^n$$

where Q_i^n are pairwise disjoint elements of $\mathcal{P}$. Consequently

$$\bigcup_{i=1}^{n} P_i = \bigcup_{i=1}^{n} \bigcup_{j=1}^{k_i} Q_j^i$$

and

$$\bigcup_{i=1}^{\infty} P_i = \bigcup_{i=1}^{n} \bigcup_{j=1}^{k_i} Q_j^i.$$

□

Definition .43. Let X be a nonempty set. *An outer measure* on X is a function $\tau : \mathcal{P}(X) \to [0, \infty]$ such that

1. $\tau(\emptyset) = 0$.
2. $\tau(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \tau(A_i)$ for every sequence $\{A_i\}_{i=1}^{\infty}$ of subsets of X.

Definition .44. Let τ be an outer measure on a set X. A subset A of X is called *measurable or τ-measurable* if

$$\tau(E) = \tau(E \cap A) + \tau(E \setminus A), \quad \text{for every } E \subset X.$$

The collection of all these sets will be denoted as $\mathcal{A}_\tau$.

Theorem .45. *Let τ be an outer measure on a set X. Then $\mathcal{A}_\tau$ constitutes a σ-algebra of X and τ restricted to $\mathcal{A}_\tau$ is a measure.*

Proof. *The empty set* belongs to $\mathcal{A}_\tau$ for if $E \subset X$ then

$$\tau(E \setminus \emptyset) + \tau(E \cap \emptyset) = \tau(E).$$

If now A is any set of $\mathcal{A}_\tau$ then

$$\tau(E) = \tau(E \cap A) + \tau(E \setminus A) = \tau(E \setminus A^c) + \tau(E \cap A^c)$$

and so A^c belongs to $\mathcal{A}_\tau$.

Now let A, B be members of $\mathcal{A}_\tau$. Then we have

$$\tau(E) = \tau(E \cap A) + \tau(E \setminus A) = \tau(E \cap A) + \tau((E \setminus A) \cap B) + \tau((E \setminus A) \setminus B).$$

Now using the subadditivity of τ we obtain

$$\tau(E \cap (A \cup B)) \leq \tau(E \cap A) + \tau((E \setminus A) \cap B).$$

Consequently

$$\tau(E) \geq \tau(E \cap (A \cup B)) + \tau(E \setminus (A \cup B)).$$

As for the opposite direction we just have to use the subadditivity of τ, so $A \cup B \in \mathcal{A}_\tau$.

Now if A and B are disjoint elements of $\mathcal{A}_\tau$, then

$$\tau(A \cup B) = \tau((A \cup B) \setminus A) + \tau((A \cup B) \cap A) = \tau(A) + \tau(B).$$

Using induction we obtain that for any finite sequence $\{A_i\}_{i=1}^{m}$ of disjoint sets of $\mathcal{A}_\tau$

$$\tau\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} \tau(A_i).$$

Now let us consider a sequence $\{A_i\}_{i=1}^{\infty}$ of pairwise disjoint sets $\mathcal{A}_\tau$ and let $A := \cup_{i=1}^{\infty} A_i$. Then for every $m \in \mathbb{N}$ we have, for every subset E of X the following,

$$\tau(E) = \tau\left(E \cap \bigcup_{i=1}^{m} A_i\right) + \tau\left(E \setminus \bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} \tau(A_i) + \tau\left(E \setminus \bigcup_{i=1}^{m} A_i\right)$$
$$\geq \sum_{i=1}^{m} \tau(A_i) + \tau(E \setminus A).$$

So by letting $i \to \infty$ and using the subadditivity of τ we obtain

$$\tau(E) \geq \sum_{i=1}^{\infty} \tau(E \cap A_i) + \tau(E \setminus A) \geq \tau\left(\bigcup_{i=1}^{\infty} (E \cap A_i)\right) + \tau(E \setminus A)$$
$$= \tau(E \cap A) + \tau(E \setminus A).$$

The opposite direction is an immediate result of the subadditivity of τ, so $A \in \mathcal{A}_\tau$.

As for the case where the A_i are not necessarily pairwise disjoint we just have to express A as

$$A = \bigcup_{i=1}^{\infty} B_i$$

where $B_1 := A_1$ and $B_{i+1} := A_{i+1} \setminus B_i$. □

Theorem .46. *Let* (X, μ) *be a* σ*-finite measure space on a semiring* $\mathcal{P}$ *and let*

$$\mu^*(A) := \inf\left\{\sum_{j=1}^{\infty} \mu(P_j) : A \subset \bigcup_{j=1}^{\infty} P_j,\ P_j \in \mathcal{P},\ \textit{for every } j \in \mathbb{N}\right\}, \quad \textit{for every } A \subset X$$

using the convention $\inf(\emptyset) = +\infty$.

Then μ^* *is an outer measure that coincides with* μ *on* $\mathcal{P}$.

Proof. Let $A \in \mathcal{P}$, then A, $\emptyset$, $\emptyset$,... constitutes a covering of A comprised with elements of $\mathcal{P}$ and so $\mu^*(A) \leq \mu(A)$.

As for the opposite direction direction, if $\{P_i\}_{i=1}^{\infty}$ is a sequence of sets of $\mathcal{P}$ such that

$$A \subset \bigcup_{i=1}^{\infty} P_i$$

then from the subadditivity of the measure we have

$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(P_i).$$

Since this holds for any random sequence $\{P_i\}_{i=1}^{\infty}$ we have

$$\mu(A) \leq \mu^*(A).$$

Now we prove the subadditivity of μ^*. Let A and $\{A_n\}_{n=1}^{\infty}$ be a sequence of subsets of X such that

$$A \subset \bigcup_{i=1}^{\infty} A_i.$$

If $\sum_{i=1}^{\infty} \mu^*(A_i) = \infty$ there is nothing to prove so we may assume that $\sum_{i=1}^{\infty} \mu^*(A_i) < \infty$. Now for a fixed $\epsilon > 0$ there exist families of subsets of X, $\{P_j^i\}_{j=1}^{\infty}$ such that

$$A_i \subset \bigcup_{j=1}^{\infty} P_j^i \quad \text{and} \quad \sum_{j=1}^{\infty} \mu(P_j^i) \leq \mu^*(A_i) + \frac{\epsilon}{2^i}$$

and so

$$A \subset \bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} P_j^i$$

and by that

$$\mu^*(A) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu(P_j^i) \leq \sum_{i=1}^{\infty} \Big(\mu(A_i) + \frac{\epsilon}{2^i}\Big) = \sum_{i=1}^{\infty} \mu(A_i) + \epsilon.$$

Now by letting $\epsilon \to 0$ we obtain the subadditivity of μ^*. □

Theorem .47. *Let X be a set and μ a measure on a semiring $\mathcal{P}$ of X. If μ^* is the associated outer measure of μ and $\mathcal{A}_{\mu^*}$ is the σ-algebra defined above then $\mathcal{P} \subset \mathcal{A}_{\mu^*}$ and μ^* restricted to $\mathcal{A}_{\mu^*}$ constitutes an extension of μ.*

Proof. We have already proved in Theorem .46 that μ^* restricted to $\mathcal{A}_{\mu^*}$ is a measure so it remains to prove the first property. Taking into consideration the subadditivity of μ^* we just have to prove that for every $A \in \mathcal{P}$ and $E \subset X$ that

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \setminus A).$$

Let us first assume that $E \in \mathcal{P}$. From the definition of the semiring $E \setminus P = \cup_{i=1}^{m} Q_i$ where $Q_i \in \mathcal{P}$ and so

$$E = (E \cap P) \cup \bigcup_{i=1}^{m} Q_i.$$

Now using the subadditivity of μ^* we have

$$\begin{aligned}
\mu^*(E) = \mu(E) = \mu(E \cap P) + \sum_{i=1}^{m} \mu(Q_i) &= \mu^*(E \cap P) + \sum_{i=1}^{m} \mu^*(Q_i) \\
&\geq \mu^*(E \cap P) + \mu^*(\bigcup_{i=1}^{m} Q_i) \\
&= \mu^*(E \cap P) + \mu^*(E \setminus P)
\end{aligned}$$

and so we proved it in case $E \in \mathcal{P}$. Now for the more general case let us first assume that $\mu^*(E) < \infty$ for else there is nothing to prove. For a fixed $\epsilon > 0$ there exists a sequence of sets $\{P_i\}_{i=1}^{\infty}$ of $\mathcal{P}$ such that

$$E \subset \bigcup_{i=1}^{\infty} P_i \quad \text{and} \quad \sum_{i=1}^{\infty} \mu(P_i) < \mu^*(E) + \epsilon.$$

From what we previously proved we have

$$\mu(P_i) = \mu^*(P_i) = \mu^*(P_i \cap P) + \mu^*(P_i \setminus P)$$

and so

$$\begin{aligned}\mu^*(E) + \epsilon > \sum_{i=1}^{\infty} \mu(P_i) = \sum_{i=1}^{\infty} \mu^*(P_i) &= \sum_{i=1}^{\infty} (\mu^*(P_i \cap P) + \mu^*(P_i \setminus P)) \\ &\geq \mu^*\left(\left(\bigcup_{i=1}^{\infty} P_i\right) \cap P\right) + \mu^*\left(\left(\bigcup_{i=1}^{\infty} P_i\right) \setminus P\right) \\ &\geq \mu^*(E \cap P) + \mu^*(E \setminus P).\end{aligned}$$

Now by letting $\epsilon \to 0$ we have our result. □

Note .48. The measure μ^* constructed above is called the *Caratheodory extension of* μ.

Theorem .49 (Uniqueness of Extension). *Let μ be the Caratheodory extension of a measure μ_0 defined on a semiring $\mathcal{P}$, $\mathcal{A}_\mu$ the σ-algebra of the measurable sets and ν a measure that extends μ_0 to a σ-algebra $\mathcal{A}'$ that contains $\mathcal{P}$. Then*

1. *$\nu(A) \leq \mu(A)$ for every set $A \in \mathcal{A}_\mu \cap \mathcal{A}'$ and if $\mu(A) < \infty$ then $\nu(A) = \mu(A)$.*
2. *If μ_0 is σ-finite the $\mu = \nu$ on $\mathcal{A}_\mu \cap \mathcal{A}'$.*

Proof. Let $\{P_i\}_{i=1}^{\infty}$ be a covering of A consisting of elements of $\mathcal{P}$. Then

$$\nu(A) \leq \sum_{i=1}^{\infty} \nu(P_i) = \sum_{i=1}^{\infty} \mu_0(P_i).$$

Since this holds for any such covering we obtain that

$$\nu(A) \leq \mu(A).$$

Now let us prove that

$$\nu(P \cap A) = \mu(P \cap A), \quad \text{for every } P \in \mathcal{P}.$$

If $\nu(P \cap A) < \mu(P \cap A)$ then we would have

$$\mu(P) = \nu(P) = \nu(P \cap A) + \nu(P \setminus A) < \mu(P \cap A) + \mu(P \setminus A) = \mu(P).$$

In case $\mu(A) < \infty$ or μ_0 is σ-finite, then A can be covered by pairwise disjoint elements P_i of $\mathcal{P}$ of finite measure and so

$$\nu(A) = \sum_{i=1}^{\infty} \nu(A \cap P_i) = \sum_{i=1}^{\infty} \mu(A \cap P_i) = \mu(A). \qquad \square$$

Definition .50. Let E be a family of sets. A set S is called E_σ if it can be written as,

$$S = \bigcup_{i=1}^{\infty} A_i$$

where $A_i \in E$ for $i \in \mathbb{N}$.

Likewise a set S is called E_δ if it can be expresses as,

$$S = \bigcap_{i=1}^{\infty} A_i$$

where $A_i \in E$ for $i \in \mathbb{N}$.

Theorem .51. *Let μ be the Caratheodory extension of a measure μ_0 defined on a semiring $\mathcal{P}$. If $\mu^*(A) < \infty$ then there exists a $\mathcal{P}_{\sigma\delta}$ set C such that*

$$A \subset C \quad \textit{and} \quad \mu^*(A) = \mu(C).$$

Proof. Using the definition of $\mu*$ we obtain for $n \in \mathbb{N}$ a sequence of sets $\{P_j^n\}_{j\in\mathbb{N}}$ of elements of $\mathcal{P}$ such that

$$\bigcup_{j=1}^{\infty} P_j^n \supset A \quad \text{and} \quad \sum_{j=1}^{\infty} \mu(P_j^n) < \mu(A) + \frac{1}{n}.$$

By setting $C_n := \cup_{j=1}^{\infty} P_j^n$ for every $n \in \mathbb{N}$ we obtain that $A \subset C_n$ and

$$\mu^*(A) \le \mu(C_n) \le \sum_{j=1}^{\infty} \mu(P_j^n) < \mu^*(A) + \frac{1}{n}$$

and we also have that C_n are $\mathcal{P}_\sigma$ sets. As a result of the above, the set $C := \cap_{n=1}^{\infty} C_n$ is a $\mathcal{P}_{\sigma\delta}$ set and also

$$A \subset C \quad \text{and} \quad \mu(C) = \mu^*(A). \qquad \square$$

Corollary .52. *Let (X, μ) be a measure space and A a measurable set such that $\mu(A) < \infty$. Then there exist $B, C \in \mathcal{B}(\mathcal{P})$ such that*

$$B \subset A \subset C \quad \textit{and} \quad \mu(C \setminus B) = 0.$$

Proof. Let C be the set of the proof of the previous Theorem and let $e := C \setminus A$. Using Theorem .51 we obtain a set $\tilde{e} \in \mathcal{B}(\mathcal{P})$ such that $\tilde{e} \supset e$ and $\mu(\tilde{e}) = 0$. By setting $B := C \setminus \tilde{e}$ we obtain that

$$B \subset C \setminus (C \setminus A) = A$$

and also $\mu(B) = \mu(C)$. $\square$

.3 Product Measure Space and Fubini's Theorem

Definition .53. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two σ-finite measure spaces. A set of the form $A \times B$ where $A \in \mathcal{A}$ and $B \in \mathcal{B}$ with $\mu(A) < \infty$ and $\nu(B) < \infty$ is called a *measurable rectangle.*

Theorem .54. *Let (X, μ), (Y, ν) be σ-finite measure spaces and $\mathcal{P}$ be the family of all the measurable rectangles of $X \times Y$. Then $\mathcal{P}$ is a semiring and the function μ_0 defined as*

$$\mu_0(A \times B) = \mu(A)\nu(B), \quad \textit{for every } A \in \mathcal{A} \textit{ and } B \in \mathcal{B}$$

is a σ-finite measure defined on $\mathcal{P}$.

Definition .55. Let $A \subset X \times Y$. We call the sets

$$A_x := \{y \in Y : (x, y) \in A\} \quad \text{and} \quad A_y := \{x \in X : (x, y) \in A\}$$

the cross-section of A of the first and of the second kind respectively.

Definition .56. Let (X, μ), (Y, ν) be σ-finite measure spaces. The product measure space $(X \times Y, \mu \times \nu, \mathcal{A} \times \mathcal{B})$ is the space $X \times Y$ with measure the standard extension of μ_0 where $\mathcal{A} \times \mathcal{B}$ is the σ-algebra where it is defined.

Before we move on with our next theorem we are first need to introduce the notions of the Borel Hull and the monotone class and then prove some theorems regarding them.

Theorem .57. *Let X be a nonempty set and $\mathcal{F} \subset \mathcal{P}(X)$. Then there exists the minimal σ-algebra which contains $\mathcal{F}$.*

Proof. We define Ω as follows,

$$\Omega := \cap\{G : \ G \text{ is a σ-algebra of } X \text{ and } G \supset \mathcal{F}\}.$$

We prove that Ω is a σ-algebra and that it contains $\mathcal{F}$. Having proved them we will have also proved that it is the least σ-algebra which contains $\mathcal{F}$ since it is the intersection of all such σ-algebras.

First since X belongs to every σ-algebra G it also belongs to Ω.

Now let $A \in \Omega$, then $A \in G$ for every σ-algebra G such that $G \supset \mathcal{F}$. So $A^c \in G$ for every σ-algebra G such that $G \supset \mathcal{F}$. Consequently $A^c \in \Omega$.

The last property is proved similarly so we omit it. □

Definition .58. Let X be a nonempty set and $\mathcal{F} \subset \mathcal{P}(X)$. The least σ-algebra which contains $\mathcal{F}$ is called the Borel hull of $\mathcal{F}$ and is denoted as $\mathcal{B}(\mathcal{F})$.

Definition .59. A family $\mathcal{F}$ of sets is called a monotone class if and only if it satisfies the following two conditions.

- If $\{A_n\}_{n\in\mathbb{N}}$ is an increasing sequence of elements of $\mathcal{F}$ (i.e. $A_{n+1} \supset A_n$) then

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}.$$

- If $\{B_n\}_{n\in\mathbb{N}}$ is an decreasing sequence of elements of $\mathcal{F}$ (i.e. $B_{n+1} \subset B_n$) then

$$\bigcap_{n=1}^{\infty} B_n \in \mathcal{F}.$$

Theorem .60. *If an monotone class contains an algebra $\mathcal{A}$ of a set X then it also contains its Borel hull $\mathcal{B}(\mathcal{A})$.*

Proof. We will prove this by proving that the minimal monotone class $\mathcal{M}$ that contains $\mathcal{A}$ also contains $\mathcal{B}(\mathcal{A})$. Such a class exists since the intersection of all such classes is still a monotone class. Since the proof is similar to the one before regarding the minimal σ-algebra we omit it.

We will prove that $\mathcal{M} = \mathcal{B}(\mathcal{A})$. Since every σ-algebra is a monotone class we immediately obtain that $\mathcal{M} \subset \mathcal{B}(\mathcal{A})$. As for the other direction we will prove that $\mathcal{M}$ is a σ-algebra. Given then the minimal condition regarding $\mathcal{B}(\mathcal{A})$ we obtain equality.

First let us prove for every $A \in \mathcal{A}$ and $B \in \mathcal{M}$, $A \cap B \in \mathcal{A}$ and $A \cap B^c \in \mathcal{A}$. For a fixed $A \in \mathcal{A}$ let,

$$\mathcal{M}_{\mathcal{A}} := \{B \in \mathcal{A} : A \cap B \in \mathcal{M} \text{ and } A \cap B^c \in \mathcal{A}\}.$$

We will prove that $\mathcal{M}_{\mathcal{A}}$ is a monotone class.

Let $\{A_n\}_{n\in\mathbb{N}}$ be an increasing sequence of sets of $\mathcal{A}$. Then we have that

$$A \cap A_n \in \mathcal{M} \quad \text{and} \quad A \cap A_n^c \in \mathcal{M}.$$

Since $\mathcal{M}$ is a monotone class we obtain that

$$\bigcup_{n=1}^{\infty} A \cap A_n \in \mathcal{M} \quad \text{and} \quad \bigcup_{n=1}^{\infty} A \cap A_n^c \in \mathcal{M}.$$

The second property is proved similarly so we omit it. Since $\mathcal{M}$ is minimal and $\mathcal{M}_{\mathcal{A}} \subset \mathcal{M}$ we obtain due to the minimality of $\mathcal{M}$ that $\mathcal{M}_{\mathcal{A}} = \mathcal{M}$.

Setting now $A = X$ gives us that if $B \in \mathcal{M}$ then $B^c \in \mathcal{M}$ and with that we proved that $\mathcal{M}$ the second property of a σ-algebra.

Now we prove that for every A, $B \in A$ $\mathcal{M}$, $A \cup B \in \mathcal{M}$. If for a fixed $B \in \mathcal{A}$ we let

$$\mathcal{N}_B := \{E \in \mathcal{M} : B \cap E \in \mathcal{M}\}$$

we obtain in the same manner as above that $\mathcal{N}_{\mathcal{B}}$ is a monotone class and again using then minimal condition of $\mathcal{M}$ we obtain that $\mathcal{N}_B = \mathcal{M}$.

Now if $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of sets of $\mathcal{M}$ then from what we proved just before we obtain that

$$\bigcup_{i=1}^{n} A_i \in \mathcal{M}, \quad \text{for every } n \in \mathbb{N}$$

and since $\mathcal{M}$ is a monotone class

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{n} A_i \in \mathcal{M}.$$

□

Theorem .61. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite complete measure spaces. Then*

1. *$A_x \in \mathcal{B}$ and $A_y \in \mathcal{A}$ for a.e. $x \in X$ and a.e. $y \in Y$ respectively.*
2. *The functions $x \mapsto \nu(A_x)$ and $y \to \mu(A_y)$ are measurable on X and Y respectively.*
3. *$m(A) := \mu \times \nu(A) = \int_X \nu(A_x)\, d\mu(x) = \int_Y \mu(A_y)\, d\nu(y)$.*

Proof. We will prove only the case for x since for y it is proved similarly. *Let us assume for now that both μ and ν are finite.*

We will prove first that the result holds for every set C belonging to the Borel Hull of $\mathcal{P}$, $\mathcal{B}(\mathcal{P})$.

Let $\mathcal{E}$ be the family of all sets C such that

1. $C_x \in \mathcal{B}$ for every $x \in X$.
2. The function $x \mapsto \nu(C_x)$ is measurable.

Since μ and ν are finite, $X \times Y \in \mathcal{E}$. Also if $C \in \mathcal{E}$ then $C^c \in \mathcal{E}$ since

1. $C^c_x = (C_x)^c$
2. $\nu(C^c_x) = \nu(Y) - \nu(C_x)$.

Furthermore $\mathcal{E}$ is a monotone class since if $\{C_n\}_{n\in\mathbb{N}}$ is an increasing sequence of sets of $\mathcal{E}$ then $\cup_{n=1}^{\infty} C_n \in \mathcal{E}$ due to the fact that

$$C_x = \bigcup_{n=1}^{\infty} C_n \quad \text{and} \quad \nu((C_n)_x) \to \nu(C_x).$$

Now since $\mathcal{P} \subset \mathcal{E}$ and $\mathcal{E}$ contains all the finite disjoint unions of its elements we obtain using Theorems .60 and .41 that $\mathcal{E} \supset \mathcal{B}(\mathcal{P})$. To prove the last one consider two disjoint elements A, $B \in \mathcal{E}$, then $(A \cup B)_x = A_x \cup B_x$ and $\nu((A \cup B)_x) = \nu(A_x) + \nu(B_x)$. If we consider the function $x \mapsto \int_X \nu(E_x)\, d\mu(x)$ defined on $\mathcal{B}(\mathcal{P})$ then we have that it is a measure that coincides with μ on $\mathcal{P}$. So by the uniqueness of the measure it holds for all the members of $\mathcal{B}(\mathcal{P})$.

Now we prove that every member of $\mathcal{B}(\mathcal{P})$ satisfies the third property.

Now let us consider a $C \in \mathcal{A} \times \mathcal{B}$ with $m(C) = 0$. From Theorem .51 we obtain a set $\tilde{C} \in \mathcal{B}(\mathcal{P})$ such that $\tilde{C} \supset C$ and $m(\tilde{C}) = 0$. Now using the previous case regarding $\tilde{C}$ we obtain that

$$\int_X \nu(\tilde{C}_x) = m(\tilde{C}) = 0,$$

so $\nu(\tilde{C}_x) = 0$ a.e.. Since $C_x \subset \tilde{C}_x$ and $\nu(C_x) = 0$ we obtain that $\nu(C_x) = 0$ a.e. due to the completeness of the measure.

Finally if $C \in \mathcal{A} \times \mathcal{B}$ then from Corollary .52, C can be expressed as

$$C = \tilde{C} \setminus e$$

where $\tilde{C} \in \mathcal{B}(\mathcal{P})$ and $m(e) = 0$. So $C_x = \tilde{C}_x \setminus e_x$ is measurable for a.e. $x \in X$. Consequently the function $x \mapsto \nu(C_x)$ which is equal to

$$x \mapsto \nu(\tilde{C}_x) - \nu(e_x),$$

which is equal to $\nu(C_x)$ for a.e. $x \in X$. As a result,

$$m(C) = m(\tilde{C}) = \int_X \nu(\tilde{C}_x), d\mu(x) = \int_X \nu(C_x) + \nu(e_x)\, d\mu(x) = \int_X \nu(C_x)\, d\mu(x)$$

where in the last equality we used the case we proved in the previous step for e.

As for the case where μ *or* ν *are infinite*, since they are σ-finite X and Y can be expressed as

$$X = \cup_{n=1}^{\infty} X_n \quad \text{and} \quad Y = \cup_{n=1}^{\infty} Y_n$$

where $\mu(X_n) < \infty$ and $\nu(Y_n) < \infty$ for $n = 1, 2, \dots$.

So now by considering a measurable subset C of $X \times Y$ we have that,

$$\mu(C) = \sum_{k=1}^{\infty} \mu(C \cap X_k) = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \mu(C \cap (X_k \cap Y_l)).$$

Now applying the result we obtained above to every set $C_{k,l} := C \cap (X_k \cap Y_l)$ gives us that

$$\mu(C_{k,l}) = \int_{X_k} \nu(Y_l \cap C_x)\, d\mu(x).$$

Since $C_x = \cup_{k=1}^{\infty} (Y_k \cap C_x)$ we obtain that that

$$\begin{aligned}\int_X \nu(C_x) &= \sum_{k=1}^{\infty} \int_{X_k} \nu(C_x)\, d\mu(x)\\ &= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \int_{X_k} \nu(Y_l \cap C_x)\, d\mu(x) = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \mu(C_{k,l}) = \mu(C).\end{aligned}$$ □

Theorem .62 (Tonneli's Theorem). *Let* (X, μ) , (Y, ν) *be two* σ*-finite measure spaces and* f *an nonnegative measurable function defined on* $X \times Y$*. Then*

1. *the functions*

$$y \mapsto f(x, y) \quad \textit{and} \quad x \mapsto f(x, y)$$

are measurable for μ*-a.e.* $x \in X$ *and for* ν*-a.e.* $y \in Y$*.*

2. *The functions*

$$y \mapsto \int_X f(x,y)\,d\mu(x) \quad \textit{and} \quad x \mapsto \int_Y f(x,y)\,d\nu(y)$$

are measurable and

$$\int_{X\times Y} f\,d\,(\mu\times\nu) = \int_Y \int_X f(x,y)\,d\mu(x)\,d\nu(y) = \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x).$$

Proof. Let us first assume that $f := \chi_A$ where $A \subset X \times Y$ measurable. We will prove only the case for x since the other is proved similarly. For every $x \in X$ and $y \in Y$ we have

$$\begin{aligned} f_x(y) = \chi_A(x,y) &= \begin{cases} 1, & \text{if } (x,y) \in A \\ 0, & \text{if } (x,y) \notin A \end{cases} \\ &= \begin{cases} 1, & \text{if } y \in A_x \\ 0, & \text{if } y \notin A_x \end{cases} \\ &= \chi_{A_x}(y). \end{aligned}$$

From Theorem .61 we have that A_x are measurable for a.e. $x \in X$ and as a result so is f_x. Integrating over Y gives us

$$\int_Y f_x\,d\nu = \nu(A_x).$$

Now from Theorem .61 we have that

$$x \mapsto \int_Y f(x,y)\,d\nu$$

is measurable and also that

$$\int_X \int_Y f_x\,d\nu\,d\mu = \int_X \nu(A_x)\,d\mu(x) = (\mu\times\nu)(A) = \int_{X\times Y} f.$$

If now f is a simple function the result still holds since it is linear combination of characteristic functions of measurable sets. As for the more general let us consider first a increasing sequence $\{f_i\}_{i=1}^{\infty}$ of measurable functions that converges pointwise to f. Then we have that

$$f_x = \lim_{i\to\infty} (f_i)_x$$

and so f_x is measurable. Now since $(f_n)_x$ is decreasing, the monotone convergence theorem gives us

$$\int_Y f_x(y)\,d\nu(y) = \lim_{i\to\infty} \int_Y (f_i)_x(y)\,d\nu(y).$$

Now since $x \mapsto \int_Y (f_n)_x(y)\,d\nu(y)$ is an increasing sequence of functions the monotone convergence theorem gives us

$$\begin{aligned} \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x) &= \lim_{i\to\infty} \int_X \int_Y f_i(x,y)\,d\nu(y)\,d\mu(x) \\ &= \lim_{i\to\infty} \int_{X\times Y} f_i\,d(\mu\times\nu) = \int_{X\times Y} f\,d(\mu\times\nu). \qquad \square \end{aligned}$$

Theorem .63 (Fubini's Theorem)**.** *Let (X,μ) , (Y,ν) be two σ-finite measure spaces and f an integrable function defined on $X\times Y$ with respect to $\mu\times\nu$. Then*

1. *The functions*
$$y\mapsto f(x,y) \quad \text{and} \quad x\mapsto f(x,y)$$
are integrable for μ-a.e. $x\in X$ and for ν-a.e. $y\in Y$.

2. *The functions*
$$y\mapsto \int_X f(x,y)\,d\mu(x) \quad \text{and} \quad x\mapsto \int_Y f(x,y)\,d\nu(y)$$
are integrable and
$$\int_{X\times Y} f\,d\,(\mu\times\nu) = \int_Y\int_X f(x,y)\,d\mu(x)\,d\nu(y) = \int_X\int_Y f(x,y)\,d\nu(y)\,d\mu(x).$$

Proof. Let us assume first that f is nonnegative. Now from Tonelli's theorem we have that
$$\int_{X\times Y} f\,d(\mu\times\nu) = \int_X\left(\int_Y f(x,y)\,d\nu(y)\right)\,d\mu(x)<\infty.$$

Now again from Tonelli's theorem we have that
$$x\mapsto\int_Y f(x,y)\,d\nu(y)$$
is measurable. Combining it with the above inequality gives us that it is in fact integrable. Consequently the above mentioned mapping is finite a.e. which as a result gives us that f_x is integrable on Y for μ-a.e. $x\in X$.

As for the general case we just have to express f as $f^+ - f^-$ and apply the above process to f^+ and f^- respectively. □

Remark .64. It is important to note here that existence of the repeated integrals of f doesn't generally guarantee the integrability of f, one can find examples of functions such that only one of the repeated integrals exists or both of them exist but have different values (check for instance Rudin [35]).

Bibliography

[1] Stefan Banach. “Sur le théorème de M. Vitali”. In: *Fundamenta Mathematicae* 5.1 (1924), pp. 130–136.

[2] Colin Bennett and Robert C. Sharpley. *Interpolation of operators*. Pure and applied mathematics v. 129. Boston: Academic Press, 1988. 469 pp. ISBN: 978-0-12-088730-9.

[3] A. S. Besicovitch. “A general form of the covering principle and relative differentiation of additive functions. II”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 42.1 (Feb. 1946), pp. 1–10.

[4] V. I. Bogachev. *Measure theory*. Berlin ; New York: Springer, 2007. 2 pp. ISBN: 978-3-540-34513-8.

[5] A. P. Calderon and A. Zygmund. “On the existence of certain singular integrals”. In: *Acta Mathematica* 88.0 (1952), pp. 85–139. DOI: 10.1007/bf02392130.

[6] Rene Erlin Castillo and Humberto Rafeiro. *An Introductory Course in Lebesgue Spaces*. OCLC: 1008636921. 2016. ISBN: 978-3-319-30034-4. URL: https://doi.org/10.1007/978-3-319-30034-4.

[7] R. Coifman and C. Fefferman. “Weighted norm inequalities for maximal functions and singular integrals”. In: *Studia Mathematica* 51.3 (1974), pp. 241–250. URL: http://eudml.org/doc/217916.

[8] Emmanuele DiBenedetto. *Real analysis*. Second edition. Birkhäuser advanced texts. New York: Birkhäuser, 2016. 596 pp. ISBN: 978-1-4939-4003-5.

[9] Javier Duoandikoetxea Zuazo. *Fourier analysis*. Graduate studies in mathematics v. 29. Providence, R.I: American Mathematical Society, 2001. 222 pp. ISBN: 978-0-8218-2172-5.

[10] C. Fefferman and E. M. Stein. “Some Maximal Inequalities”. In: *American Journal of Mathematics* 93.1 (Jan. 1971), p. 107. ISSN: 00029327. DOI: 10.2307/2373450. URL: https://www.jstor.org/stable/2373450?origin=crossref.

[11] R. Fefferman. “Strong Differentiation with Respect to Measures”. In: *American Journal of Mathematics* 103.1 (Feb. 1981), p. 33. DOI: 10.2307/2374188.

[12] Zoltan Furedi and Peter A. Loeb. “On the Best Constant for the Besicovitch Covering Theorem”. In: *Proceedings of the American Mathematical Society* 121.4 (1994). Publisher: JSTOR, p. 1063. DOI: 10.2307/2161215. URL: https://doi.org/10.2307%2F2161215.

[13] José García-Cuerva and J.-L. Rubio de Francia. *Weighted norm inequalities and related topics*. North-Holland mathematics studies 116. Amsterdam ; New York : New York, N.Y., U.S.A: North-Holland ; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co, 1985. 604 pp. ISBN: 978-0-444-87804-5.

[14] L. Grafakos and J. Kinnunen. “Sharp inequalities for maximal functions associated with general measures”. In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 128.4 (1998), pp. 717–723. DOI: 10.1017/s0308210500021739.

[15] Loukas Grafakos. *Classical Fourier analysis*. New York: Springer, 2014. ISBN: 978-1-4939-1194-3.

[16] Loukas Grafakos and Stephen Montgomery-Smith. “Best Constants for Uncentred Maximal Functions”. In: *Bulletin of the London Mathematical Society* 29.1 (Jan. 1997), pp. 60–64. DOI: 10.1112/s0024609396002081.

[17] Miguel de Guzmán. *Real variable methods in Fourier analysis*. OCLC: 1162291670. Amsterdam; New York; New York: North-Holland Pub. Co. : sole distributors for the U.S.A. and Canada, Elsevier North-Holland, 1981. ISBN: 9781281797247 9786611797249 9780080871578.

[18] G. H. Hardy and J. E. Littlewood. “A maximal theorem with function-theoretic applications”. In: *Acta Mathematica* 54.0 (1930), pp. 81–116. DOI: 10.1007/bf02547518.

[19] Edwin Hewitt and Karl Stromberg. *Real and Abstract Analysis: a modern treatment of the theory of functions of a real variable*. 1965. ISBN: 978-3-642-88044-5. URL: https://doi.org/10.1007/978-3-642-88044-5.

[20] Sergei V. Hruscev. “A Description of Weights Satisfying the A \$ınfty\$ Condition of Muckenhoupt”. In: *Proceedings of the American Mathematical Society* 90.2 (1984). Publisher: JSTOR, p. 253. DOI: 10.2307/2045350. URL: https://doi.org/10.2307%2F2045350.

[21] Richard A. Hunt. “An extension of the Marcinkiewicz interpolation theorem to Lorentz spaces”. In: *Bulletin of the American Mathematical Society* 70.6 (Nov. 1, 1964), pp. 803–808. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1964-11242-8. URL: http://www.ams.org/journal-getitem?pii=S0002-9904-1964-11242-8.

[22] B. M. Makarov and Anatolii Podkorytov. *Real analysis: measures, integrals and applications*. Universitext. OCLC: ocn852197538. London ; New York: Springer, 2013. 772 pp. ISBN: 978-1-4471-5121-0.

[23] Lech Maligranda. “Marcinkiewicz Interpolation Theorem and Marcinkiewicz Spaces”. In: *Wiadomości Matematyczne* 48.2 (June 11, 2012), p. 157. ISSN: 2543-991X, 2080-5519. DOI: 10.14708/wm.v48i2.328. URL: http://wydawnictwa.ptm.org.pl/index.php/wiadomosci-matematyczne/article/view/328.

[24] J. MARCINKIEWICZ. “Sur I’interpolation d’operations”. In: *C. R. Acad. Sci. Paris.* 208 (1939), pp. 1272–1273.

[25] Antonios Melas. "The best constant for the centered Hardy–Littlewood maximal inequality". In: *Annals of Mathematics* 157.2 (Mar. 1, 2003), pp. 647–688. ISSN: 0003-486X. DOI: 10.4007/annals.2003.157.647. URL: http://annals.math.princeton.edu/2003/157-2/p08.

[26] Antonios D. Melas. "The Bellman functions of dyadic-like maximal operators and related inequalities". In: *Advances in Mathematics* 192.2 (Apr. 2005), pp. 310–340. DOI: 10.1016/j.aim.2004.04.013.

[27] Anthony P. Morse. "Perfect Blankets". In: *Transactions of the American Mathematical Society* 61.3 (1947), pp. 418–442. ISSN: 00029947. URL: www.jstor.org/stable/1990381.

[28] Benjamin Muckenhoupt. "Weighted norm inequalities for the Hardy maximal function". In: *Transactions of the American Mathematical Society* 165 (1972), pp. 207–207. DOI: 10.1090/s0002-9947-1972-0293384-6.

[29] Benjamin Muckenhoupt. "The equivalence of two conditions for weight functions". In: *Studia Mathematica* 49.2 (1974), pp. 101–106. URL: http://eudml.org/doc/217835.

[30] F. Nazarov, S. Treil, and A. Volberg. "Bellman function in stochastic control and harmonic analysis". In: *Systems, Approximation, Singular Integral Operators, and Related Topics*. Birkhäuser Basel, 2001, pp. 393–423. DOI: 10.1007/978-3-0348-8362-7_16.

[31] F. L. Nazarov and S. R. Treil'. "The hunt for a Bellman function: Applications to estimates of singular integral operators and to other classical problems in harmonic analysis." In: *St. Petersburg Mathematical Journal* 8.5 (1996). Publisher: American Mathematical Society (AMS), Providence, RI, pp. 32–162. ISSN: 1061-0022; 1547-7371/e.

[32] Eleftherios Nikolidakis. "The geometry of the dyadic maximal operator". In: *Revista Matemática Iberoamericana* 30.4 (2014), pp. 1397–1411. DOI: 10.4171/rmi/819.

[33] Eleftherios N. Nikolidakis. "The Bellman function of the dyadic maximal operator in connection with the Dyadic Carleson Imbedding Theorem and related inequalities". In: (May 17, 2019). arXiv: 1905.08091v3.

[34] Eleftherios N. Nikolidakis and Antonios D. Melas. "A sharp integral rearrangement inequality for the dyadic maximal operator and applications". In: *Applied and Computational Harmonic Analysis* 38.2 (Mar. 2015), pp. 242–261. DOI: 10.1016/j.acha.2014.03.008.

[35] Walter Rudin. *Real and complex analysis*. 3rd ed. New York: McGraw-Hill, 1987. 416 pp. ISBN: 978-0-07-054234-1.

[36] Barry Simon. *Harmonic analysis*. A comprehensive course in analysis part 3. Providence, Rhode Island: American Mathematical Society, 2015. 759 pp. ISBN: 978-1-4704-1102-2.

[37] John M. Sullivan. "Sphere packings give an explicit bound for the Besicovitch Covering Theorem". In: *Journal of Geometric Analysis* 4.2 (June 1994), pp. 219–231. DOI: 10.1007/bf02921548.

[38] G. Vitali. “Sull’integrazione per serie”. In: *Rendiconti del Circolo Matematico di Palermo* 23.1 (Dec. 1907), pp. 137–155. DOI: 10.1007/bf03013514.

[39] Alexander Volberg. “Bellman function technique in Harmonic Analysis. Lectures of INRIA Summer School in Antibes, June 2011”. In: *arXiv:1106.3899 [math]* (June 20, 2011). arXiv: 1106.3899. URL: http://arxiv.org/abs/1106.3899.

[40] Norbert Wiener. “The ergodic theorem”. In: *Duke Mathematical Journal* 5.1 (Mar. 1939), pp. 1–18. DOI: 10.1215/s0012-7094-39-00501-6.

[41] A. ZYGMUND. “On a theorem of Marcinkiewicz concerning interpolation of operators”. In: *J. de Math.* 35 (1956), pp. 223–248. URL: https://ci.nii.ac.jp/naid/10009823157/en/.

Index